

The journal has had 40 points in Minister of Science and Higher Education of Poland parametric evaluation. Annex to the announcement of the Minister of Education and Science of 05.01.2024 No. 32318. Has a Journal's Unique Identifier: 201159. Scientific disciplines assigned: Physical culture sciences (Field of medical and health sciences); Health Sciences (Field of medical and health sciences).

Punkty Ministerialne 40 punktów. Załącznik do komunikatu Ministra Nauki i Szkolnictwa Wyższego z dnia 05.01.2024 Lp. 32318. Posiada Unikatowy Identyfikator Czasopisma: 201159. Przypisane dyscypliny naukowe: Nauki o kulturze fizycznej (Dziedzina nauk medycznych i nauk o zdrowiu); Nauki o zdrowiu (Dziedzina nauk medycznych i nauk o zdrowiu). © The Authors 2025;

This article is published with open access at Licensee Open Journal Systems of Nicolaus Copernicus University in Torun, Poland

Open Access. This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author (s) and source are credited. This is an open access article licensed under the terms of the Creative Commons Attribution Non commercial license Share alike.

(<http://creativecommons.org/licenses/by-nc-sa/4.0/>) which permits unrestricted, non commercial use, distribution and reproduction in any medium, provided the work is properly cited.

The authors declare that there is no conflict of interests regarding the publication of this paper.

Received: 09.09.2024. Revised: 20.01.2025. Accepted: 21.01.2025. Published: 25.01.2025.

POLYGENIC DISEASES

Authors:

Dolomatov Sergey Igorevich, V.I. Vernadsky Crimean Federal University, PhD in Biology, Associate Professor

Zukow Walery, Nicolaus Copernicus University, Torun, PhD in Medicine, Associate Professor

Zhukova Anna Alexandrovna, V.I. Vernadsky Crimean Federal University, PhD in Biology, Associate Professor

Smirnova Svetlana Nikolaevna, V.I. Vernadsky Crimean Federal University, PhD in Biology, Associate Professor

Abstract

Background: Recent advances in molecular genetics have revolutionized our understanding of hereditary diseases, particularly polygenic (multifactorial) conditions. This comprehensive review examines the genetic basis of common polygenic diseases, including diabetes, preeclampsia, breast cancer, and factors influencing athletic performance.

Material and Methods: Analysis of current literature covering genetic aspects of polygenic diseases, focusing on gene polymorphisms, single nucleotide polymorphisms (SNPs), and their role in disease development. The review encompasses nutrigenetics, sports genomics, and genetic markers associated with athletic performance and injury risk.

Results: Gene polymorphisms significantly impact disease susceptibility and progression. For type 2 diabetes, over 600 candidate genes have been identified. In preeclampsia, multiple genes influence maternal-fetal interactions and vascular function. Breast cancer involves complex interactions between high and moderate penetrance genes, with over 300 SNPs identified. Athletic performance is influenced by both genetic factors (approximately 66%) and environmental conditions, with specific genetic markers associated with endurance, strength, and injury risk.

Conclusions: Polygenic diseases result from complex interactions between multiple genes and environmental factors. Understanding these interactions enables more effective personalized approaches to prevention, diagnosis, and treatment. Integration of genetic testing into clinical practice offers opportunities for improved patient outcomes through targeted interventions and personalized medicine strategies.

Keywords: Polygenic diseases; Gene polymorphism; Single nucleotide polymorphisms; Type 2 diabetes; Preeclampsia; Breast cancer; Athletic genetics; Nutrigenetics; Sports genomics; Personalized medicine; Genetic markers; DNA methylation; Epigenetics; Exercise genetics; Sports injuries

Реферат

Цель исследования: Систематизировать и обобщить современные представления о молекулярно-генетических основах полигенных (мультифакторных) заболеваний человека.

Методы: Проведен аналитический обзор современной научной литературы, посвященной генетическим аспектам мультифакторных заболеваний. Рассмотрены механизмы влияния полиморфизма генов на патогенез наиболее распространенных полигенных заболеваний.

Результаты: Представлены современные данные о роли полиморфизма генов в патогенезе полигенных заболеваний. Показано, что мультифакторные заболевания возникают в результате сложного взаимодействия генетических и средовых факторов. Описаны основные гены предрасположенности к различным мультифакторным заболеваниям, включая преэклампсию, рак молочной железы, сахарный диабет 2-го типа. Рассмотрены вопросы персонализированного питания на основе генетического тестирования. Представлены данные о генетических маркерах спортивных достижений.

Заключение: Понимание генетических основ мультифакторных заболеваний имеет важное значение для разработки методов их ранней диагностики, профилактики и лечения. Использование современных молекулярно-генетических технологий позволяет выявлять гены предрасположенности к различным заболеваниям и разрабатывать персонализированные подходы к их профилактике и лечению.

Ключевые слова: полигенные заболевания, мультифакторные заболевания, гены предрасположенности, однонуклеотидный полиморфизм, преэклампсия, рак молочной железы, сахарный диабет 2-го типа, нутригенетика, спортивная генетика, персонализированная медицина.

Content

CHAPTER 1. ROLE OF HEREDITARY APPARATUS DISORDERS IN THE PATHOGENESIS OF HEREDITARY DISEASES	6
INTRODUCTION	6
MOLECULAR DIAGNOSTICS OF MONOGENIC AND POLYGENIC DISEASES	8
MONOGENIC DISEASES IN HUMANS	9
GENE MODIFIERS OF MONOGENIC DISEASES	10
OLIGOGENIC AND MULTIFACTORIAL DISORDERS	10
OLIGOGENIC ETIOLOGY EXEMPLIFIED BY CARDIOMYOPATHIES	12
MULTIFACTORIAL (POLYGENIC) DISEASES	12
MEDICAL ASPECTS OF GENE MUTATIONS AND CHROMOSOMAL ABERRATIONS EXEMPLIFIED BY THE DIAGNOSIS OF HEREDITARY CAUSES OF INFERTILITY	13
CHROMOSOME STRUCTURE AND NUMBER ABNORMALITIES: A CAUSE OF INFERTILITY	13
MONOGENIC CAUSES OF INFERTILITY	15
MONOGENIC CAUSES OF SPERMATOGENIC INSUFFICIENCY	15
MONOGENIC CAUSES OF SPERM MORPHOLOGY DEFECTS	16
MONOGENIC CAUSES OF SPERM MOTILITY DEFECTS	16
POLYGENIC CAUSES OF MALE INFERTILITY	16
RECOMMENDED LITERATURE FOR CHAPTER 1. ROLE OF GENETIC APPARATUS DISORDERS IN THE PATHOGENESIS OF HEREDITARY DISEASES	17
CHAPTER 2. GENE POLYMORPHISM	19
INTRODUCTION	19
MUTATIONS – THE CAUSE OF GENE POLYMORPHISM	19
SNP, SINGLE NUCLEOTIDE POLYMORPHISM	20
SNPS AND HEREDITARY DISEASES	23
SNP AND DIABETES MELLITUS	24
SNP AND ONCOLOGY	25
VNTR (SSR) - VARIATIONS IN THE NUMBER OF TANDEM REPEATS OF NUCLEOTIDES	25
MUTATIONAL MECHANISMS OF SSR VARIATION	27
SSR AND TRANSCRIPTION	27
SSR AND TRANSLATION	28
MEDICAL SIGNIFICANCE OF SSR	28
RECOMMENDED LITERATURE FOR CHAPTER 2. GENE POLYMORPHISM	28
CHAPTER 3. POLYGENIC (MULTIFACTORIAL) DISEASES. TYPE 2 DIABETES AS AN EXAMPLE OF POLYGENIC DISEASES	30
INTRODUCTION	30
THE ROLE OF GENE NETWORKS IN THE PATHOGENESIS OF MULTIFACTORIAL DISEASES	31

THE ROLE OF MULTIFACTORIAL DISEASES OF THE HEART, KIDNEYS, LUNGS, AND LIVER IN CARCINOGENESIS	33
THE CONNECTION BETWEEN CANCER AND HEART FAILURE	33
THE CONNECTION BETWEEN ONCOLOGY AND CHRONIC KIDNEY FAILURE..	33
DEVELOPMENT OF LUNG CANCER IN CHRONIC OBSTRUCTIVE PULMONARY DISEASE (COPD)	34
PROGRESSION OF METABOLIC DYSFUNCTION-ASSOCIATED FATTY LIVER DISEASE AND ONCOLOGICAL CONDITIONS	35
THE POLYGENIC NATURE OF MULTIFACTORIAL DISEASES: A CASE STUDY OF TYPE 2 DIABETES	35
TYPE 2 DIABETES MELLITUS AS AN EXAMPLE OF A POLYGENIC DISEASE: PREDISPOSITION GENES	36
GENES OF ANTIOXIDANT ENZYMES	40
RECOMMENDED LITERATURE FOR CHAPTER 3. POLYGENIC DISEASES	40
CHAPTER 4. PREECLAMPSIA AS A POLYGENIC (MULTIFACTORIAL) DISEASE	43
INTRODUCTION	43
HEREDITARY NATURE OF THE DISEASE	45
GENES PREDISPOSING TO PREECLAMPSIA	45
GENES REGULATING ENDOTHELIAL FUNCTION AND BLOOD PRESSURE IN THE PATHOGENESIS OF PRE-ECLAMPSIA	46
GENE AGT AND GENE NOS3 ANALYSIS	47
GENES REGULATING LIPID METABOLISM IN THE PATHOGENESIS OF PREECLAMPSIA	48
APOE GENE	49
ADIPOQ GENE	49
GENES OF THE BLOOD COAGULATION SYSTEM REGULATION IN THE PATHOGENESIS OF PE	49
IMMUNE SYSTEM FACTORS IN THE PATHOGENESIS OF PRE-ECLAMPSIA	50
RECOMMENDED LITERATURE FOR CHAPTER 4. PREECLAMPSIA	55
CHAPTER 5. POLYGENIC BASIS OF MULTIFACTORIAL ONCOLOGICAL DISEASES IN HUMANS: THE CASE OF BREAST CANCER	57
INTRODUCTION	57
POLYGENIC NATURE OF ONCOLOGICAL DISEASES: THE ROLE OF GENE POLYMORPHISM IN THE DEVELOPMENT OF CANCER PATHOLOGIES	57
GENETIC RISK FACTORS FOR BREAST CANCER	58
GENETIC MARKERS FOR BREAST CANCER	60
RECOMMENDED LITERATURE FOR CHAPTER 5. BREAST CANCER	64
CHAPTER 6. NUTRITIONAL GENETICS	66
NUTRIGENETICS, NUTRIGENOMICS, AND DIETARY EPIGENETICS	66
GENETIC BASIS OF PERSONALIZED NUTRITION	67
PRACTICAL APPLICATION OF GENETIC MARKERS IN FOOD CONSUMPTION RECOMMENDATIONS	69
GENETIC FOUNDATIONS OF BODY WEIGHT CONTROL	70

INTERACTION OF GENE VARIANTS WITH DIET LEADING TO OBESITY	71
RECOMMENDED LITERATURE FOR CHAPTER 6. GENETIC NUTRITION	71
CHAPTER 7. ATHLETE'S GENETIC MAP	73
INTRODUCTION	73
GENETIC FOUNDATIONS OF SPORTING ACHIEVEMENTS	73
GENES DETERMINING SUCCESS IN SPORTS	74
IDENTIFICATION OF CANDIDATE GENES IN SPORTS GENOMICS	75
GENETICS AND INDIVIDUAL DIFFERENCES IN RESPONSE TO TRAINING	77
SPORTS NUTRIGENOMICS	78
GENETIC VARIANTS RESPONSIBLE FOR ATHLETE ENDURANCE	80
GENE VARIANTS RESPONSIBLE FOR ATHLETE STATUS IN STRENGTH SPORTS	81
GENETICS OF SPORTS INJURIES	82
MITOCHONDRIAL DNA AND ATHLETIC PERFORMANCE	85
EPIGENETICS IN SPORTS MEDICINE	86
RECOMMENDED LITERATURE FOR CHAPTER 7. ATHLETE'S GENETIC MAP	86

Содержание

ПОЛИГЕННЫЕ ЗАБОЛЕВАНИЯ	89
РАЗДЕЛ 1. РОЛЬ НАРУШЕНИЙ НАСЛЕДСТВЕННОГО АППАРАТА КЛЕТКИ В ПАТОГЕНЕЗЕ НАСЛЕДСТВЕННЫХ ЗАБОЛЕВАНИЙ	89
РАЗДЕЛ 2. ПОЛИМОРФИЗМ ГЕНОВ	100
РАЗДЕЛ 3. ПОЛИГЕННЫЕ (МУЛЬТИФАКТОРНЫЕ) ЗАБОЛЕВАНИЯ. САХАРНЫЙ ДИАБЕТ 2-го ТИПА, КАК ПРИМЕР ПОЛИГЕННОГО ЗАБОЛЕВАНИЯ	111
РАЗДЕЛ 4. ПРЕЭКЛАМПСИЯ, КАК ПОЛИГЕННОЕ (МУЛЬТИФАКТОРНОЕ) ЗАБОЛЕВАНИЕ	123
РАЗДЕЛ 5. ПОЛИГЕННЫЕ ОСНОВЫ ОНКОЛОГИЧЕСКИХ МУЛЬТИФАКТОРНЫХ ЗАБОЛЕВАНИЙ ЧЕЛОВЕКА НА ПРИМЕРЕ РАКА МОЛОЧНОЙ ЖЕЛЕЗЫ (РМЖ)	136
РАЗДЕЛ 6. ГЕНЕТИКА ПИТАНИЯ	145
РАЗДЕЛ 7. ГЕНЕТИЧЕСКАЯ КАРТА СПОРТСМЕНА	151

CHAPTER 1. ROLE OF HEREDITARY APPARATUS DISORDERS IN THE PATHOGENESIS OF HEREDITARY DISEASES

INTRODUCTION

In contemporary medical practice, genetic testing methods are increasingly in demand within the framework of predictive medicine. It is important to distinguish between the terms "genetic test" (GT) and "genetic (molecular) assay" (GA). The former refers to already identified and confirmed diagnostic gene markers. In contrast, the results of a GA are prognostic in nature and require further confirmation before they attain the status of a genetic test. Molecular analyses of polymorphisms in individual genes conducted on a sufficiently large sample of patients, supported by a robust control group, even within a single population or regional population, approach the criteria required for a genetic test.

Genetic testing during the presymptomatic period provides the opportunity to identify hereditary tendencies towards the development of future diseases that currently exist only in the genome (not yet phenotypically expressed). Consequently, a genetic test allows, based on current medical experience, the identification of disease monitoring pathways and the development of early prevention methods. Information about various degrees of risk for developing these specified diseases can be obtained through genetic testing. As a result, a physician, taking into account the results of molecular genetic analysis, can develop a strategy for pathogenetically justified anticipatory therapy, thereby making the necessary correction to congenital metabolic defects.

Evidently, these studies result in the accumulation of data on both the genomes of individuals and entire families, gradually forming personal and familial DNA databases. This DNA database constitutes a "genetic passport".

Thus, a genetic passport represents an individual's DNA database, reflecting each person's unique genetic characteristics and predispositions to certain hereditary diseases.

Currently, based on existing data, the testing of many gene markers, whose association with serious diseases is considered proven, is entirely justified.

The information contained in a genetic passport aids in avoiding potential health issues related to ignoring individual genomic features and specific hereditary characteristics. On the other hand, such data enable a person to more fully realize their genetic potential and hold undeniable value for future generations.

According to V. S. Baranov and O. S. Glotov, an individual genetic map (genetic passport) includes the following main sections:

1. Personal passport data (name, year of birth, nationality);
2. The patient's karyotype;
3. A unique genetic identification number obtained through genomic fingerprinting, consisting of a 12-digit number;
4. Results of testing for asymptomatic carrier status (heterozygosity) of mutations for the most common monogenic diseases;
5. Presymptomatic diagnosis of diseases with late onset;
6. Results of testing for gene markers of common multifactorial diseases;
7. Medical-genetic conclusions derived from genetic counseling before and after genetic testing;
8. Recommendations based on the genetic testing results for the individual and the attending physician.

The complete version of a genetic map (genetic passport) should encompass not only the results of research on predisposition genes but also information on the asymptomatic carrier status

of mutations in the genes responsible for the most common hereditary diseases (heterozygosity for genes linked to hemophilia, cystic fibrosis, phenylketonuria, among others). Testing for hereditary predisposition related to one or several specific conditions is practically useful when preliminary medical-genetic counseling indicates potential hereditary risks for the patient.

The genetic map (genetic passport) for reproductive health, developed by the D. O. Ott Research Institute of Obstetrics and Gynecology, Northwest Branch of the Russian Academy of Medical Sciences (St. Petersburg), has proven effective. In addition to analyzing the karyotype and testing for carrier status of severe hereditary diseases (heterozygosity) in couples planning for a child, it holds significant prognostic value to analyze a woman using gene panels for diseases that complicate pregnancy, fetal development, childbirth, and the postpartum period (such as preeclampsia, habitual miscarriage, varicose veins, and feto-placental insufficiency).

Considering the significance of a genetic passport for a child's health, harmonious personality development, rational nutrition, effective education, sports activities, optimal career guidance, and the ability to prevent the development of certain late-onset diseases, conducting such a genetic test at an early age appears to be well justified.

Furthermore, in families at high risk for type 1 and type 2 diabetes, bronchial asthma, sudden death syndrome, conduction and rhythm abnormalities, metabolic syndrome, and obesity, as well as in several other medical conditions, preventive genetic testing at an early age is considered justified. Naturally, such testing can be conducted only with parental consent, at the referral of a pediatrician, and following a consultation with a geneticist. Additionally, information is being accumulated about genetic markers, testing of which can assess a teenager's suitability for various sports.

Exome sequencing technology (WES) enables the study of the structure, genetic polymorphisms, and functions of various gene variants within the genome across specific populations. This technology aids in understanding the hereditary mechanisms of specific monogenic or multifactorial (polygenic) diseases, allowing for accurate diagnosis, improved prevention, and treatment efficacy. This field stimulates the rapid development of molecular predictive medicine and the formation of genetic passports as a step towards a clinical genetic passport, derived from exome sequencing data (WES technology). For instance, in reproductive medicine, particularly in pregnancy planning, a comprehensive approach is required, based on a range of molecular, genetic, cytogenetic, and embryological methods.

Clinical-genetic passports, genetic mapping, and next-generation sequencing are now integral to all areas of medical science, enabling clinical medicine to address significant reproductive challenges. Specifically, these advancements are vital in fields such as non-invasive prenatal testing (NIPT) or the screening of monogenic and oligogenic diseases by detecting pathogenic variants in probands and families at high risk; pregnancy planning through preimplantation genetic diagnosis (PGD) and treatment; and diagnostic confirmation, which is increasingly relevant for assessing the risk of multifactorial and infectious diseases, in addition to identifying phenotypic characteristics in individuals. It is important to note the critical updates in genetic terminology regarding these recent methods. In 2015 and 2017, the American College of Medical Genetics and Genomics (ACMG) and the Russian Society of Medical Genetics, respectively, recommended replacing the terms "mutation" and "polymorphism" with "nucleotide sequence variant," accompanied by the following modifiers: (1) P - pathogenic, (2) LP - likely pathogenic, (3) of uncertain significance, (4) likely benign, or (5) benign.

The WES (Whole Exome Sequencing) technology allows for the rapid and efficient identification of multiple hereditary diseases in any individual. This is particularly relevant in cases of co-inherited X-linked and autosomal dominant disorders such as

ichthyosis, Wilson's disease, and hemochromatosis. By understanding the molecular defects leading to disease development, patients can benefit from more appropriate follow-up care.

The majority of diseases are not monogenic; therefore, before evaluating the degree of disease risk, it is essential to accurately understand the nature of the disease (i.e., monogenic, oligogenic, or multifactorial condition), which is not always straightforward. There is an emerging opportunity to identify individuals at high genetic risk for the most common pathologies (such as diabetes, cardiovascular diseases, oncopathologies, etc.) at any age. The situation with multifactorial diseases is somewhat more complex, as genomic changes influence the etiology of the disease, with a large set of genes predisposing individuals to the condition (the phenomenon of additivity). Moreover, besides the hereditary factor, susceptibility to the disease is shaped by numerous environmental factors, while inheritance cannot be explained solely by Mendel's laws. Indeed, identical diagnoses may be caused by different risk factors and etiologies in different individuals. In these cases, Whole Exome Sequencing (WES) technology offers rational algorithms to identify genetic markers of complex diseases, even in limited samples.

In conclusion, the concept of predictive medicine—the use of a clinical-genetic passport for addressing preconception screening, the birth of healthy offspring, diagnosis, and prevention of multifactorial as well as infectious diseases—should rely on Next-Generation Sequencing (NGS) technologies as a fundamental tool, utilizing specialized databases, algorithms, bioinformatics, and genetic user concepts of expression and penetrance.

MOLECULAR DIAGNOSTICS OF MONOGENIC AND POLYGENIC DISEASES

The application of Next-Generation Sequencing (NGS) in medical research is multifaceted. Sequencing methods can be categorized into the following groups based on their purpose: (1) whole genome analysis (Whole Genome Sequencing, WGS); (2) analysis of protein-coding genes in the genome (Whole Exome Sequencing, WES); (3) analysis of gene sequences associated with specific diseases (from exomes covering around 4,000–5,000 clinically significant genes to panels targeting small regions of one to three genes); (4) transcriptome sequencing (RNA sequencing); and (5) analysis of bacterial microbiome diversity. Given that the objective is to focus research on protein-coding nucleotide sequences to identify rare pathological SNPs, insertions, deletions that may underlie diseases, or discover new genetic markers for oligogenic and multifactorial diseases, WES is a more cost-effective method than WGS. WES holds a distinctive position among NGS tests. Considering its relatively low cost, WES is currently more attractive in clinical settings. Furthermore, WES significantly reduces the size of the analyzed database to 6 GB, compared to WGS (90 GB).

Due to these qualities, WES is becoming a priority method for complex research and practical objectives. WES/WGS enables the identification of new unknown or rare variants, which are linked to diseases or very rare pathologies. This category also includes susceptibility genes that exist in certain populations but have not yet been considered in differential diagnosis and were not previously recognized as causes of disease.

Whole Exome Sequencing (WES), a study that deciphers protein-coding gene sequences and identifies pathogenic (harmful) genetic variants inherited from parents or occurring spontaneously, is used for mass screening of structural and regulatory genes to determine allele frequencies of disease-associated polymorphisms in different populations. This helps

detect pathogenic genetic changes (mutations or polymorphisms) contributing to protein sequence disorders. With its extensive capabilities, exome sequencing today enables the diagnosis of monogenic diseases and the examination of seemingly healthy populations to identify a wide range of potential risks before disease manifestation. A genetic passport is defined as a new clinical concept based on WES for identifying new candidate genes, gene variants, and molecular mechanisms in the diagnosis, prognosis, and treatment of monogenic, oligogenic, and multifactorial diseases.

It is also necessary to note that cytogenetic research methods remain relevant in the creation of a genetic passport. Certain types of deletions (microdeletions) and translocations (balanced translocations) are relatively common and may be viewed as serious medical issues, for instance, in reproductive medicine or oncology (loss of heterozygosity).

Therefore, the data from cytogenetic analysis is an important component of the genetic profile.

MONOGENIC DISEASES IN HUMANS

As of June 1, 2023, the OMIM database contains entries for 7,377 hereditary diseases and syndromes, including their molecular associations. This includes 6,305 phenotypes linked to a single gene, indicating a monogenic nature of these genetic traits or syndromes. This advancement has largely been achieved through the active implementation of Whole Exome Sequencing (WES).

WES technology provides a reliable method for diagnosing monogenic diseases in humans. Specifically, WES is employed as a first-line examination tool for genetic diagnosis and in idiopathic disorders where probands present with a "blurred" phenotype. For patients with unclear clinical presentations, differential diagnosis using WES is essential in identifying the root cause of the disease.

For instance, a study conducted in China utilized WES to investigate 169 children with congenital developmental anomalies (median age of 10.5 months) and monogenic diseases. Monogenic disorders were diagnosed in 43 (25%) patients. The pathologies with the highest frequency included metabolic (33%), neuromuscular (19%) diseases, and multiple deformations (14%). The diagnostic efficiency for children with metabolic disorders, growth impairments, or visual analyzer anomalies was significantly enhanced by the availability of detailed clinical data. WES data allowed for adjustments in 30 (70%) cases, including disease monitoring initiation in 41.9% (18 cases), rehabilitation and palliative care in 27.9% (12 cases), modification of current treatment in 25.6% (11 cases), and other comprehensive assessment procedures in 7% (3 cases).

Another example of verification through WES of a complex form of monogenic disease is the analysis of a hotspot region in the promoter of the RNA processing endoribonuclease gene (RMRP) in a proband with the extremely rare autosomal-recessive skeletal chondrodysplasia, anauxetic dysplasia (AD). Heterozygous variants of the nucleotide sequence rs387906533 (n.91_92delinsGC) (chr9:35657924-35657925delCTinsGC) were found in exon 1 of the RMRP gene, and an unknown insertion n.-6_-5insTCTCAGCTTCAC (chr9:g.35658020-35658021insTCTCAGCTTCAC) was discovered in the promoter region of the gene. The variant is a 12-nucleotide insertion between the TATA box and the transcription start site. It was found that the n.-6_-

5insTCTCTCAGCTTCAC mutation was paternally derived, while the n.91_92delinsGC mutation was maternally derived. No prior evidence existed regarding the insertion in the RMRP promoter region as a cause of Alzheimer's disease without extraskeletal manifestations typical for carriers of such mutations.

Thus, WES technology enables the rapid and efficient identification of multiple hereditary diseases in any individual. This is particularly relevant for co-inherited X-linked and autosomal-dominant pathological forms such as ichthyosis, Wilson's disease, and hemochromatosis. By understanding the molecular defects leading to the development of these diseases, patients can benefit from the most appropriate follow-up care

GENE MODIFIERS OF MONOGENIC DISEASES

Nevertheless, even with the knowledge of pathological genes and the nature of mutations causing diseases, it is not always possible to accurately forecast key aspects of the phenotypic manifestation of a disease, such as the age of onset, the rate of progression, severity, and the expression of other comorbid conditions. Numerous reports in the literature indicate that patients with the same primary mutation exhibit distinct differences in pathological phenotype expression, most likely due to genetic determinants acting as modifiers. Accumulated clinical and genetic data currently suggest that the view that phenotype expressivity is solely a property of dominant or recessive mutations in a single locus, even in classic monogenic disorders, may be overly simplistic. In this context, phenomena described by terms such as epistasis or genetic interaction/modification are effectively synonyms of the same process, namely, the influence of one gene on the phenotypic outcome of a second gene. A distinction that may help better define these genetic phenomena lies in determining whether the primary locus is both necessary and sufficient for the onset of the disease. If so, the presence of an allele(s) in a second locus/gene plays a purely modifying role on the severity of the phenotype, determined by the rate of disease progression, the severity of the phenotype concerning its pleiotropy, or the expression of endophenotypes that require genetic interactions. Conversely, if not, the requirement of an allele in the second gene for the manifestation of pathology defines a case of digene or oligogenic inheritance (the simplest form of a polygenic disease).

A genetic modifier can influence the phenotype established by mutation(s) in the primary locus in several ways. Depending on the nature of the modifier, this interaction may lead to a reduction in the penetrance of certain endophenotypes and/or disorders, variable expressivity with the modifying locus resulting in either exacerbated phenotypes (i.e., earlier onset and faster disease progression), or conversely, a milder clinical picture (i.e., later onset, slowed disease progression, and absence/protection from certain endophenotypes).

Modifiers that enhance the expression of certain endophenotypes may lead to the description of the given monogenic disease as a new form of the condition.

OLIGOGENIC AND MULTIFACTORIAL DISORDERS

Hereditary disorders caused by pathogenic variants in multiple genes are more prevalent, known as oligogenic hereditary disorders. Oligogenic diseases represent an intermediate state between monogenic diseases, which are linked to a defect in a single gene, and polygenic diseases, which are influenced by multiple genes and environmental factors. Oligogenic models, even those involving only two genes (i.e., digenic or bilocus), remain challenging to detect and validate due to the fact that the involved gene variants might be relatively common in the general population, may have a lesser individual effect on gene function, and may be located in genes not always known to biologically interact. Data concerning digenic diseases, the simplest form of oligogenicity, were first collected in the Digenic Diseases Database (DIDA). Digenic inheritance (DI) involves pathologies with the simplest form of multigenic etiology, engaging more than one gene (and possibly environmental factors). A true DI occurs when biallelic mutations in two distinct genes, in cis- or trans-position, are necessary and sufficient to induce a pathology with a specific diagnosis. The DIDA database facilitated the development of a new generation of prognostic tools aimed at combinations of variants associated with digenic disease. The database also provides an initial credibility assessment for each oligogenic combination based on evidence supporting their association with the disease.

A correct functional evaluation of how variants and genes synergistically interact to cause a pathological phenotype is also crucial for a true understanding of the mechanisms leading to the disease. A "true oligogenic" scenario of an assumed oligogenic pathology corresponds to a situation where all variants need to be present simultaneously to exhibit any symptoms of the disease, whereas in a "monogenic plus modifier" scenario, a variant in the primary gene might still cause milder symptoms or a different subphenotype. Moreover, it is vital to not only consider the presence of individual variants located in different genes but also the instances where genes contain homozygous or heterozygous compound variants, as it is necessary to have evidence indicating that these biallelic events in one gene are insufficient to cause the expected pathological phenotype. In other words, evidence supporting the oligogenic nature of a disease should demonstrate that biallelic events (heterozygosity) in one gene are insufficient to cause the disease phenotype, while individuals carrying mutations in only one gene are either healthy or exhibit less severe symptoms. This implies that a combination of digenic variants is linked to the disease phenotype within a true digenic model, where the simultaneous presence of both variants in an individual is necessary for the development of the disease phenotype.

Another approach to identifying oligogenic diseases involves providing functional evidence of gene interactions indicating an oligogenic combination in the disorder and developing the biological mechanisms in which the genes and their variants are involved. To obtain functional oligogenic evidence, emphasis should be placed on establishing causality between multiple gene variants and the disease phenotype under study, and demonstrating a synergistic or additive effect (i.e., evidence that the effect of the combination is different from the individual effects of the variants) using *in vivo* and/or *in vitro* experiments. Therefore, considering the pathogenicity of an oligogenic

combination at the gene level is crucial for understanding their synergistic action in causing the disease

OLIGOGENIC ETIOLOGY EXEMPLIFIED BY CARDIOMYOPATHIES

Understanding the hereditary factor's role in the pathogenesis of a disease is crucial. Whether a disorder is mono- or oligogenic, or multifactorial, the answer is often not readily apparent. Whole Exome Sequencing (WES) technology assists in addressing this question. WES was utilized to identify six new P (pathogenic) or LP (likely pathogenic) gene variants in 40 patients with hypertrophic (HCM, n=14) and dilated cardiomyopathy (DCM, n=26). Hypertrophic cardiomyopathy, caused by gene variants encoding sarcomeric proteins—heavy chain β -myosin (MYH7) and myosin-binding protein C (MYBPC3)—accounts for up to 50% of all clinical cases: myosin light chain 2 (MYL2), myosin light chain 3 (MYL3), and cardiac troponin T (TNNT2) in 5–10%, cardiac troponin I (TNNI3) in 5%, cardiac troponin C (TNNC1) in <1%, cardiac α -actin (ACTC1) in <1%, α -tropomyosin (TPM1) in 1.5%, and cysteine and glycine-rich protein 3 (CSRP3).

As a result, frameshift mutations (11:47372858, c.A224insG+) in the MYBPC3 gene and missense mutations (rs193922390, c.5135 G>A, p.R1712Q) in the MYH7 gene were identified as pathogenic variants for HCM. Missense variants such as rs138049878 (c.2608 C>T, p.R870C), rs727503260 (c.2302 G>C, p.G768R) and rs397516088 (c.1063 G>A, p.A355T) in the MYH7 gene, and rs199476306 (c.188 C>T, p.A63V) in the TPM1 gene, were assessed as likely pathogenic for HCM. The diagnostic yield of WES for HCM was 43% (with six variants in fourteen patients). This is a relatively high diagnostic efficiency, supporting the involvement of a limited number of genes in the pathology's development.

MULTIFACTORIAL (POLYGENIC) DISEASES

Multifactorial diseases (polygenic diseases) encompass almost all the most prevalent chronic human ailments, such as atherosclerosis, diabetes, obesity, bronchial asthma, osteoporosis, endometriosis, malignant tumors, as well as neuropsychiatric and cardiovascular diseases. These conditions arise from the interaction of multiple genes with adverse environmental factors. Currently, the International Classification of Diseases (ICD) includes over 55,000 nosological units. The overwhelming majority of these are classified as multifactorial diseases. As of March 11, 2023, more than 12,000 human diseases, 220,322 SNPs, and 493,105 genomic associations have been registered. Presently, there are three primary approaches to identifying candidate genes: functional mapping (candidate gene analysis), genetic linkage in high-risk families, and genome-wide association studies (GWAS), including genome sequencing. GWAS are actively employed for the analysis and testing of samples from various national biobanks. According to current knowledge, the influence of genetic factors on the expression and penetrance of phenotypic traits is either due to genetic polymorphisms of point mutations with strong effects or frequent SNPs with weak effects. For understanding the genotype-phenotype correlation, the preferred approach is the investigation of candidate genes associated with multifactorial diseases.

MEDICAL ASPECTS OF GENE MUTATIONS AND CHROMOSOMAL ABERRATIONS EXEMPLIFIED BY THE DIAGNOSIS OF HEREDITARY CAUSES OF INFERTILITY

Infertility is defined as the inability of a couple to conceive a child after attempting for two years without success. Approximately 15% of couples are infertile, with the male factor being responsible for about 50% of these cases. Often, though not always, male infertility is a multifactorial syndrome encompassing a wide range of disorders. In more than 50% of infertile men, the cause is unknown (idiopathic) and may be congenital or acquired. Male infertility can be initially diagnosed through semen analysis. Semen analyses of infertile men can reveal various abnormal conditions, such as azoospermia, oligozoospermia, teratozoospermia, asthenozoospermia, etc. Current estimates suggest that about 30% of men seeking medical assistance for infertility face oligozoospermia or azoospermia of unknown etiology. The causes of infertility are known in less than 50% of cases, with genetic or hereditary conditions and specific chromosomal anomalies being the main factors. Approximately 10-20% of men without sperm in the ejaculate have a deletion in the Y chromosome. This deleted region includes the azoospermia factor (AZF) locus, located in the Yq11 region, divided into four repeat-deleted non-overlapping subregions known as AZFa, AZFb, AZFc, and AZFd. Each subregion may be associated with specific testicular histology, and several candidate genes have been identified in these areas. The Deleted in Azoospermia (DAZ) gene family is reported to be the most frequently deleted candidate gene in AZF, located in the AZFc region. Additionally, the DAZ gene has an autosomal homologue, DAZL (DAZ-Like), on the short arm of chromosome 3 (3p24), and a mutated autosomal DAZL might be responsible for a spermatogenic defect. Random attacks on exposed mitochondrial DNA (mtDNA) of spermatozoa by reactive oxygen species or free radicals inevitably cause oxidative damage or mutation of the mitochondrial genome, with pathological consequences, leading to male infertility. A key nuclear enzyme involved in the extension and repair of mtDNA strands is DNA polymerase gamma, mapped to the long arm of chromosome 15 (15q25), comprising a CAG repeat region. Its mutation affects ATP production. The introduction of molecular methods has significantly advanced the understanding of the genetics of infertility. However, the comprehension of the genetic causes of male infertility remains limited.

CHROMOSOME STRUCTURE AND NUMBER ABNORMALITIES: A CAUSE OF INFERTILITY

It is widely recognized that cytogenetic defects are a significant factor in disrupting spermatogenesis. The connection between hereditary factors and male infertility was first established in 1959 when it was shown that men with Klinefelter Syndrome (KS) have an additional X chromosome, leading to a 47,XXY karyotype. KS is a common condition (~1–2 men per 1000) and one of the leading genetic causes of male infertility (3–4% of cases). Patients with KS typically exhibit small testes, non-obstructive azoospermia (NOA), and elevated levels of FSH and LH hormones, a condition known as hypergonadotropic hypogonadism. Chromosomal abnormalities are confirmed in 5% of patients with severe oligozoospermia and in 10%-15% of patients with azoospermia. Chromosome analysis is usually performed by culturing lymphocytes (72 hours) using cytogenetic methods. During routine analysis, 20 cells are examined. In cases of chromosomal mosaicism or abnormalities, 30 cells are analyzed.

Besides KS, up to 1.7% of patients with male infertility have autosomal abnormalities, such as translocations or inversions. Individuals with karyotype abnormalities affecting both

the number and structure of chromosomes have a predisposition to infertility due to unsuccessful pregnancies, often resulting from recurrent miscarriages. Key chromosomal abnormalities that can lead to reduced fertility include chromosomal translocations or inversions, aneuploidy, and Y chromosome microdeletions. In patients with structural anomalies (primarily concerning translocations), the formation of normal bivalents during meiosis is disrupted, leading to expected meiotic disturbances and halting spermatogenesis maturation. Most patients with structural chromosome changes exhibit oligozoospermia. Therefore, the frequency of Robertsonian translocations, reciprocal translocations, and inversions is higher in men with oligozoospermia compared to men with azoospermia and the general male population.

The nature of infertility treatment depends on understanding the precise chromosomal rearrangement and the degree of spermatogenesis disruption. Preimplantation and prenatal genetic testing options may be employed to prevent transmitting genetic defects to offspring. For instance, carriers of Robertsonian translocations have a high risk of conceiving offspring with aneuploidy, while asymptomatic carriers of microdeletions or duplications associated with clinical syndromes risk fathering a child with severe developmental disorders. Chromosomal translocations are the most common structural abnormalities in men, with a frequency of 1.23 per 1000, and their prevalence is ten times higher in infertile populations. Chromosomal translocations are classified as balanced and unbalanced. A balanced reciprocal translocation involves the exchange of genetic material between two or more chromosomes. There are both autosomal and sex chromosome translocations in balanced reciprocal translocations. Depending on breakpoints, approximately 60% of autosomal translocation carriers have at least one abnormal parameter in sperm analysis. Although the frequency of sex chromosome translocations is low, some reports have linked Y chromosome translocations to azoospermia.

Robertsonian translocation is the most common form of unbalanced chromosomal translocation in humans and a frequent cause of male infertility. Robertsonian translocations occur in 0.9–3.4% of infertile men with severe spermatogenic dysfunction. They can arise in five pairs of acrocentric chromosomes (chromosomes 13, 14, 15, 21, 22) and result in breaks at centromere regions, with the fusion of two long arms to form a single large chromosome. Consequently, individuals with Robertsonian translocation have 45 chromosomes. The remnants of the short arms of the two fused chromosomes are typically lost. Despite this genetic anomaly, Robertsonian translocation carriers are phenotypically normal, as the short arms of the acrocentric chromosomes do not contain important genes. However, carriers are at increased risk of sperm aneuploidy, potentially leading to miscarriage or offspring with severe genetic abnormalities. Theoretically, one-sixth of the sperm of carriers has a normal karyotype, another one-sixth carries the Robertsonian translocation, and the remaining two-thirds are in an unbalanced state, either nullisomy or disomy for the chromosomes involved in the translocation. Prader-Willi syndrome and Angelman syndrome are associated with maternal and paternal uniparental disomy of chromosome 15.

One prevalent cause of male infertility is microdeletions of the Y chromosome's long arm (Yq). The long arm of the Y chromosome (Yq) contains numerous ampliconic and palindromic sequences, making it prone to self-recombination during spermatogenesis and, thus, susceptible to intrachromosomal deletions. These deletions alter the copy number of Y chromosome genes, leading to male infertility. Three common Yq deletions recurring in infertile men are known as AZF (azoospermia factor) microdeletions, specifically AZFa, AZFb, and AZFc. Estimates based on studies of nearly 40,000 Y chromosomes suggest that the global prevalence of Yq microdeletions is 7.5% among infertile men; however, infertile European men are less susceptible to Yq microdeletions, with the highest prevalence

observed in American and East Asian infertile men. Additionally, partial deletions of the AZFc locus have been associated with infertility, but the effect appears to be ethnicity-dependent.

MONOGENIC CAUSES OF INFERTILITY

The issue of monogenic causes of infertility is general in nature, affecting a wide array of medical circumstances and is not an isolated concern. More than 50% of infertile men suffer from one or multiple chronic conditions or significant health issues not directly related to reproduction. Recent meta-analyses have clearly demonstrated that infertile men are at a higher risk of early mortality. Furthermore, the risk of early death increases with the severity of spermatogenesis disorders and decreased sperm quality. The exact mechanisms of this correlation, observed in epidemiological studies, remain unknown.

A significant portion of registered monogenic forms of spermatogenic defects involves homozygous autosomal-recessive pathogenic variants, mapped in consanguineous families. Infertility is a widespread clinical problem, affecting 8-12% of couples globally, with approximately 75% of cases diagnosed as idiopathic infertility due to undetected causal factors. Indeed, 40-60% of spermatogenesis disorders remain unexplained, and among cases of moderate oligozoospermia (sperm count 10-39 million per ejaculate), this share is close to 80%. In recent years, considerable efforts have been made to investigate new candidate genes responsible for male infertility driven by monogenic mutations.

Quantitative or structural chromosomal abnormalities have long been associated with male infertility. Monogenic mutations have only recently been considered in the pathogenesis of this condition. Mutations in specific genes involved in meiosis, mitosis, or spermiogenesis lead to spermatogenic insufficiency, resulting in conditions such as insufficient (oligozoospermia) or absent (azoospermia) sperm production, limited progressive and/or total sperm motility (asthenozoospermia), altered sperm morphology (teratozoospermia), or combinations thereof.

It is important to note that monogenic causes of infertility can be due to genes located on both the Y chromosome and autosomes. For instance, as mentioned, Yq microdeletions result from intrachromosomal non-allelic homologous recombination events (due to the highly repetitive structure of the Y chromosome) and are generally present in 5%-10% of patients with azoospermia. In this case, the pathological effect on male fertility is caused by the loss of several key regulatory genes of spermatogenesis.

Currently, active developments are underway for genetic panels aimed at detecting monogenic causes of male infertility, driven by spermatogenic insufficiency and androgen insensitivity. These are based on testing genes located on both autosomes and the Y chromosome.

For diagnosing monogenic causes of infertility in men, a polygenic NGS panel is typically used to detect nucleotide variations in the coding exons and flanking introns of the aforementioned genes. Pathogenic variants may include missense, nonsense, splicing, or small indels. Large deletions/duplications have been reported in genes defining male infertility, such as USP9Y, DPY19L2, SPATA16, SUN5, and CFAP43.

MONOGENIC CAUSES OF SPERMATOGENIC INSUFFICIENCY

Defects in sperm quantity include azoospermia and oligozoospermia. Azoospermia, the absence of sperm in seminal fluid, affects 1% of the male population and accounts for 20% of all male infertility cases. Recent discoveries have identified point mutations causing

azoospermia in the following genes: NR5A1, SYCP3, ZMYND15, TAF4B, TEX11, NANOS1, PLK4, MEIOB, SYCE1, USP9Y, SOHLH1, TEX15, HSF2, and KLHL10. Notably, frameshift mutations in ZMYND15 cause the SPGF14 phenotype. The protein encoded by this gene is involved in the temporally normal haploid gene expression during spermatogenesis.

A homozygous mutation in SYCE1 is associated with the SPGF15 phenotype. This gene encodes a component of the synaptonemal complex, a structure that physically binds homologous chromosomes during meiosis I.

Mutations in TEX11 are linked to meiotic arrest and azoospermia with a frequency of 1-15% in men with azoospermia. The protein encoded by TEX11 regulates the pairing of homologous chromosomes during double-strand DNA repair through the formation of a synaptonemal complex and chiasma during crossover processes. SYCP3, which also plays a similar role, has been found to be mutated in infertile men.

The SOHLH1 gene mutates in some cases of azoospermia and encodes a testicular transcription factor essential for spermatogenesis. A mutation in the NR5A1 gene, encoding steroidogenic factor 1, has been reported in a patient with meiotic arrest and normal levels of follicle-stimulating and luteinizing hormones, as well as testosterone.

Lastly, a patient with spermatogenesis blocked at the spermatocyte stage exhibited a dominant-negative mutation in the HSF2 gene, which encodes the transcription factor heat shock factor 2.

MONOGENIC CAUSES OF SPERM MORPHOLOGY DEFECTS

Teratozoospermia represents a heterogeneous group of disorders. The morphological assessment of sperm involves evaluating the major functional regions (head, midpiece, and tail), which may exhibit abnormalities in shape and size. The phenotype may present as a single type of malformation in one patient or different types in the same patient. Recent studies involving families with pathological cases and small phenotypically homogeneous cohorts have identified autosomal recessive cases of teratozoospermia. Currently, the genes associated with sperm morphology defects include: AURKC, ZBPB, DPY19L2, SPATA16, PICK1, BRDT, and SUN5.

MONOGENIC CAUSES OF SPERM MOTILITY DEFECTS

Asthenozoospermia is a condition leading to reduced sperm motility due to flagellar defects. The axoneme, outer dense fibers, mitochondria, or fibrous sheath of the flagellum might be affected. Ultrastructural defects in the 9+2 axoneme structure can involve outer or inner dynein arms, central microtubules, and radial spokes. Genes related to sperm motility defects include: SLC26A8, CATSPER1, SEPT12, CFAP43, CFAP44, DNAH1, and PLCZ1.

POLYGENIC CAUSES OF MALE INFERTILITY

Despite significant progress made through cytogenetic studies and next-generation sequencing (NGS) technologies in uncovering the genetic causes of male infertility, a large proportion of unexplained heritability remains to be explored. Indeed, in many patients with spermatogenesis disorders, the etiology remains unknown. Meanwhile, increasing evidence highlights idiopathic non-obstructive azoospermia as a complex and polygenic condition. Over the past three decades, efforts have been made to identify polymorphisms in candidate

genes associated with complex forms of idiopathic infertility. In this context, SNPs have been the most analyzed variations compared to studies on VNTR and CNV.

Candidate genes for predisposition have been selected based on their involvement in hormone production, cell cycle regulation, and spermatogenesis. However, substantial progress in these studies has been elusive, primarily due to (1) the extremely small size of analyzed case/control cohorts, likely leading to numerous errors in statistical hypothesis testing, and (2) inadequate clinical characterization of the included patients.

Several population studies have suggested the productive analysis of approximately 2000 genes directly involved in spermatogenic processes, with over 600 exhibiting specific expression in male germ cells. Such a vast number of candidate genes complicates the visualization of the overall molecular picture of azoospermia, causing skepticism among some researchers about transferring molecular genetic insights into clinical healthcare practice.

Some research groups have begun regularly analyzing the gene sequences of NR5A1, DMRT1, and TEX11 in patients with idiopathic variants of non-obstructive azoospermia, identifying causative mutations in 5% of cases and increasing diagnostic efficiency to 25%. In routine clinical practice, about 75% of cases with severe spermatogenesis disorders are of idiopathic origin. A list of 60 candidate genes is proposed for diagnosing male infertility caused by spermatogenesis disorders. It is suggested that next-generation sequencing studies in large cohorts of patients with clearly idiopathic infertility will provide new, intriguing data on their racial and ethnic prevalence among infertile patients, likely enhancing diagnostic effectiveness.

RECOMMENDED LITERATURE FOR CHAPTER 1. ROLE OF GENETIC APPARATUS DISORDERS IN THE PATHOGENESIS OF HEREDITARY DISEASES

1. Genetic Passport as the Basis of Individual and Predictive Medicine. Edited by V. S. Baranov. St. Petersburg: N-L Publishing, 2009. 528 pages.

2. Glotov O.S., Chernov A.N., Glotov A.S. Human Exome Sequencing and Prospects for Predictive Medicine: Analysis of International Data and Own Experience. *Journal of Personalized Medicine*, 2023; 13(8): 1236. doi: 10.3390/jpm13081236

3. Guerri G., Maniscalchi T., Barati S., et al. Non-syndromic Monogenic Male Infertility. *Acta Biomedica*, 2019; 90(Suppl 10): 62–67. doi: 10.23750/abm.v90i10-S.8762

4. Laan M., Kasak L., Punab M. Translational Aspects of Novel Findings in Genetics of Male Infertility—Status Quo 2021. *British Medical Bulletin*, 2021; 140(1): 5–22. doi: 10.1093/bmb/ldab025

5. Colaco S., Modi D. Genetics of the Human Y Chromosome and Its Association with Male Infertility. *Reproductive Biology and Endocrinology*, 2018; 16: 14. doi: 10.1186/s12958-018-0330-5

6. Harton G.L., Tempest H.G. Chromosomal Disorders and Male Infertility. *Asian Journal of Andrology*, 2012; 14(1): 32–39. doi: 10.1038/aja.2011.66

7. Kuroda S., Usui K., Sanjo H., et al. Genetic Disorders and Male Infertility. *Reproductive Medicine and Biology*, 2020; 19(4): 314–322. doi: 10.1002/rmb2.12336

8. Massart A., et al. Genetic Causes of Spermatogenic Failure. *Asian Journal of Andrology*, 2012; 14(1): 40–48. doi: 10.1038/aja.2011.67

9. Cerván-Martín M., et al. Genetic Landscape of Nonobstructive Azoospermia and New Perspectives for the Clinic. *Journal of Clinical Medicine*, 2020; 9(2): 300. doi: 10.3390/jcm9020300

10. Papadimitriou S., Gravel B., Nachtegaele C., et al. Toward Reporting Standards for the Pathogenicity of Variant Combinations Involved in Multilocus/Oligogenic Diseases. HGG Advances, 2023; 4(1): 100165. doi: 10.1016/j.xhgg.2022.100165

11. Kousi M., Katsanis N. Genetic Modifiers and Oligogenic Inheritance. Cold Spring Harbor Perspectives in Medicine, 2015; 5(6): a017145. doi: 10.1101/cshperspect.a017145

CHAPTER 2. GENE POLYMORPHISM

INTRODUCTION

A genomic map is constructed based on the results of genome DNA sequencing. By determining the sequences of nitrogenous bases, we can identify genes and the distances between them. By analyzing the nucleotide sequence of DNA involved in encoding a given protein, we can assess the functions performed by this nucleotide sequence. It is hypothesized that natural selection prevents the accumulation of harmful mutations in sequences that code for functional products (proteins). Therefore, it can be assumed that an intact coding sequence will lead to the formation of a functionally active polypeptide. By comparing the DNA sequence of the wild type (non-mutated, original gene) with the sequence of a mutant gene, it is possible to determine the nature of the mutation and its exact location in the sequence. The initial Mendelian view of the genome classified alleles as either wild type or mutant. It was later recognized that multiple alleles for a single gene could exist within a population, each having varying effects on the phenotype. In some cases, it may even be inappropriate to designate any one allele as the wild type.

The coexistence of multiple alleles at a single gene locus within a population is referred to as genetic polymorphism. Any site where multiple alleles exist as stable components of the population is, by definition, polymorphic. A locus is typically defined as polymorphic if two or more alleles are present with a frequency greater than 1% in the population. Human eye color is a good example of phenotypic polymorphism resulting from underlying genetic polymorphism. There is no singular "normal" eye color; many different colors occur among different individuals, with minimal differences in visual function between them. What underlies the polymorphism among different alleles? They possess various mutations that can alter the function of their product, thereby causing changes in the phenotype. The population dynamics of these different alleles are partially determined by their selective effects on the phenotype. The wild type itself can be polymorphic. Multiple versions of the wild type allele can be distinguished by sequence differences that do not affect their function and, consequently, do not lead to phenotypic variants.

From a molecular mechanism perspective, the three main types of variations at the DNA level (DNA polymorphism markers) are, firstly, single nucleotide changes - SNPs; secondly, insertions or deletions (Indels) of varying lengths ranging from 1 to several hundred base pairs; and thirdly, VNTR - variations in the number of tandem nucleotide repeats. Mutation mechanisms lead either to transitions: purine-purine (A↔G) or pyrimidine-pyrimidine (C↔T) exchanges, or transversions: purine-pyrimidine or pyrimidine-purine (A↔C, A↔T, G↔C, G↔T) replacements. There is a noticeable predominance of transitions. What accounts for the growing popularity of SNPs in terms of the research or diagnostic genetic information provided? The fact is that the more modern concept of SNP primarily emerged due to the need for very high-density genetic markers for the study of multifactorial diseases and recent advances in methods for detecting polymorphism and genotyping (Vignal A. et al., 2002).

MUTATIONS – THE CAUSE OF GENE POLYMORPHISM

Genetic variability, defined as differences in the primary structure of DNA caused by SNPs, Indels, or VNTRs, within a single biological species is referred to as genetic polymorphism. The primary DNA structures (nucleotide sequences) in humans (except for monozygotic twins) differ significantly. These differences, whether at the population, ethnic, or individual level, include both coding regions (exons) and non-coding sequences (intergenic

sequences, introns, regulatory regions, and more) and are attributed to various mutations that cause gene polymorphism. Genetic polymorphism is generally considered a Mendelian trait, occurring in a population with at least two gene variants, each with a frequency of no less than 1%. An allelic gene is defined as a gene variant present in a population with a frequency of at least 1%. In contrast, a gene with a frequency below 1% is termed a rare gene variant. Rare gene forms can have significant medical implications if they are pathological or predispose individuals to certain diseases.

Mutations that lead to gene polymorphism can be categorized as qualitative or quantitative. Qualitative mutations involve changes in nucleotide substitutions while maintaining the same number of nucleotides in the DNA chain. Quantitative mutations involve changes in the number of nucleotides in the DNA strand, such as variations in the length of nucleotide sequence repeats. Both causes of gene polymorphism can occur in gene-coding regions, introns, promoters, and non-genic DNA sequences.

Qualitative genetic polymorphism is mainly represented by single nucleotide polymorphisms (SNPs) and occurs most frequently. Comparative genomic studies across different human races and ethnic groups have established deep genetic kinship among all humans, providing valuable insights into anthropogenesis, human migration, and the formation of ethnic groups. Approximately 70% of SNPs are neutral, meaning they do not significantly affect the protein's function, reduce individual viability, fertility, or offspring vitality. Nevertheless, neutral mutations form the basis for individual traits, contributing to individual organism diversity. Meanwhile, the remaining 30% of SNPs are involved in the pathogenesis of monogenic or polygenic diseases. Thus, whole-genome SNP analysis is crucial for both fundamental biology and practical medicine.

Quantitative gene polymorphism can be illustrated by variations in the number of tandem repeats (STRs – Short Tandem Repeats), which are categorized into microsatellite or minisatellite DNA depending on the repeat group size. VNTR (Variable Number Tandem Repeats) refers to sequences with longer repeated nucleotide groups. Gene-coding regions (exons) mainly have trinucleotide repeats as they do not shift the reading frame. These trinucleotide repeats often consist of the CAG (cytosine–adenine–guanine) sequence, encoding glutamic acid. In most cases, the increase or decrease in repeats is neutral and promotes individual diversity, similar to neutral SNPs.

However, such polymorphisms are characteristic of "expansion diseases" genes. In the case of dynamic mutations (changes primarily in the number of trinucleotide repeats), reaching a critical repeat count transforms genetic polymorphism from functionally neutral to a pathological mutation type known as "dynamic mutations." Pathological repeat increases can occur in coding or regulatory gene parts, resulting in dysfunction. These mutations can lead to pathological protein formation or gene expression disruption. Dynamic mutations are considered causes of numerous neurodegenerative diseases (Huntington's disease, Kennedy disease, spinocerebellar ataxia, and others). Clinical features of such diseases include late-onset symptoms, increased severity in subsequent generations, and a lack of effective treatment methods.

SNP, SINGLE NUCLEOTIDE POLYMORPHISM

A single nucleotide polymorphism is the substitution of one nitrogenous base in a DNA sequence compared to the original form of the gene. A single nucleotide polymorphism occurs at a specific nucleotide site. For instance, some DNA molecules in the same population may have the base pair T–A at a specific nucleotide site, while other DNA molecules in the same population may have the base pair C–G at the same location. This difference constitutes an SNP

(single nucleotide polymorphism). An SNP is defined by two alleles for which there can be three genotypes among individuals in a population—homozygous chromosomes (or heterozygous chromosomes with T–A in one chromosome and C–G in the homologous chromosome).

Thus, when alleles at the same locus are compared, the identified difference at one nucleotide is called a single nucleotide polymorphism (SNP). On average, one SNP occurs approximately every 1,330 bases in the human genome. Due to the specific SNP landscape in a DNA sequence, each person is unique. The presence of SNPs can be detected by directly comparing sequence data from different individuals. One of the goals of genetic mapping is to create a catalog of common gene variants. The observed SNP frequency per genome predicts that there should be more than 10 million SNPs in the human population overall (considering the genomes of all individuals in the population) that occur with a frequency greater than 1%—hence, are polymorphic.

SNPs account for about 90% of all human genetic variations. At a certain locus, there can be many different sequence variants. Some of these sequences affect the phenotype, while others are "hidden" because they have no visible effect. In addition to allele frequency information, the genomic location of SNPs and their presumed functional significance (coding, intronic, promoter, etc.) are important for SNP selection (Xu H. et al., 2005). There can be numerous changes at a locus, including those that alter the DNA sequence without changing the polypeptide sequence, those that change the polypeptide sequence without altering its function, those that lead to polypeptides with different functions, and those that result in altered polypeptides that are non-functional. It is generally considered that the complete human DNA sequence contains at least a million SNPs, both in coding regions and in introns and promoters.

There is no doubt that genetic variability is a ubiquitous phenomenon and forms the genetic foundation of species diversity. With the completion of the Human Genome Project, the study of genetic variations has become one of the cornerstones of biomedical research. This is not only because gene variants influence the diversity of human anthropometric characteristics, but also because they affect the risk of disease development and the body's response to environmental challenges. The information obtained from the study of hereditary variability not only deepens our understanding of human genes and evolution but also assists in the identification and treatment of human hereditary diseases.

Since the most common variants are SNPs (Single Nucleotide Polymorphisms), this term is often used as a synonym for genetic sequence variation. However, here we restrict its use to the formal definition of SNP: a single nucleotide position change. As a result of high-throughput genotyping methods, millions of human SNPs have been registered in recent years. For example, several studies have evaluated how SNPs influence phenotype. In particular, the biological effects of amino acid substitutions caused by non-synonymous coding SNPs are observed. The impact of non-synonymous coding SNPs on protein function is studied, induced by changes in amino acid sequences in the protein's primary structure.

Another genomic area essential for gene functioning is the regulatory regions of genes. Gene promoters, through the binding of specific transcription factors, directly participate in the initiation and regulation of gene transcription. Thus, sequence changes in gene promoters can alter the identification and binding of transcription factors, which, in turn, can affect gene expression and exert biological effects. For example, one allele of the HLA-G gene (human leukocyte antigen gene), whose products inhibit the maternal anti-fetal immune response, is significantly associated with an increased risk of intrauterine fetal death. One possible explanation is that an SNP falls into the binding site of the interferon response factor-1 (IRF-1), influencing IRF-1 binding and downregulating the transcription of the HLA-G gene. Therefore, an independent research direction is the study of SNP distribution in the regulatory regions of

human genes. Putative transcription factor binding sites in gene promoters have been computed and compared with previously identified SNPs. The results indicate that SNPs have specific distribution patterns, both in regulatory regions of genes and in transcription factor binding sites, compared to the frequency of SNP occurrence in other DNA regions.

More than 35,000 SNPs have been identified in the regulatory regions of genes. Over 99.8% of these SNPs are represented by two alleles. The frequency of nucleotide substitution in promoter SNPs with two alleles is shown in Table 1. Literature suggests that the high frequency of SNP occurrence in the promoter region may be due to their role in altering gene expression during evolution by influencing the sequences of promoter binding sites. The effects of SNPs on transcription factor binding sites are not subject to straightforward interpretation. While some sequence changes introduced by SNPs can completely disrupt gene expression, others may only affect the level of expression (Y. Guo and D. C. Jamison, 2005).

Table 1: Frequency of Nucleotide Substitution in Promoter SNPs with Two Alleles in the Promoter Region. Cited from Y. Guo and D. C. Jamison, 2005.

Purine	A ↔ G: 33.10%
Pyrimidine	C ↔ T: 33.10%
Purine ↔ Pyrimidine	C ↔ G: 8.93%
Purine ↔ Pyrimidine	A ↔ C: 8.77%
Purine ↔ Pyrimidine	G ↔ T: 8.82%
Purine ↔ Pyrimidine	A ↔ T: 7.42%

During the course of evolution, the necessity for adapting specific levels of gene expression to certain environmental factors leads to natural selection for gene regulation. In this regard, single nucleotide polymorphisms (SNPs) found in transcription factor binding sites have a direct impact on gene expression levels. Consequently, the fact that a greater number of SNPs are located in transcription factor binding sites may reflect varying requirements for the expression of different genes under different conditions. Thus, promoter SNPs may serve as active factors in natural selection when choosing the mode of gene transcription.

It is important to emphasize that SNPs associated with disease, according to genome-wide association studies (GWAS), are most frequently found in non-coding specific regulatory regions of DNA—transcription factor binding sites (TFBS) (Chen C. et al., 2014). An analysis of differential affinity for transcription factor binding between regular alleles and risk alleles identified functional SNPs with potential disruption or enhancement of transcription factor binding. Researchers identified regulatory SNP candidates, one of which exhibits borderline susceptibility to lung cancer and significant susceptibility to breast cancer.

Many functional studies have focused on SNPs located in coding regions, which can affect phenotype by altering the primary structure of encoded proteins. Additionally, SNPs can lead to premature termination of translation and mRNA decay. Another function of SNPs is their influence on splicing sites, resulting in alternative splicing.

Whole-genome sequencing of individuals is now widely used, enabling the assessment of individual variations at the DNA level, including both neutral SNPs and those associated with diseases or disease susceptibility. To date, hundreds of individual human genomes from all major racial groups have been sequenced, including genomes of Denisovans (a species of Homo from the Paleolithic era, living over 30,000 years ago) and Neanderthals (over 25,000 years old). The "1000 Genomes Project" was conducted from 2008 to 2015 with the goal of identifying common human genetic variants through deep sequencing of at least 1,000 human genomes; the final number was actually 2,504 anonymous human genome sequences representing 26 human populations. A large proportion of polymorphic sites indicates that each individual has a unique

set of SNPs. A specific combination of sites found in a certain region is called a haplotype and represents a small part of the complete genotype. The term "haplotype" was originally introduced to describe the genetic content of the major histocompatibility locus in humans, an area that determines proteins playing crucial roles in immune mechanisms. Today, this term has been expanded to describe a specific combination of alleles or any other genetic markers present in a defined region of the genome. Using SNPs, a detailed map of human genome haplotypes has been created; this allows researchers to more easily identify disease-causing genes.

SNPS AND HEREDITARY DISEASES

Human hereditary diseases are caused by various types of gene mutations, including insertions/deletions, genomic structural variations, and pathogenic single nucleotide polymorphisms (SNPs), which are closely linked to an increased risk of diseases. Among these, monogenic diseases pose a significant challenge for gene therapy. Patients with such conditions inherit one pathogenic allele causing the disease phenotype, particularly in a dominant-negative manner, along with one normal allele. Therefore, treatment strategies typically involve the specific suppression of pathogenic alleles while sparing the wild-type alleles. Genetic diseases are caused entirely or partially by an anomaly in an individual's genetic makeup. In the human genome, the most frequent variations are SNPs. The average human genome contains approximately 4-5 million SNPs, occurring every 1,000-2,000 nucleotides, providing sufficient density for comprehensive haplotype analysis. As SNPs are the most common form of polymorphism in the human genome, they are widely used in anthropogenetic studies, hereditary disease research, and pharmacogenetic studies.

In recent years, a large number of genes contributing to monogenic and multifactorial (polygenic) diseases have been identified. Consequently, specialized bioinformatics databases and software related to the collection and analysis of genetic data have significantly increased in quantity, size, and scope. Single nucleotide polymorphisms (SNPs), as the most prevalent type of genetic variability, now form the primary basis for most genetic research and databases. While other types of variations, including indels, copy number variants, and epigenetic markers, remain important considerations and can influence disease, SNPs are significantly easier to detect and serve as the most useful and widely applied markers in modern genetic research (Johnson A.D., 2009). Generally, many SNPs do not affect cellular function, but some SNPs have been reported to be closely associated with diseases or influence cellular responses to drugs. Although more than 99% of human DNA sequences are identical across populations, certain SNPs can have a substantial impact on how individuals react to environmental factors (pathogenic bacteria, viruses, toxins, and chemicals), as well as drugs and other treatments. This makes SNPs highly valuable for biomedical research, pharmaceutical development, and medical diagnostics.

Genetic markers provide a foundation for a highly productive method of identifying genetic variants at the molecular level. A typical issue involves a mutation with known effects on the phenotype, where the relevant genetic locus can be placed on a genetic map, but where there is no knowledge about the corresponding gene or its product. Many serious and even fatal human diseases fall into this category. For instance, cystic fibrosis demonstrates recessive Mendelian inheritance, but the molecular nature of the mutant protein was unknown until it was identified based on the characterization of the gene. In the genome, SNPs may appear close to or within a specific target gene. Researchers can identify such markers due to their close association with the gene responsible for the mutant phenotype. By comparing the DNA of patients suffering from a disorder with that of healthy individuals, certain markers that are always present (or always absent) in patients can be identified.

A pathological phenotype is not always determined by a single gene; hereditary pathology is often polygenic or multifactorial. Additionally, some SNPs associated with diseases may not have a functional relation to the disorder. However, their usefulness in the analysis of complex genetic traits can be illustrated by two SNPs linked to disease phenotype components. These include a mutation that results in a gain of function in the interleukin-4 receptor gene, associated with severe asthma, and a mutation in the β 2-adrenergic receptor gene (ADRB2), linked to greater resistance to adrenergic drugs. Advancements in molecular biology technologies have led to understanding the importance of identifying SNPs associated with widespread polygenic diseases such as osteoporosis, diabetes, cardiovascular and inflammatory diseases, mental disorders, and most types of cancer (Bin Alwi Z., 2005).

The widespread familial nature of complex polygenic diseases clearly indicates a significant genetic component. However, unlike the monogenic inheritance of diseases, this genetic element consists of numerous gene variants, each contributing a small effect. When a certain set of susceptibility genes is present in an individual's genotype, their combined contribution can provoke pathology, especially in conjunction with environmental factors. It is important to note that SNPs are numerous and more stable than tandem repeats due to their lower mutation frequency. Many SNPs can manifest in the phenotype if they occur in coding or regulatory regions of a gene. Therefore, using SNP markers, it is often possible to verify a direct link between the phenotype and the functional variant. For these reasons, SNPs are preferred for constructing high-density genetic marker maps, essential for one of the main research directions in human genetics: identifying complex genetic traits.

SNP AND DIABETES MELLITUS

Type 2 diabetes is a hereditary polygenic (multifactorial) disease. Recent research findings indicate that heredity is a significant factor contributing to this disease. Firstly, certain ethnic minorities and indigenous groups with low admixture populations (such as the Pima Indians, residents of Micronesia and other Pacific Islands, Australian Aboriginals, and Mexican Americans) exhibit exceptionally high prevalence rates of type 2 diabetes (up to 21% among the Pima Indians). Secondly, there are family groups predisposed to type 2 diabetes. Members of these biologically related families are generally more predisposed to developing the disease compared to the population in which they reside. Finally, twin studies in medical genetics have shown a notably higher correlation of type 2 diabetes risk among monozygotic twins compared to dizygotic twins (70% versus 10%) (Staiger H. et al., 2009). Like other polygenic diseases, type 2 diabetes does not follow simple Mendelian inheritance and is considered a polygenic disorder. Complex diseases, such as type 2 diabetes, arise from the simultaneous occurrence of common DNA sequence variations in many genes in combination with specific environmental factors. It is suggested that each of these DNA changes has a moderate impact on the function of the affected genes. Nevertheless, collectively, these variations confer increased susceptibility to adverse environmental factors. Single nucleotide polymorphisms (SNPs), substitutions in single base pairs, account for approximately 90% of sequence variations in the human genome. For this reason, they are considered primary determinants of individual susceptibility to such complex diseases.

In the past decade, at least 600 (confirmed and potential) genes associated with diabetes predisposition have been identified. The greatest success in the genetics of type 2 diabetes has been achieved through case-control studies in large cohorts of patients and control subjects. Most of the genes can be considered validated across various ethnic groups, including TCF7L2, SLC30A8, HHEX, CDKAL1, CDKN2A/B, IGF2BP2, and FTO. Leptin and ghrelin are two important peptides that regulate appetite and energy balance. It has been shown that common

polymorphisms in the genes encoding these peptides and their associated receptors are linked to body mass, various obesity markers, and metabolic disorders that are considered precursors to type 2 diabetes (Ghalandari H. et al., 2015). Adiponectin is a protein hormone secreted from adipocytes that increases insulin sensitivity and improves diabetic conditions by enhancing fatty acid oxidation and glucose uptake in skeletal muscles, while reducing glucose production in the liver. Low serum concentrations of adiponectin are associated with diabetes, obesity, insulin resistance, and metabolic syndrome. The adiponectin gene is located on chromosome 3q27. It is hypothesized that single nucleotide polymorphisms (SNPs) in the adiponectin gene (ADIPOQ) are associated with diabetes. SNPs in ADIPOQ help assess the relationship between common variants and adiponectin levels, as well as the risk of developing diabetes. Research findings suggest that rs1501299, rs2241766, rs266729, rs17366743, rs17300539, rs182052, rs822396, rs17846866, rs3774261, and rs822393 are substantially associated with type 2 diabetes, which is the predominant form of the disease. Additionally, rs2241766 and rs266729 are widely associated with gestational diabetes, a condition that develops in women during pregnancy. Consequently, not a specific single mutation, but a range of SNPs in the adiponectin gene may be a risk factor for diabetes globally. This study indicates that adiponectin plays a significant role in the pathogenesis of type 2 diabetes and gestational diabetes (Howlader M. et al., 2021).

SNP AND ONCOLOGY

According to some authors, the use of genetic polymorphism markers, particularly single nucleotide polymorphisms (SNPs), promises to provide a comprehensive tool for analyzing the human genome and identifying genes and genomic regions that contribute to carcinogenesis (Engle L.J. et al., 2006). The authors analyze various methods for determining SNPs, taking into account how each of these methodologies can be utilized to locate genes and genomic areas of interest for various types of cancer. It is suggested that, as with other multifactorial hereditary diseases, more than 90% of SNPs that could induce cancer are located in non-protein-coding regions of DNA. This fact complicates the deciphering of the functional impacts of SNPs on the pathology development process. Numerous mechanisms have been proposed to explain how non-coding genomic variants can influence human disease, such as disruptions in splicing, translation, or the stability of protein-coding genes. In addition to protein-coding genes, risk-associated SNPs identified by GWAS also affect key non-coding genes for miRNAs and lncRNAs (Liu S. et al., 2017). Recent studies have shown that among 996 cancer-associated SNPs identified in GWAS, 575 SNPs are located in fragments that positively regulate gene expression, and 758 SNPs in fragments with negative regulatory activity. Among these, 70 variants are regulatory variants for which two alleles provide different regulatory activity. Analysis of two regulatory variants—rs11055880, associated with breast cancer risk, and rs12142375, associated with leukemia risk—allowed for the determination of their regulatory activity concerning the expression of the ATF7IP and PDE4B genes. The value of these studies may lie in identifying regulatory variants associated with cancer susceptibility and studying their molecular functions, which could enhance the reliability of cancer risk assessment.

It is suggested that, in addition to the accumulation of inherited polymorphisms (including SNPs), the loss of one allele (loss of heterozygosity - LOH) through various possible mechanisms is significant in carcinogenesis processes. In turn, loss of heterozygosity is the most common molecular-genetic alteration observed in human cancer, predominantly involving the inactivation of tumor suppressor genes (Zheng H.-T. et al., 2005).

VNTR (SSR) - VARIATIONS IN THE NUMBER OF TANDEM REPEATS OF NUCLEOTIDES

The human genome contains numerous repetitive DNA sequences that vary in complexity, ranging from a single nucleotide to an entire gene. These repetitive sequences can be distributed widely or appear as simple tandem arrays. Arrays with repeat units up to 5 or 6 nucleotides are known as simple tandem repeats and are widely dispersed and highly polymorphic. Some simple tandem repeat groups, such as trinucleotide repeats, may experience an increase in copy number through the mutation process. Microsatellites, or simple sequence repeats (SSR), are prevalent in genomes and exhibit high polymorphism levels.

The human genome has many nucleotide sequences that occur repeatedly. These repetitive sequences vary from full genes (such as ribosomal RNA genes) to simple sequences composed of one or more base pairs. The organization of these sequences can range from widely dispersed copies of relatively long, complex sequences to tandem arrays of simple sequence compositions. Among the simplest and most common repeats are dinucleotide repeats, primarily featuring bases AC on one DNA strand and GT on the other. Other simple tandem repeats (STR) include mono-, tri-, tetra-, and pentanucleotide repeating units.

Genomic microsatellites (simple sequence repeats; SSR), iterations of nucleotide motifs ranging from 1-6 base pairs, have been identified in the genomes of every organism analyzed to date. SSRs constitute a substantial portion of non-coding DNA and are relatively rare in protein-coding regions. For instance, all observed mono-, di-, and tetranucleotide SSRs were located in non-coding regions in studies involving 54 plant species. All types of SSRs (from mono- to hexanucleotide repeats) were found in significant quantities in the non-coding regions of the genomes of invertebrates, plants, and primates (Li Y.-C. et al., 2002).

According to experts, the fact that all SSR types (excluding triplet and hexanucleotide repeats) occur far less frequently in protein-coding sequences compared to non-coding DNA may be attributed to negative selection against mutations in coding regions. However, triplet repeats associated with diseases primarily occur in the coding regions of the human genome. Some triplet repeats do not persist over time, even when forming part of protein-coding sequences, as long triplet repeats (e.g., CAG arrays) may become destabilized during meiosis (crossing over).

Most SSRs (48-67%) identified in various species are dinucleotides. However, in primates, mononucleotide SSRs (mostly poly(A/T) tracks) represent the most common SSR class. Dinucleotide SSRs can be found in introns, such as in the *Adh1* gene intron of mice, which contains (TA)₁₄ (TG)₈ and (TA)₁₉, and in the *IL-5* gene intron, which contains (AT)₁₇. An increase in di- or tetranucleotide SSR sizes within introns may lead to the disruption of the original protein or the formation of new genes through frameshift mutations.

When examining SSR occurrence in various functional genome regions, most demonstrate significantly higher densities in non-coding areas. The exceptions to this pattern are trimers and hexamers, which are nearly twice as frequent in exons compared to introns and intergenic regions. Their high frequency in coding regions can be explained because they do not alter reading frames or affect the properties of the genetic code. Consequently, this mutation type does not exert a substantial negative impact compared to other SSR types. Positive selection of these in exons also suggests a potential function for the repeats (Jasinska A., Krzyzosiak W.J., 2004).

Loci with longer repeat units appear subject to stronger selection due to size differences, especially in genome areas with high recombination rates. These findings also indicate that differences in SSR frequency between coding and non-coding genome regions arise from specific selection against frameshift mutations in coding areas caused by length changes in non-triplet repeats. Nevertheless, 14% of all proteins contain repetitive sequences, with eukaryotic repeats observed to be threefold higher than in prokaryotes. Families of prokaryotic and

eukaryotic repeats are grouped within non-homologous proteins, possibly indicating that repetitive sequences emerged after these two kingdoms diverged. Eukaryotes with more repeats may hold an evolutionary advantage in adapting more rapidly to new conditions. It is suggested that the differences between coding and non-coding SSR frequencies result from selective pressure against frameshift mutations in coding regions.

MUTATIONAL MECHANISMS OF SSR VARIATION

The genomic abundance of Simple Sequence Repeats (SSR) and their various functions and effects (whether presumed or conclusively established) are closely associated with mutation frequency, as the mutation rate of SSRs (ranging from 10^{-2} to 10^{-6} events per locus per generation) is significantly higher compared to point mutation rates in coding gene loci. Although the mutation process appears to exhibit distinct variations across species, repeat types, loci, alleles, age, and gender, instability primarily manifests in the form of changes in the number of SSR repeats. Two mutational mechanisms can be utilized to explain such high mutation rates. The first mechanism is related to DNA slippage during DNA replication. The second involves recombination between DNA strands. The effectiveness of these two mechanisms might hypothetically depend on environmental conditions. Various factors have been identified to influence the mutation rate at SSR loci, including the repeat motif, allele size, chromosome location, GC content in flanking DNA, cell division type (mitotic versus meiotic), gender, and genotype (e.g., mutations in MMR genes). The balance in the effectiveness of these two mechanisms may be influenced by environmental conditions.

SSR AND TRANSCRIPTION

Numerous studies suggest that SSRs located in promoter regions can affect gene activity. It has been found that (TC) n repeating nucleotide sequences in promoter areas serve as a transcriptional element for the hsp26 heat shock protein gene in various species. Deletion of different di-, tri-, and tetra-SSR repeats significantly altered transcriptional activity. For instance, the transcriptional activity of promoters was sharply reduced by removing (TCCC) n repeats from the promoter regions of the TGF- β 3 gene in a CAT expression system. Conversely, a (GT) n repeat may enhance gene activity, particularly if the GT repeat is situated closer to the promoter sequences.

Only a small portion of the total RNA transcribed in human cells becomes mature mRNA. The ensemble of mRNA constitutes the human transcriptome, which can vary depending on the stage of ontogeny, physiological state, and pathological processes. A small fraction of different repeat sequences, which comprise more than half of the human genome, is retained in mature transcripts and influences their function. In addition to mRNA, the transcriptome also includes various non-coding RNAs that perform either structural or regulatory functions within cells. Consequently, there are thousands of different transcriptomes across hundreds of different cell types and organs, each in various physiological and pathological states. The diversity of transcripts compared to genes increases by approximately 50% due to alternative splicing, numerous antisense transcripts are synthesized, and all non-coding RNAs, including the recently recognized microRNA family, significantly contribute to the human transcript pool.

The analysis of the presence of triplet repeats in mRNA has shown that at least 718 repeats were identified in 619 mRNAs (Jasinska A., Krzyzosiak W.J., 2004). Short Tandem Repeats (STR) frequently appear in many transcripts; however, in certain cases, pathological events occur when a critical repeat length is reached. This phenomenon, for instance, is observed in various neurological disorders. The pathological effects of these repeats are triggered by

transcripts of mutant RNA and/or the mutant proteins encoded by these RNAs, which are determined by the localization of expanded repeats in non-coding or coding regions. It has been demonstrated that RNA structures formed by these mutant repetitive RNA tracts exert toxic effects on cells (Ciesiolka A. et al., 2017).

SSR AND TRANSLATION

Numerous studies indicate that SSRs can influence gene translation. For instance, the reduction in the number of (CA)_n repeats enhances mRNA translation in *Escherichia coli* cells. Moderate expansion of CGG repeats can significantly diminish the translation of the CAT gene (pSVsCAT) when human CGG repeats are inserted. The distribution of tetramers AGCT in the genomes of *E. coli* and *Bacillus subtilis* predicts translational frameshift in several genes (infB, ACEF/PDHC, eno, RPLI, OmpA, OMF, and TOLA). The binding of the CUG repeat (CUGBP1) interacts with the 5' region of C/ebpβ mRNA and regulates the translation of C/EBPBETA isoforms. Triplet repeats UGG have a strong inhibitory effect on mRNA translation in *E. coli*, which has been demonstrated in the reporter gene CAT.

MEDICAL SIGNIFICANCE OF SSR

Approximately 15% of sporadic cases of colorectal cancer, as well as other malignancies, exhibit SSR instability. Progressive accumulation of SSR instability may also contribute to the development of gastric cancer. Examples of SSR alterations accompanying various types of human cancer can be divided into two subtypes: Type A, which shows relatively minor changes within six base pairs, and Type B, which shows abrupt changes involving eight base pairs. In certain types of cancer, a link has been proposed between SSR Type B instability and familial predisposition. It has been shown that 14 neurological disorders result from expansions of unstable trinucleotide repeats. Such SSR triplet disorders include cases with changes either in non-coding (untranslated) or coding (exon) sequences. Expanded repeats can confer genetic instability, likely contributing to transcription suppression. The human FMR1 repeat array (CGG)_n may exhibit genetic instability characterized by progressive expansion over generations, leading to gene repression and the development of Fragile X syndrome. Myotonic dystrophy, an autosomal dominant neurological disorder, is caused by the expansion of CTG repeats in the DMPK locus, where affected individuals have $n \geq 50$ repeats (CTG). Abnormal expansion (≥ 39) of (CAG) repeats, which results in polyglutamine expansions, causes Machado-Joseph disease. It has also been established that the length, rather than the specific allele (CA)_n of the aldose reductase gene repeat, is associated with diabetic retinopathy, and a polymorphism of the (CA)_n allele in the first intron of the human interferon-gamma gene is associated with lung graft fibrosis. For instance, the (CCTTT)₁₄ allele in the promoter of the NOS2A gene significantly limited diabetic retinopathy.

Furthermore, it has been reported that simple repeat expansions may cause certain types of cancer. This can occur through alteration of tumor suppressor gene expression and changes in the properties of encoded proteins. Demonstrated meiotic instability may link this mutation mechanism to familial cancer syndromes. The established instability of STR in several loci in hereditary non-polyposis colorectal cancer suggests that sequence instability may be a factor in cancer progression. Further identification of candidate genes containing triplet repeats should facilitate testing the hypothesis that unstable simple repeat sequences may cause cancer.

RECOMMENDED LITERATURE FOR CHAPTER 2. GENE POLYMORPHISM

1. Guo Y., Jamison D. C. The distribution of SNPs in human gene regulatory regions. *BMC Genomics*, 2005, Vol. 6, p. 140. doi:10.1186/1471-2164-6-140
2. Kim B.-C., Kim W.-Y., Park D. SNP@Promoter: a database of human SNPs (Single Nucleotide Polymorphisms) within the putative promoter regions. *BMC Bioinformatics*, 2008, Vol. 9(Suppl 1), p. S2. doi:10.1186/1471-2105-9-S1-S2
3. Vignal A., Milan D., Magali SanCristobal, Eggen A. A review on SNP and other types of molecular markers and their use in animal genetics. *Genet. Sel. Evol.*, 2002, Vol. 34, pp. 275–305. DOI: 10.1051/gse:2002009
4. Johnson A.D. SNP bioinformatics: a comprehensive review of resources. *Circ Cardiovasc Genet.*, 2009, Vol. 2, No. 5, pp. 530–536. doi:10.1161/CIRCGENETICS.109.872010
5. Bin Alwi Z. The use of SNPs in pharmacogenomics studies. *Malaysian Journal of Medical Sciences*, 2005, Vol. 12, No. 2, pp. 4-12.
6. Chen C., Chang I.-S., Hsiung C.A., Wasserman W.W. On the identification of potential regulatory variants within genome-wide association candidate SNP sets. *BMC Medical Genomics*, 2014, Vol. 7, p. 34. <http://www.biomedcentral.com/1755-8794/7/34>
7. Xu H., Gregory S.G., Hauser E.R. et al. SNPselector: a web tool for selecting SNPs for genetic association studies. *Bioinformatics*, 2005, Vol. 21, N22, pp. 4181–4186. doi:10.1093/bioinformatics/bti682
8. Ghalandari H., Hosseini-Esfahani F., Mirmiran P. The Association of Polymorphisms in Leptin/Leptin Receptor Genes and Ghrelin/Ghrelin Receptor Genes with Overweight/Obesity and the Related Metabolic Disturbances: A Review. *Int. J. Endocrinol. Metab.*, 2015, Vol. 13, No. 3, p. e19073. DOI: 10.5812/ijem.19073v2
9. Howlader M., Sultana I., Akter F., Hossain M. Adiponectin gene polymorphisms associated with diabetes mellitus: A descriptive review. *Heliyon*, 2021, Vol. 7, C8, p. e07851. doi: 10.1016/j.heliyon.2021.e07851
10. Staiger H., Machicao F., Fritsche A., Haring H.-U. Pathomechanisms of Type 2 Diabetes Genes. *Endocrine Reviews*, 2009, Vol. 30, No. 6, pp. 557–585. doi: 10.1210/er.2009-0017
11. Liu S., Liu Y., Zhang Q. et al. Systematic identification of regulatory variants associated with cancer risk. *Genome Biology*, 2017, Vol. 18, p. 194. DOI 10.1186/s13059-017-1322-z
12. Zheng H.-T., Peng Z.-H., Li S., He L. Loss of heterozygosity analyzed by single nucleotide polymorphism array in cancer. *World Journal of Gastroenterology*, 2005, Vol. 11, No. 43, pp. 6740-6744. <http://www.wjgnet.com/1007-9327/11/6740.asp>
13. Engle L.J., Simpson C.L., Landers J.E. Using high-throughput SNP technologies to study cancer. *Oncogene*, 2006, Vol. 25, pp. 1594–1601. doi:10.1038/sj.onc.1209368
14. Li Y.-C., Korol A.B., Fahima T. et al. Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review. *Molecular Ecology*, 2002, Vol. 11, pp. 2453–2465. DOI: 10.1046/j.1365-294x.2002.01643.x
15. Jasinska A., Krzyzosiak W.J. Repetitive sequences that shape the human transcriptome. *FEBS Letters*, 2004, Vol. 567, pp. 136–141. doi:10.1016/j.febslet.2004.03.109
16. Ciesiolka A., Jazurek M., Drazkowska K., Krzyzosiak W.J. Structural Characteristics of Simple RNA Repeats Associated with Disease and their Deleterious Protein Interactions. *Front. Cell. Neurosci.*, 2017, Vol. 11, p. 97. doi: 10.3389/fncel.2017.00097
17. Panzer S., Kuhl D.P.A., Caskey C.T. Unstable Triplet Repeat Sequences: A Source of Cancer Mutations? *STEM CELLS*, 1995, Vol. 13, pp. 146-157.

CHAPTER 3. POLYGENIC (MULTIFACTORIAL) DISEASES. TYPE 2 DIABETES AS AN EXAMPLE OF POLYGENIC DISEASES

INTRODUCTION

The alteration in gene structure (polymorphism) results in a hereditary predisposition to various common chronic diseases. The identification of thousands of genes in the human genome, elucidation of the molecular mechanisms underlying hereditary and polygenic (multifactorial) diseases, and the establishment of the genetic uniqueness of each individual are significant achievements that laid the scientific groundwork for molecular medicine. Molecular medicine distinguishes itself from traditional (classical) medicine in two key aspects: individualization (each human genome is unique) and its preventative approach (genomes can be analyzed before the onset of disease). One of the foundations of molecular medicine, predictive medicine, involves analyzing allele variants of genes that predispose individuals to polygenic diseases.

Current data indicate that the genetic burden in populations of Europe and Russia is approximately at the same level, around 5.5%. Of this, about 1% is attributable to monogenic diseases (pathological mutations of a single gene). Approximately 0.5% are due to chromosomal aberrations, and 3-4% correspond to polygenic (multifactorial) diseases with a strong hereditary factor. Therefore, among hereditary diseases, polygenic diseases exhibit the most widespread prevalence. Typically, the pathogenesis of polygenic diseases involves multiple genes rather than just one. For instance, two genes (BRCA1, BRCA2) are known to contribute to hereditary forms of breast cancer. Additionally, at least three genes (PS1, PS2, AAP) are involved in familial forms of Alzheimer's disease, and about 600 genes are considered candidate genes in the diagnosis of diabetes. A key feature of this group of hereditary diseases is that the phenotypic manifestation of gene mutations responsible for predisposition always depends on the influence of external precipitating factors on the organism. For this reason, this crucial group of medical hereditary diseases is known as multifactorial hereditary diseases. Multifactorial diseases have a complex genetic architecture represented by a broad spectrum of allele variants. These genes vary in origin (germline or somatic, de novo) and are present in the population with varying frequency (ranging from widespread to rare, with mutation-level frequencies). These genes can contribute different levels of effect on the risk of developing a disease in an individual.

Obesity, osteoporosis, endometriosis, myocardial infarction, most tumors, psychiatric, and cardiovascular diseases are included in the group of polygenic hereditary diseases based on the interaction of multiple genes with adverse environmental factors. This implies that the pathogenesis of multifactorial diseases is determined by gene mutations (gene polymorphism) across many genes. Consequently, these diseases will phenotypically manifest only with the combination of hereditary predisposition (hereditary risk of developing the disease) and corresponding adverse environmental factors. Based on the facts presented above, it can be concluded that laboratory analysis (e.g., PCR method) of genetic markers and possible phenotypic correction of their functions can significantly reduce the risk of developing multifactorial diseases (such as atherosclerosis, diabetes, myocardial infarction, hypertension, and others). Presymptomatic identification of individuals in high-risk groups for multifactorial pathology and timely prevention are the main directions of predictive medicine.

In conclusion, multifactorial diseases result from the unfavorable combination of genetic factors and environmental influences. Therefore, mutant genes (alleles) that are compatible with birth and life but contribute to the development of multifactorial diseases under certain unfavorable conditions are termed "predisposition genes." In addition to genes, various risk

factors accumulate with age, contributing to the pathogenesis of this group of diseases. Cardiovascular diseases, chronic kidney disease, chronic lung diseases, liver pathologies, and cancer are the most common multifactorial diseases and pose a significant burden on the healthcare system.

Multifactorial diseases, also referred to as polygenic or complex hereditary diseases, are influenced by numerous genetic and environmental factors. It is important to note that the alleles involved in the pathological process are not typically pathological themselves; rather, they increase an individual's susceptibility to certain diseases. Confirming the role of susceptibility genes in the pathogenesis of polygenic diseases and determining the allelic composition of these genes in each population are highly complex tasks. Generally, these studies follow an algorithm consisting of three main stages: Proving that the disease is familial, Demonstrating that this familial tendency is due to genetic factors, Identifying the genes involved in the pathogenesis.

Mapping and identifying susceptibility genes for multifactorial diseases are typically conducted either through studying families with an increased risk of developing the disease or by analyzing groups of unrelated patients with the condition compared to healthy individuals (case/control comparison). The primary focus of the search is often the mapping of single nucleotide polymorphisms (SNPs) located both within and outside structural genes that are most closely correlated with the disease.

This field of research requires significant investments of time and resources. For instance, in large-scale studies seeking susceptibility genes for prostate cancer, DNA samples from 1,150 patients and 1,150 control samples were used. The research began with the analysis of 500,000 SNPs, which were later reduced to 200 SNPs. This extensive work led to the identification of 15 candidate genes, the study of which continues.

Over the past decade, genetic loci (primarily single nucleotide polymorphisms - SNPs) associated with common complex diseases have been identified.

THE ROLE OF GENE NETWORKS IN THE PATHOGENESIS OF MULTIFACTORIAL DISEASES

According to Professor N. A. Kolchanov, a gene network is a group of genes that function in a coordinated manner to form the phenotypic traits of an organism (molecular, biochemical, physiological). On one hand, each of these local gene networks acts as discrete and autonomous communities of genes and the products of their expression. On the other hand, local gene networks are integrated into a single global network of the organism. Thanks to advancements in bioinformatics, numerous gene networks have been identified, responsible for regulating glucose levels in the body, synthesizing steroid hormones, regulating adipocyte function (fat tissue cells), maintaining erythropoiesis mechanisms, and oxidative stress gene networks that regulate organ and tissue responses to reactive oxygen and nitrogen species.

As previously mentioned, all gene networks closely interact with each other and are integrated into the organism's unified gene network. This integration can be horizontal or vertical (hierarchical). Horizontal integration refers to the integration of gene networks at the same level. An example of horizontal integration is the paracrine integration of insulin and glucagon gene networks that together regulate blood glucose levels. A vertical integration example is the gene network controlling steroid hormone synthesis. The hierarchical nature of this vertical network is conditioned by three levels of hierarchy: the hypothalamus, pituitary gland, and peripheral endocrine glands.

It is important to note that gene networks are also under external control. Gene networks are subject to regulation by specific integrator mechanisms. Specific individual gene networks

can act as such integrators. An example of such an integrator is the gene network regulating the circadian (daily) rhythm, which perceives environmental signals and manages the rhythm of numerous gene networks based on these signals, ranging from the regulation of metabolic processes to mitotic cycles. Another integration mechanism for gene networks can be various metabolites (glucose, reactive oxygen species, etc.) and neurohumoral factors (hormones, transcription factors). From this, we can conclude that the integration of gene networks is a crucial mechanism for providing an adaptive response of the entire organism and its life-support systems to environmental influences.

Analyzing the above facts, we can conclude that the efficiency of local gene networks, as well as the organism's adaptability to constantly changing environmental conditions, depends on the unique characteristics of the genes within these networks. Each signal entering the organism elicits a response from a particular gene network, which, through corresponding integrators, transmits the signal to other gene networks. Malfunctions in gene networks, caused by gene mutations or their functional inadequacies due to allelic variants, can significantly distort the work of the entire integral gene network, affecting the whole organism. The pathogenesis of each disease involves various functionally interconnected genes within a specific local gene network. Alongside major genes initiating a disease, numerous modifier genes also exist, whose phenotypic effects are largely determined by environmental factors. Identifying such genes, understanding the nature of functional disorders at the level of local and integral gene networks, and gene-gene interactions in multifactorial diseases is a significant practical task for bioinformatics.

It is evident that due to gene polymorphism and the individual uniqueness of allelic composition, the gene networks of each person function differently. Since the genes forming the basis of each gene network are polymorphic, their allelic variants are functionally distinct among different individuals. As a result of gene polymorphism, each person has a unique biochemical profile, an individual norm of biochemical reactions, and unique adaptive capabilities in response to adverse environmental factors. Therefore, individual reactions to external influences, infections, injuries, geophysical conditions, food products, toxins, and medications are shaped by gene polymorphism (individual genetic composition).

This is directly related to the pathogenesis of multifactorial diseases. Modern medical literature has frequently documented cases of familial (hereditary) predispositions to atherosclerosis, diabetes, heart diseases, lung, kidney, liver diseases, mental disorders, and a high tendency towards oncological and allergic diseases. The foundation of these hereditary diseases is genes. Since an individual's phenotype (including pathological) is an expression of the specific features of their genotype, realized with the participation of local and integral gene networks within specific environmental conditions. The allele set in a human genotype is unique, but gene polymorphism is not always neutral. Sometimes the presence of certain gene variants can lead to the formation of proteins with altered physicochemical properties, accompanied by disruptions in protein functional activity. As previously noted, single nucleotide polymorphisms (SNPs) in gene-coding regions can impact not only the primary structure of the protein but also its tertiary structure, thus affecting its functional activity. Under certain environmental conditions (geophysical factor, diet, etc.), the effects determined by gene polymorphism can either promote or hinder the development of various diseases. Genes whose allelic variants, in the presence of specific environmental factors, predispose to certain diseases are termed susceptibility genes. Thus, susceptibility genes are mutant genes (alleles) that are compatible with birth and life in the postnatal period, but under certain unfavorable conditions, contribute to the development of specific diseases.

It is possible that the boundary between monogenic and polygenic diseases is not as significant, and there is a close relationship between these groups of hereditary diseases. This

opinion is based on the fact that "common" genes for Mendelian and multifactorial diseases are closely linked at the level of protein-protein interactions and have higher expression levels in cells and tissues than genes "only for multifactorial diseases." Additionally, "common" genes for Mendelian and multifactorial diseases have a large number of transcripts and encode longer proteins. In association studies, variants in the "common" genes for Mendelian and multifactorial diseases have higher risk values than the genes "only for multifactorial diseases." It is suggested that studies of multifactorial diseases should primarily focus on examining gene variants caused by mutations in the promoter regions of Mendelian disease genes or in non-coding regions of the genome adjacent to these promoters.

THE ROLE OF MULTIFACTORIAL DISEASES OF THE HEART, KIDNEYS, LUNGS, AND LIVER IN CARCINOGENESIS

Polygenic disease is a complex phenomenon. Moreover, the presence of such a disease may lead to other pathologies, such as oncopathologies. An additional challenge arises from the fact that the simultaneous manifestation of two or more diseases from this group in a single patient presents various clinical problems. Additionally, therapeutic correction of one disease may be incompatible with the treatment of another. Furthermore, the increased risk of developing oncological diseases should be considered among patients already diagnosed with multifactorial diseases. In other words, the widespread prevalence of multifactorial diseases may explain the increased risk of spreading oncological diseases. According to the Population Division of the United Nations Department of Economic and Social Affairs, as of 2019, it is expected that by 2050, the global population over the age of 65 will exceed 1.5 billion. Therefore, considering that age is one of the key risk factors for the manifestation of polygenic diseases, it can be assumed that there will be an increased manifestation of multifactorial diseases, including those contributing to the development of oncopathologies.

THE CONNECTION BETWEEN CANCER AND HEART FAILURE

On one side, cardiotoxicity induced by chemotherapy for oncopathologies is becoming increasingly evident, as treatment based on cytostatics can lead to irreversible myocardial damage. On the other side, there is an argument that heart failure may be an oncogenic condition. Indeed, population studies have shown that the prevalence of oncological diseases is higher among patients with heart failure compared to the general population. The risk of developing cancer in individuals with heart failure may be 24-68% higher compared to healthy individuals. Specifically, a correlation has been established between heart failure and the progression of colorectal cancer. This correlation is most likely independent of changes in hemodynamic parameters and is determined by physiologically active substances released by the damaged myocardium. Chronic low-grade systemic inflammation is considered a primary pathophysiological process linking heart failure with the multistep process of carcinogenesis. In addition to colorectal cancer, heart failure may contribute to the development of breast and lung cancer. It should be noted that the pathogenesis of both heart failure and cancer is based on common pathways, such as inflammation, oxidative stress, and somatic mutations. Therefore, the existence of shared fundamental mechanisms may explain the coexistence of these two nosologies.

THE CONNECTION BETWEEN ONCOLOGY AND CHRONIC KIDNEY FAILURE

In relation to the risk of oncological diseases in the context of chronic kidney failure, it should be noted that, firstly, the risk of cancer is observed even in patients with moderate kidney function impairments. Secondly, this trend was only observed in men. It is noteworthy that regardless of smoking and the age of patients, a decline in the glomerular filtration rate (GFR) by 10 ml/min/1.73 m² was associated with a 29% increase in cancer risk. This association was even more significant when evaluating the risk of malignancies in the lungs, colon, kidneys, and urinary tract. It can be hypothesized that a decrease in GFR results in higher levels of carcinogenic compounds and nitrogen-containing substances in the plasma of patients with chronic kidney failure. Moreover, it is known that carcinogenic compounds, such as 2-amino-6-methyldipyrido[1,2-a:3',2'-d]imidazole (Glu-P-1) and 2-aminodipyrido[1,2-a:3',2'-d]imidazole (Glu-P-2), bind to DNA and cause mutations. The risk level and location of oncological diseases in the progression of kidney failure may depend on the severity of the underlying disease. Indeed, among dialysis patients, the risk of cancer in the urinary tract, endocrine system, and digestive tract was significantly increased.

Given these findings, some authors draw attention to neutrophil gelatinase-associated lipocalin (NGAL). This is a 25 kDa protein primarily secreted by immune cells such as neutrophils, macrophages, and dendritic cells, and its production is stimulated in response to inflammation. NGAL is primarily known as a biomarker for acute kidney injury and is released following damage to the nephron tubules, as well as during renal regeneration processes. NGAL levels are also elevated in patients with chronic kidney disease. The presence of NGAL enhances inflammatory responses, and its level increases with arterial hypertension, obesity, diabetes, and metabolic complications such as insulin resistance. It also participates in carcinogenesis.

Regarding the mechanisms of carcinogenesis induction, one of the main processes is chronic inflammation. On one hand, a chronic inflammation site is a source of reactive oxygen species, which contribute to various mutations—the fundamental mechanism of carcinogenesis. In addition, the production of proinflammatory cytokines, such as interleukin-1 β (IL-1 β), interleukin-6 (IL-6), and tumor necrosis factor-alpha, enables malignant cells to evade the body's immune surveillance, stimulating angiogenesis, tumor growth, and invasiveness.

DEVELOPMENT OF LUNG CANCER IN CHRONIC OBSTRUCTIVE PULMONARY DISEASE (COPD)

Chronic Obstructive Pulmonary Disease (COPD) and lung cancer are two major public health challenges and causes of morbidity and mortality. Both diseases share common risk factors, such as exposure to smoking and genetic predisposition. Currently, it has been established that certain families of genes involved in proteases, detoxifying enzymes, and inflammatory cytokines may play a significant role in the pathogenesis of COPD. Additionally, some candidate genes, such as Glycophorin A (GYPA) and Hedgehog Interacting Protein (HHIP), have been identified and may be linked to susceptibility to both COPD and lung cancer. COPD is characterized by chronic inflammation of the lungs, suggesting that inflammation may be the connecting link between COPD and lung cancer, although the exact underlying mechanism remains unknown. Inflammation mediators contribute to the development of inflammation-induced cancer by promoting cell proliferation, resistance to apoptosis, invasion, and metastasis.

Another significant mechanism in the pathogenesis of COPD is oxidative stress, which contributes to cell proliferation in lung cancer. Mitochondrial dysfunction within the lung parenchyma may impact the pathogenesis of COPD, including oxidative stress imbalance, potentially enhancing chronic inflammation and contributing to carcinogenesis.

PROGRESSION OF METABOLIC DYSFUNCTION-ASSOCIATED FATTY LIVER DISEASE AND ONCOLOGICAL CONDITIONS

Metabolic Dysfunction-Associated Fatty Liver Disease (MAFLD), which occurs in individuals who do not abuse alcohol, is a specific case of steatogepatosis. This pathology is driven by insulin resistance and metabolic syndrome. The metabolic disturbances observed in MAFLD patients are the most significant risk factors for cancer, primarily liver oncopathology - hepatocellular carcinoma (HCC). Additionally, the presence of MAFLD in patients with chronic Hepatitis B further increases the risk of developing HCC. In recent years, MAFLD has become the fastest-growing cause of HCC in the United States. Obesity and type 2 diabetes play crucial roles in the development and progression of HCC in MAFLD. Diabetes may influence carcinogenesis through several mechanisms, including hyperinsulinemia, hyperglycemia, oxidative stress, and chronic inflammation. In turn, chronic inflammation stimulates the secretion of IL-6 and TNF- α , which are associated with oncogenesis. Additionally, steatosis may lead to mitochondrial dysfunction and endoplasmic reticulum (ER) stress, resulting in increased production of reactive oxygen species. Consequently, oxidative stress leads to DNA damage and activation of oncogenic signaling pathways such as NF- κ B.

It is essential to emphasize that MAFLD is not only associated with an increased risk of HCC but also with several extrahepatic cancers, including colorectal cancer, kidney cancer, thyroid cancer, and breast cancer. Notably, the risk of kidney and thyroid cancer remained significant even after correcting for metabolic processes. The connection between MAFLD and colorectal cancer has garnered increasing attention in recent years. Population studies have identified MAFLD as an important risk factor for colorectal cancer.

THE POLYGENIC NATURE OF MULTIFACTORIAL DISEASES: A CASE STUDY OF TYPE 2 DIABETES

Diabetes mellitus is classified among a group of metabolic diseases characterized by chronic elevation of blood plasma glucose levels. According to various estimates, more than 90% of all diabetes cases are attributed to type 2 diabetes (T2D), previously known as "adult-onset diabetes" or "non-insulin-dependent diabetes." The incidence of type 2 diabetes has reached the proportions of a global epidemic. Presently, approximately 537 million adults aged 20–79 are affected by diabetes, and the total number of individuals diagnosed with the condition is continually rising. Projections indicate that the prevalence of T2D will increase to 643 million by 2030 and 783 million by 2045. The pathogenic mechanisms of T2D include insulin resistance, impaired insulin secretion, and increased hepatic glucose production. Complications may manifest as multi-organ failure, including nephropathy, neuropathy, retinopathy, periodontal disease, and accelerated cardiovascular disease development. In 2010, in the United States, 1.9 million people aged 20 and older were newly diagnosed with type 2 diabetes. T2D is becoming increasingly prevalent in many regions across the globe. The prevalence of T2D among adolescents and young people aged 10 and older is 8.5 per 100,000 population, whereas, among children under 10, the prevalence is 0.4 per 100,000 people. In the United States, 26.9% of all individuals aged 65 and older are affected by this disease. Projections suggest that factors such as population aging and the rising proportion of overweight individuals will further increase the prevalence of T2D. Meanwhile, lifestyle changes, such as weight loss, dietary optimization, and increased physical activity, can help prevent or significantly delay diabetes onset.

Furthermore, accumulating evidence indicates that genetic factors play a significant role in the pathogenesis of T2D. Notably, there is a widely recognized correlation between disease risk and belonging to a certain ethnic group. Moreover, the disease often has a familial aspect. It is important to emphasize that a family history of diabetes is a major risk factor for the disease and is often included in tools for quantitative diabetes risk assessment. At the age of 80, the risk of developing the disease is approximately 3.5 times higher if close relatives have been diagnosed with T2D compared to individuals whose relatives have not had the disease.

T2D is a polygenic disease with a hereditary risk ranging from 30% to 70%. T2D qualifies as a multifactorial (polygenic) disease. Both genetic and environmental factors influence the manifestation of T2D symptoms. The degree of contribution from genetic factors and environmental factors may vary among individuals. However, the disease develops when the combined effects of both factors exceed a certain threshold. The role of genetic factors is undisputed. According to twin and genealogical methods in medical genetics, the heritability of T2D complications is approximately 40%. Additionally, the frequency of susceptibility gene polymorphisms varies between patient groups and control groups. Population-specific features related to the predisposition genes for T2D have been identified. While excess body weight precedes the development of type 2 diabetes in Europeans, in East Asian populations, diabetes develops at lower body weights compared to Europeans, due to differing levels of insulin secretion.

TYPE 2 DIABETES MELLITUS AS AN EXAMPLE OF A POLYGENIC DISEASE: PREDISPOSITION GENES

The primary tool for studying pathological variants of genes predisposing to Type 2 Diabetes (T2D) is next-generation sequencing. Most of these genes are associated with the regulation of the functional state of pancreatic β -cells. Genetic polymorphisms affecting key proteins involved in glucose metabolism and insulin secretion may also influence predisposition to T2D. Polygenic genomic assessments for T2D can predict the risk of developing diabetes. Some authors argue that identifying genes predisposed to T2D for preventive measures is too cumbersome, costly, and redundant. An additional significant challenge is selecting marker genes for a specific population. According to them, the most feasible and effective screening is based on plasma glucose levels, glycated hemoglobin, blood pressure, and body mass index. However, the volume of research on the genetic aspects of T2D predisposition is steadily increasing worldwide. This trend is driven by the global scale of the problem, the rapid increase in the number of T2D patients, the growing magnitude of its impact, and the urgent need for effective methods for early diagnosis and treatment. It is important to emphasize that the primary direction in the search for T2D predisposition genes involves gene mutations, predominantly SNPs.

One of the first widely accepted genetic markers for T2D is the TCF7L2 gene, located on the long arm of chromosome 10. Currently, over 600 loci susceptible to T2D have been identified among European and Asian populations. The TCF7L2 gene has shown the same diagnostic value in both European and East Asian populations. Similarly, the KCNQ1 gene—identified as a susceptibility locus with a relatively high risk ratio of 1.4 times for T2D in the Japanese population—is also highly reproducible in European populations. In contrast, the genes DGKD and ZNF257 have diagnostic value only in East Asian populations.

Contemporary literature places significant emphasis on the prevalence of T2D predisposition genes across different populations. As an example, consider the dynamics of obesity and T2D levels in Middle Eastern countries. Several patterns emerge: firstly, there is a rapid increase in the number of individuals with excess body weight—from threshold levels

(less than 5% of the population) in the mid-1970s to more than 20% today. In Kuwait, the proportion of individuals with excess body weight reaches nearly 40%. Secondly, in most Arab countries, marriages between cousins can reach up to 30%, making the Arab genome architecture unique in terms of susceptibility to various diseases, including both Mendelian and polygenic diseases. Thirdly, among approximately 100 gene variants associated with excess body weight, two were described as unique to Arabs, as they had not been previously recorded in other global populations. Meanwhile, 19 alleles were registered as distinct markers of excess body weight in Arabs but did not have pathogenic effects in non-Arab populations.

Attention is also drawn to the existence of gender specificity in genetic markers. For instance, the ALDH2 gene is closely associated with the risk of developing Type 2 diabetes in men but not in women.

Below, we provide an overview of the most relevant SNPs associated with the development of T2D in humans.

The TCF7L2 Gene (rs7903146): This gene encodes a transcription factor playing a crucial role in the Wnt signaling pathway. The protein is involved in blood glucose homeostasis. Genetic variants of this gene are associated with an increased risk of type 2 diabetes. Several transcript variants for this gene have been identified, encoding various isoforms. Dysregulation of the Wnt signaling pathway plays a key role in the development of insulin resistance. The rs7903146 genotype is a well-studied risk marker for type 2 diabetes in European and Asian populations. Patients with type 2 diabetes are more likely to carry the rs7903146 variant across different populations.

The KLF14 Gene (rs4731702): This non-intronic gene encodes a protein that is part of the Kruppel-like family of transcription factors. The encoded protein functions as a transcription corepressor and is induced by transforming growth factor-beta (TGF-beta) to suppress the expression of the TGF-beta receptor II gene. This gene regulates gene networks involved in lipid metabolism. High fasting insulin levels and insulin resistance are associated with the rs4731702 risk allele. Moreover, it may influence the expression of genes that determine body mass index.

The KCNQ1 Gene (rs2237895): This gene encodes the alpha subunit of the voltage-gated potassium channel Kv7.1, a member of the Kv channel superfamily, located on chromosome 11p15.5. The protein encoded by the KCNQ1 gene serves as the pore-forming alpha subunit of the KCNQ1/KCNE1, KCNQ1/KCNE2, and KCNQ1/KCNE3 potassium channels. KCNQ1 is expressed not only in the myocardium but also in adipose tissue, the pancreas, and the brain. In pancreatic islets, it plays an important role in glucose homeostasis, functioning as a regulator of insulin secretion. The intronic rs2237895 in KCNQ1 is associated with a reduced risk of abdominal obesity in patients with type 2 diabetes.

The FTO Gene: This gene encodes a nuclear protein from the non-heme iron and 2-oxoglutarate-dependent oxygenase superfamily, related to AlkB. Like other non-heme iron enzymes of this family, it regulates oxidative demethylation processes of DNA. High expression levels of this gene are observed in the nervous system, adrenal glands, and myocardium. A strong association with body mass index, obesity risk, and type 2 diabetes has been identified. It has been shown that rs1558902 in FTO is closely correlated with the incidence of type 2 diabetes in individuals, with high correlation values remaining significant even after adjustments for current age and body mass index.

The HNF4A Gene: This gene is a member of the steroid hormone receptor superfamily, primarily expressed in the kidneys, liver, pancreas (including β -cells), and small intestine, and it influences lipid metabolism and transport. It also participates in the regulation of liver function and hepatocyte differentiation. HNF4A is essential for glucose metabolism, as well as for insulin expression and secretion, while in the liver, HNF4A is required for hepatic

gluconeogenesis. Currently, it is known that non-coding variants of the HNF4A gene, rs6017317 and rs4812829, as well as the coding missense variant rs1800961, are risk factors for type 2 diabetes. Pathogenic variants of the HNF4A gene may contribute to insulin secretion disorders, as observed in patients with type 2 diabetes and MODY1. The MODY1 form of diabetes is a type of maturity-onset diabetes of the young, characterized by an early age of onset (usually in adolescence or childhood), dominant inheritance, and β -cell function defects. MODY1 arises from mutations in the HNF4A transcription factor gene.

The IGF2BP2 gene encodes a protein that binds to the 5' UTR of the mRNA of insulin-like growth factor 2 (IGF2) and regulates its translation. This gene is characterized by high expression levels in the islets of the pancreas. In adipose tissue and the pancreas, IGF2BP2 can decrease the expression of IGF2, which plays a crucial role in the regulation of pancreatic histogenesis and adipogenesis. The gene is essential in metabolic processes regulation, and its pathological variants are associated with susceptibility to diabetes. Polymorphisms in the IGF2BP2 gene affect the functional state of β -cells in the pancreas. Certain forms of the gene are closely linked to overweight and obesity. Some forms of the IGF2BP2 gene may increase the risk of developing Type 2 Diabetes (T2D) through alterations in adipose tissue or impaired β -cell function, leading to reduced insulin secretion. Specifically, higher levels of fasting plasma glucose, total cholesterol, and serum insulin are observed in T2D patients carrying the C allele of rs1470579, suggesting a potential role of IGF2BP2 in the pathogenesis of insulin resistance.

The CDKN2A and CDKN2B genes (rs10811661) are tumor suppressor genes involved in cell apoptosis, tumor formation, and proliferation regulation. The CDKN2A gene encodes the cyclin-dependent kinase inhibitor p16INK4A, with the highest expression in adipose tissue and adrenal glands. The CDKN2B gene encodes p15INK4B and is located near the CDKN2A tumor suppressor gene in a region frequently mutated and subject to deletions in various tumors. This gene encodes a cyclin-dependent kinase inhibitor, p15INK4B, which forms complexes with CDK4 or CDK6 and prevents CDK kinase activation. Therefore, the encoded protein functions as a growth regulator controlling the cell cycle progression at G1 phase. Polymorphisms in the CDKN2A/B genes influence protein metabolism, which determines the state of β -cells, their insulin secretion, and the proliferation of this cell population. Population studies in Asia and Europe have confirmed that polymorphism in the CDKN2A/B genes is closely associated with the risk of developing T2D. Additionally, in obesity, macrophages penetrate adipose tissue, leading to chronic low-grade inflammation. These adipose tissue macrophages stimulate the secretion of pro-inflammatory cytokines and subsequently contribute to insulin resistance. SNP rs10811661, linked to the expression of the long non-coding RNA known as antisense non-coding RNA in the INK4 locus (ANRIL or CDKN2B-AS1), is closely related to diabetes risk in humans. Notably, rs10811661 in CDKN2A/B shows significant correlation with increased breast cancer risk and may hold potential value as a prognostic factor for personalized medical care.

The KCNJ11 gene (rs5219; p.Glu23Lys) encodes an integral membrane protein Kir6.2 with the function of a potassium channel. The encoded protein is controlled by G-proteins and is responsible for potassium influx into the cell. Elevated glucose levels stimulate the opening of the Kir6.2 channel and the influx of K^+ ions into the cell. Increased intracellular K^+ levels depolarize the cell membrane and induce Ca^{2+} channels to increase free intracellular Ca^{2+} levels. Ca^{2+} ions trigger insulin secretion mechanisms, releasing the hormone from secretory granules. Polymorphism in the KCNJ11 gene may lead to reduced insulin production due to decreased or absent Kir6.2 protein expression. Therefore, mutations in this gene are responsible for various forms of diabetes, both monogenic and polygenic, such as familial persistent hyperinsulinemic hypoglycemia of infancy (PHHI), an autosomal recessive disorder

characterized by unregulated insulin secretion. Defects in this gene may also contribute to autosomal dominant insulin-independent Type 2 Diabetes (NIDDM), transient neonatal diabetes mellitus type 3 (TNDM3), and permanent neonatal diabetes mellitus (PNDM). Specifically, the KCNJ11 rs5219 allele may reduce the channel's sensitivity to ATP. A close association has been established between rs5219 polymorphisms and T2D susceptibility in East Asian and European populations. Initially, with rs5219 in the human genotype, insulin sensitivity is increased, thus maintaining normal glucose tolerance. However, over time, as insulin secretion decreases or insulin resistance develops further, glucose levels rise, leading to Type 2 Diabetes.

The CDKAL1 gene (rs7756992) encodes the CDK5 regulatory subunit-associated protein-1. This protein exhibits high expression levels in neural tissue and pancreatic beta cells and is involved in the glucose-dependent regulation of insulin secretion in the pancreatic islets. Previous studies have shown that rs7756992 is associated with the risk of impaired insulin secretion. The polymorphism rs775480 is located in intron 5 of the CDKAL1 gene, and this SNP is associated with reduced glucose sensitivity and insulin secretion in β -cells. Furthermore, the rs10946398 polymorphism of the CDKAL1 gene has been proposed as a marker for impaired insulin secretion, as the CC/CA genotypes (presence of the C allele) contribute to the susceptibility to type 2 diabetes (T2D) in individuals with obesity. Currently, the alleles rs4712524, rs10946398, and rs775480 in CDKAL1 are established risk factors for T2D.

The PPAR γ 2 gene (rs1801282) encodes PPAR γ 2 (peroxisome proliferator-activated receptor gamma 2), a ligand-activated transcription factor of the nuclear hormone receptor superfamily. The PPAR γ 2 gene plays a crucial role in glucose homeostasis, lipid metabolism, obesity, insulin sensitivity, and the regulation of the expression of various adipocyte-specific genes, directly involved in the pathogenesis of type 2 diabetes. The pathological variant of the gene rs1801282, associated with obesity and the development of T2D, was one of the first identified. Several studies have shown that the Pro12 PPAR γ allele is associated with a 1.25-fold increase in T2D risk. Furthermore, the Pro12Ala polymorphism has been associated with T2D risk in a multiethnic study.

The SLC30A8 gene (rs13266634) encodes the ZNT8 protein, a zinc ion transporter involved in zinc accumulation in intracellular vesicles. It is a proton-coupled zinc ion antiporter that mediates zinc uptake into the secretory granule lumen of pancreatic beta cells, thereby regulating insulin secretion. This gene is highly expressed only in the pancreas, particularly in the islets of Langerhans. The encoded protein co-localizes with insulin in the secretory pathway granules of insulin-secreting cells. Low expression of the ZNT8 protein leads to reduced insulin production by β -cells. The SLC30A8 rs13266634 polymorphism has been closely linked to an increased risk of developing T2D. Additionally, rs13266634 in SLC30A8 has been associated with age as a risk factor for T2D development.

The POMC gene encodes the proopiomelanocortin proprotein, which undergoes extensive, tissue-specific post-translational processing. The proprotein has several potential cleavage sites, and depending on the tissue type and available convertases, processing can yield up to ten biologically active peptides involved in various cellular functions. The encoded protein is primarily synthesized in corticotrophic cells of the anterior pituitary gland, which use four cleavage sites, including adrenocorticotropin essential for steroidogenesis regulation in the adrenal glands, and beta-lipotropin. In other tissues, including the hypothalamus, placenta, and epithelium, all cleavage sites may be utilized, resulting in the formation of endorphins, melanocyte-stimulating hormone, and immunomodulation. These include several distinct melanotropins, lipotropins, and endorphins. In patients with T2D and obesity linked to mutations in the POMC gene, a missense mutation in the signal peptide was observed, leading to a heterozygous arginine-to-glycine substitution at A15G-POMC (codon 15), which

suppressed the production and secretion of the POMC protein. In humans, POMC deficiency can cause insulin resistance and hyperinsulinemia.

The ADIPOQ gene encodes the regulatory peptide adiponectin in adipose tissue, an adipocytokine. Adipose tissue plays a significant role in the development of diabetes and obesity. Adiponectin, the primary secretory protein of adipocytes in human plasma, functions as an energy regulator and is involved in glucose tolerance. Adiponectin from adipose tissue exists in various forms, such as monomers, trimers, and hexamers. It acts through its receptors, adipoR1 and adipoR2, affecting multiple metabolic processes. Activation of adipoR1 primarily promotes mitochondrial biogenesis and glucose uptake through the AMPK and p38 MAPK pathways. Activation of adipoR2 is involved in fatty acid oxidation via the PPAR α pathway and provides cytoprotective and anti-inflammatory effects by modulating NF- κ B. Collectively, these pathways contribute to the regulation of energy metabolism and inflammatory responses in various tissues, including muscles, liver, kidneys, brain, and bones. The ADIPOQ gene is located on human chromosome 3q27 and is reported as a susceptibility locus for T2D. Polymorphisms in the ADIPOQ gene, such as rs1501299, rs182052, and rs7627128, show a significant association with T2D. Additionally, the ADIPOQ gene polymorphism, particularly the rs2241766 allele, is shown to increase the risk of developing T2D and is involved in the pathogenesis of diabetic nephropathy.

GENES OF ANTIOXIDANT ENZYMES

The imbalance between antioxidants and reactive oxygen species (ROS) leads to increased oxidative stress, which may result in diabetes. The formation and accumulation of ROS in β -cells can cause β -cell dysfunction, insulin production defects, and impaired function, leading to diabetes. However, the impact of oxidative stress can be reduced or altered by enzymatic antioxidants, including catalase (CAT), glutathione-S-transferase (GST), glutathione peroxidase (GPx), superoxide dismutase (SOD), nitric oxide synthase, and NADPH oxidase. Individuals with polymorphisms affecting the genetic regulation of these six enzymes are at an increased risk of developing Type 2 Diabetes (T2D). Moreover, the risk of developing T2D rises with increased genetic variation in the genes regulating antioxidant enzymes.

In addition to the aforementioned genes, susceptibility genes for Type 2 Diabetes include ANK1/NKX6-3 (rs33981001, rs62508166), HHEX (rs1111875), WFS1 (rs10010131), MTNR1B (rs1387153), and DGKB (TMEM195) (rs10244051).

RECOMMENDED LITERATURE FOR CHAPTER 3. POLYGENIC DISEASES

1. Genetic Passport — The Basis of Individual and Predictive Medicine, edited by V. S. Baranov. — St. Petersburg: Publishing House N-L, 2009. — 528 pages.
2. Shi C. et al. Multifactorial Diseases of the Heart, Kidneys, Lungs, and Liver and Incident Cancer: Epidemiology and Shared Mechanisms. *Cancers (Basel)*. 2023 Feb; 15(3): 729. doi: 10.3390/cancers15030729.
3. Zamorano J.L., Lancellotti P., Rodriguez Munoz D. et al. 2016 ESC Position Paper on cancer treatments and cardiovascular toxicity developed under the auspices of the ESC Committee for Practice Guidelines: The Task Force for cancer treatments and cardiovascular toxicity of the European Society of Cardiology (ESC). *Eur. Heart J.* 2016; 37:2768–2801. doi: 10.1093/eurheartj/ehw211.
4. Hasin T., Gerber Y., McNallan S.M. et al. Patients with heart failure have an increased risk of incident cancer. *J. Am. Coll. Cardiol.* 2013; 62:881–886. doi: 10.1016/j.jacc.2013.04.088.

5. Meijers W.C., Maglione M., Bakker S.J.L. et al. Heart Failure Stimulates Tumor Growth by Circulating Factors. *Circulation*. 2018; 138:678–691. doi: 10.1161/CIRCULATIONAHA.117.030816.
6. Wong G., Hayen A., Chapman J.R. et al. Association of CKD and cancer risk in older people. *J. Am. Soc. Nephrol.* 2009; 20:1341–1350. doi: 10.1681/ASN.2008090998.
7. Wong G., Staplin N., Emberson J. et al. Chronic kidney disease and the risk of cancer: An individual patient data meta-analysis of 32,057 participants from six prospective studies. *BMC Cancer*. 2016; 16:488. doi: 10.1186/s12885-016-2532-6.
8. Locatelli F., Canaud B., Eckardt K.U. et al. Oxidative stress in end-stage renal disease: An emerging threat to patient outcome. *Nephrol. Dial. Transpl.* 2003; 18:1272–1280. doi: 10.1093/ndt/gfg074.
9. Igo R.P. Jr., Kinzy T.G., Bailey J.N. Genetic Risk Scores. *Curr Protoc Hum Genet*. 2019; 104(1): e95. doi: 10.1002/cphg.95.
10. Palomaki G.E., Melillo S., Marrone M., Douglas M.P. Use of genomic panels to determine risk of developing type 2 diabetes in the general population: a targeted evidence-based review. *Genet Med*. 2013; 15(8): 600–611. doi: 10.1038/gim.2013.8.
11. Akhlaghipour I., Bina A.R., Mogharrabi M.R. Single-nucleotide polymorphisms as important risk factors of diabetes among the Middle East population. *Hum Genomics*. 2022; 16: 11. doi: 10.1186/s40246-022-00383-2.
12. Shojima N., Yamauchi T. Progress in genetics of type 2 diabetes and diabetic complications. *J Diabetes Investig*. 2023; 14(4): 503–515. doi: 10.1111/jdi.13970.
13. Witka B.Z., Oktaviani D.J., Marcellino M. et al. Type 2 Diabetes-Associated Genetic Polymorphisms as Potential Disease Predictors. *Diabetes Metab Syndr Obes*. 2019; 12: 2689–2706. doi: 10.2147/DMSO.S230061.
14. Younes S., Ibrahim A., Al-Jurf R., Zayed H. Genetic polymorphisms associated with obesity in the Arab world: a systematic review. *Int J Obes (Lond)*. 2021; 45(9): 1899–1913. doi: 10.1038/s41366-021-00867-6.
15. Ausoni S., Azzarello G. Development of Cancer in Patients With Heart Failure: How Systemic Inflammation Can Lay the Groundwork. *Front Cardiovasc Med*. 2020; 7: 598384. doi: 10.3389/fcvm.2020.598384.
16. Romejko K., Markowska M., Niemczyk S. The Review of Current Knowledge on Neutrophil Gelatinase-Associated Lipocalin (NGAL). *Int J Mol Sci*. 2023; 24(13): 10470. doi: 10.3390/ijms241310470.
17. del Bosque-Plata L., Martínez-Martínez E., Espinoza-Camacho M.Á., Gragnoli C. The Role of TCF7L2 in Type 2 Diabetes. *Diabetes*. 2021; 70(6): 1220–1228. doi: 10.2337/db20-0573.
18. Lorenz K., Voight B.F. Dissecting an adiposity locus with an arsenal of genomics. *Genome Biol*. 2018; 19: 74. Published online 2018 Jun 7. doi: 10.1186/s13059-018-1460-y.
19. Akash M.S.H., Rasheed S., Rehman K. et al. Biochemical Activation and Regulatory Functions of Trans-Regulatory KLF14 and Its Association with Genetic Polymorphisms. *Metabolites*. 2023; 13(2): 199. doi: 10.3390/metabo13020199.
20. Yu X., Liao M., Zeng Y. et al. Associations of KCNQ1 Polymorphisms with the Risk of Type 2 Diabetes Mellitus: An Updated Meta-Analysis with Trial Sequential Analysis. *J Diabetes Res*. 2020; 2020: 7145139. doi: 10.1155/2020/7145139.
21. Popović A.-M., Huđek Turković A., Žuna K. et al. FTO Gene Polymorphisms at the Crossroads of Metabolic Pathways of Obesity and Epigenetic Influences. *Food Technol Biotechnol*. 2023; 61(1): 14–26. doi: 10.17113/ftb.61.01.23.7594.
22. Radi S.H. et al. HNF4 α isoforms: the fraternal twin master regulators of liver function. *Front Endocrinol (Lausanne)*. 2023; 14: 1226173. doi: 10.3389/fendo.2023.1226173.

23. Zhang P., Wu W., Ma C. et al. RNA-Binding Proteins in the Regulation of Adipogenesis and Adipose Function. *Cells*. 2022; 11(15): 2357. doi: 10.3390/cells11152357.
24. Davidson H.W., Wenzlau J.M., O'Brien R.M. Zinc transporter 8 (znt8) and beta cell function. *Trends Endocrinol Metab*. 2014; 25(8): 415–424. doi: 10.1016/j.tem.2014.03.008.
25. Yi B., Huang G., Zhou Z. Different role of zinc transporter 8 between type 1 diabetes mellitus and type 2 diabetes mellitus. *J Diabetes Investig*. 2016; 7(4): 459–465. doi: 10.1111/jdi.12441.
26. Hamidi Y. et al. A Meta-analysis of ADIPOQ rs2241766 polymorphism association with type 2 diabetes. *J Diabetes Metab Disord*. 2022; 21(2): 1895–1901. doi: 10.1007/s40200-022-01086-0.
27. Baldelli S. et al. The Role of Adipose Tissue and Nutrition in the Regulation of Adiponectin. *Nutrients*. 2024; 16(15): 2436. doi: 10.3390/nu16152436.

CHAPTER 4. PREECLAMPSIA AS A POLYGENIC (MULTIFACTORIAL) DISEASE

INTRODUCTION

Preeclampsia (PE) is the most significant health concern for women during pregnancy worldwide. PE is a pregnancy-specific condition characterized by the development of new-onset hypertension and proteinuria in the second half of pregnancy. This disease represents a serious obstetric challenge, contributing significantly to maternal and perinatal morbidity and mortality globally. PE is described as a placental disorder, as most women typically regain normal blood pressure after childbirth. The risk of developing PE can vary between 3 to 14–21% across different populations. The global trend indicates an increasing risk of PE during pregnancy among women worldwide. As a result of preeclampsia, there is an elevated risk of premature birth, perinatal mortality, and developmental disorders of the nervous system, as well as cardiovascular and metabolic diseases in later life. In Russia, PE occurs in up to 11–16% of cases and ranks third among the causes of maternal mortality. Currently, there is no universal international classification or definition system for PE. However, hypertensive disorders of pregnancy are typically divided into four categories: Gestational Hypertension: Symptoms include blood pressure of $\geq 140/90$ mm Hg after the 20th week of pregnancy in a previously normotensive woman. Preeclampsia: Symptoms include blood pressure of $\geq 140/90$ mm Hg after the 20th week of pregnancy in a woman who previously had normal blood pressure. Proteinuria, characterized by protein excretion in urine of ≥ 300 mg/L or 500 mg/24 h, in the absence of urinary tract infection. Eclampsia: Occurs in women with preeclampsia and is characterized by seizures not related to other causes. Superimposed Preeclampsia: Chronic hypertension with the development of proteinuria during pregnancy.

Initially, PE was divided into two subgroups based on histological, clinical, biological, hemodynamic, and epidemiological data. These two subgroups were identified as early (<34 weeks) and late (≥ 34 weeks) PE, depending on gestational age at diagnosis or delivery. Histological placental lesions associated with maternal vascular pathology result from impaired remodeling of spiral arteries and fetal growth restriction. Such scenarios are more commonly seen in early PE, and extracting the fetus and placenta before 34 weeks of pregnancy increases the risk of neonatal mortality. In early placental PE, the remodeling of spiral arteries is incomplete, and the arteries retain their contractility, which in late pregnancy might lead to reduced fetal blood supply—placental insufficiency—and consequently intrauterine growth retardation. The fetus, due to reduced oxygen delivery, signals the maternal body, which responds by raising blood pressure. When blood pressure rises, fetal blood supply should normally increase, but this does not occur if the spiral arteries of the uterus retain their contractile ability. Unlike late PE, the hemodynamic state in early PE is characterized by peripheral vasoconstriction without increased cardiac output.

In placental or early-stage preeclampsia, the etiology is linked to abnormal placentation under hypoxic conditions. Placental or early-onset preeclampsia carries a significantly higher risk of complications for both the mother and fetus, with a higher prevalence of placental lesions, especially between 28 and 32 weeks of pregnancy. Consequently, placental or early-onset preeclampsia is associated with fetal growth restriction and adverse maternal and neonatal outcomes.

In maternal preeclampsia, or late-stage preeclampsia, the issue arises from the interaction between an ostensibly normal placenta and maternal factors provoked by endothelial dysfunction, creating precursors for microvascular damage. Maternal preeclampsia or late onset appears to be a decompensated response to oxidative stress in the placenta, caused by

dysfunctional maternal endothelium. Endothelial dysfunction, an aspect of the systemic inflammatory response in the mother, can lead to generalized vasoconstriction and reduced blood flow to several organs, including the heart, kidneys, and brain. On the other hand, late PE, also known as the maternal subtype, develops after 34 weeks of pregnancy. It is hypothesized that this subtype results from maternal genetic predisposition to cardiovascular and metabolic diseases associated with chronic systemic inflammation combined with an altered balance between maternal blood perfusion and the increased metabolic demands of the placenta and fetus. These disturbances can lead to placental dysfunction and associated placental hypoperfusion without the placentation abnormalities observed in early PE.

However, since the pathology's cause does not appear to reside in the placenta, there tends to be a lower incidence of fetal involvement and more favorable perinatal outcomes. Despite the clear pathophysiological differences between these PE subtypes, they are often difficult to distinguish because the two subtypes may co-occur, for example, in late primiparous women with vascular diseases who exhibit abnormal placentation. Therefore, while distinguishing between types can be useful for understanding and predicting the condition, most patients with preeclampsia exhibit elements of both pathologies.

Common complications of pre-eclampsia (PE) include damage to internal organs such as the kidneys, liver, and brain. PE can lead to HELLP syndrome (hemolysis, elevated liver enzymes, and low platelet count), cerebral edema, liver, kidney, or heart dysfunction, placental abruption, intrauterine growth restriction, preterm birth, and even maternal and fetal death.

The pathogenic mechanisms underlying PE are yet to be fully understood; however, abnormalities in placental development and trophoblast invasion, placental ischemia, oxidative stress, inflammation, immune maladaptation, and thrombosis are considered key factors in disease development. It is hypothesized that the primary cause of PE is a disruption in the formation of placental tissue during early gestation due to abnormal cytotrophoblast invasion and incomplete remodeling of the uterine spiral arteries. It is widely accepted that the pathogenesis of PE begins during trophoblast invasion and spiral artery remodeling in the first and early second trimesters of pregnancy. Early PE is characterized by abnormal trophoblastic invasion due to incomplete transformation of spiral arteries, leading to placental ischemia, subsequent inflammatory response, and increased oxidative stress. Vascular endothelial growth factor (VEGF), endoglin, and placental growth factor (PlGF) are angiogenic factors significantly implicated in the etiology of PE. These factors contribute to a systemic imbalance between anti-angiogenic and pro-angiogenic factors, resulting in a clinical syndrome. Consequently, increased endothelial activation in the mother and a general hyper-inflammatory state precede the appearance of clinical symptoms of PE.

Women with PE are in a potentially life-threatening health condition due to the increased risk of long-term complications such as cardiovascular diseases. It has been established that these components are largely determined by genetic factors, which may play a role in the pathogenic changes occurring. The familial nature of PE has been known for many years, and extensive genetic research has been conducted in this area. Twin studies have shown that the heritability of PE is about 55%, with contributions from both maternal and fetal genes. The prevailing theory views PE as a multifactorial disease based on heredity and environmental factors.

Despite the extensive research on the relationship between genetic polymorphisms and PE, there is no basis to claim that a single genetic polymorphism plays a leading role. Most researchers believe that it is likely only multiple combinations of polymorphic gene variants, in conjunction with environmental factors, lead to the development of this dangerous pregnancy complication. Environmental factors, including age, stress, excessive salt intake, chronic kidney and cardiovascular diseases, hypoxia, physical inactivity, obesity, diabetes, alcohol, and

traumatic brain injury, ultimately determine whether low-penetrance gene polymorphisms of predisposition result in phenotypic manifestations of the disease. It is probable that no single environmental factor or genetic variant is universal and can explain all cases of PE. Meanwhile, the diversity of their combinations may influence the diverse phenotypic manifestation of specific forms of disease progression (for instance, PE with concurrent intrauterine growth restriction).

HEREDITARY NATURE OF THE DISEASE

In some rare cases, familial preeclampsia (PE) appears to be inherited according to Mendelian patterns of genetic inheritance. In such instances, PE is determined by a rare monogenic pathogenic variant of a gene with high penetrance. However, in most cases, PE appears to be a complex genetic disorder, resulting from numerous common gene variants across different loci that individually have minor effects but collectively contribute to a pregnant individual's susceptibility to the disease.

To identify genetic predisposition to PE, it is crucial to analyze both maternal and fetal genotypes. The role of the placenta in the primary pathogenesis of PE indicates the contribution of fetal genes to disease susceptibility. The close association of severe early PE with cases of fetal chromosomal abnormalities also points to this. Furthermore, there is an established increased risk of PE in individuals whose fathers were born under PE conditions. Based on this, a hypothesis of genetic conflict has been developed to explain genetic predisposition to PE. According to this hypothesis, fetal (including paternal) genes exert a stimulating influence on placental blood flow to enhance nutrient transfer to the fetus from the maternal body. In contrast, maternal genes strive to limit nutrient transfer within a certain optimal range permissible for the maternal body. In other words, fetal genes increase maternal blood pressure to enhance uteroplacental blood flow, while maternal genes act oppositely. Thus, endothelial dysfunction in individuals with PE can be interpreted as a compensatory attempt by the fetus to counter insufficient uteroplacental nutrient supply.

Initially, most genetic studies of PE predominantly considered the maternal genotype. Additionally, the reduced risk of developing PE in subsequent pregnancies complicated the analysis of fetal genotype contribution. Nevertheless, significant attention is now given to genes expressed in the placenta. According to current data, approximately 35% of PE cases can be attributed to maternal genotype specificity, 20% to fetal genotype specificity, 13% to paternal genotype specificity, and 32% to environmental factors.

Gene expression analysis is a vital and efficient tool for identifying genes that are differentially expressed between tissues in normotensive and preeclamptic pregnancies. Recently, more than 100 polymorphic gene variants have been linked to the pathogenesis of PE. The genetic map of the disease is quite diverse and includes groups of genes related to blood coagulation systems, metabolism, major histocompatibility complexes, cytokines and growth factors, as well as genes of the endothelium and vascular system.

GENES PREDISPOSING TO PREECLAMPSIA

The identification of genes predisposing to PE is primarily conducted using next-generation sequencing methods (sometimes along with PCR techniques) by comparatively analyzing gene polymorphisms between individuals with PE and those experiencing physiological pregnancies. Most studies analyze a single polymorphism in a candidate gene, although literature data exist for several predisposition genes. Candidate genes are selected based on the current understanding of the pathophysiological mechanisms of PE. Population

studies based on the PE/control case principle not only identify genes predisposing to PE but also those that, conversely, have a restraining (protective) effect.

On the other hand, population analysis of the frequency of genetic factors with high or low risk can improve the quality of PE risk prediction. Furthermore, this strategy facilitates more effective early prenatal screening and interventions in populations with higher susceptibility to PE.

GENES REGULATING ENDOTHELIAL FUNCTION AND BLOOD PRESSURE IN THE PATHOGENESIS OF PRE-ECLAMPSIA

Adverse outcomes associated with hypertensive disorders during pregnancy result from excessive inflammation and endothelial damage, which include eclampsia, stroke, retinal detachment, acute renal failure, placental abruption, pulmonary edema, liver hematoma, disseminated intravascular coagulation, and cerebrovascular bleeding. Fetal complications encompass stillbirth, intracranial hemorrhage, oligohydramnios, and fetal growth restriction. The prognosis for maternal and fetal health generally converges around the diagnosis of pre-eclampsia. Consequently, one of the significant symptoms of pre-eclampsia is elevated blood pressure during pregnancy. Therefore, genes involved in blood pressure regulation and markers of endothelial dysfunction are included in all genetic panels analyzing predisposition to pre-eclampsia. It is possible that placental disorder, primarily characterized by uteroplacental ischemia secondary to vascular remodeling disruption, is the first anomaly in the pathogenesis of pre-eclampsia. There is accumulating evidence suggesting that a cardiovascular disorder might be the initial anomaly leading to placental ischemia, which in turn initiates pre-eclampsia.

VEGFA GENE The VEGFA gene encodes the Vascular Endothelial Growth Factor A, located on chromosome 6 at cytogenetic band 6p21.3. The gene belongs to the PDGF/VEGF growth factor family and encodes a heparin-binding protein that exists as a disulfide-linked homodimer. This growth factor induces proliferation and migration of vascular endothelial cells and is essential for both physiological and pathological angiogenesis. The polymorphism of this gene is associated with severe diseases. Notably, it is activated in many tumors, with expression levels correlating with the tumor stage and progression. Elevated levels of this protein are found in patients with POEMS syndrome, also known as Crow-Fukase syndrome. Allelic variants of this gene are linked to microvascular complications of type 1 diabetes (MVCD1) and atherosclerosis. Vascular Endothelial Growth Factor A is critical for placental development and vascularization. VEGF is vital for the formation of trophoblasts, the embryonic vascular network, and maternal and fetal blood cells. Abnormal vascular growth and endothelial dysfunction may be part of the pathogenesis of pre-eclampsia. It is known that VEGFA is expressed at significantly higher levels in the placentas of women with pre-eclampsia compared to healthy women (49.6% versus 14.9%, respectively). Data suggest that two VEGF gene polymorphisms (rs3025039, rs2010963) are associated with an increased risk of pre-eclampsia in various ethnic groups, indicating that these targets could serve as useful genetic markers for early pre-eclampsia prediction. Specifically, several reports indicate that rs3025039, located in the 3'-untranslated region of VEGFA and capable of modulating gene expression, is considered a predisposing factor for pre-eclampsia in populations from China, Southern Europe, and Northeast Africa, increasing disease risk. Regarding the association of rs3025039 (NC_000006.12:g.43784799C>T) with pre-eclampsia and its potential role in its pathogenesis, contradictory opinions exist in the literature. Nevertheless, there is reason to believe that a significant association between the VEGFA gene polymorphism and pre-eclampsia may exist in certain populations. The VEGF gene polymorphism rs2010963, c.-94C>G, may also be

associated with an increased risk of pre-eclampsia in various ethnic groups, indicating that this SNP could be a useful genetic marker for early pre-eclampsia prediction.

ACE GENE This gene encodes the Angiotensin-Converting Enzyme (ACE), which is involved in regulating blood pressure and fluid-salt balance. It catalyzes the conversion of angiotensin I to the physiologically active octapeptide angiotensin II. Angiotensin II is a potent vasopressor and aldosterone-stimulating peptide that controls blood pressure and water-electrolyte balance. The Angiotensin-Converting Enzyme (ACE) also inactivates the vasodilator protein bradykinin. Thus, the enzyme encoded by the gene increases blood pressure and is a target for pharmaceutical ACE inhibitors, commonly prescribed to lower blood pressure. Gene polymorphisms (mutations in this gene) are involved in a wide range of conditions, including cardiovascular diseases, psoriasis, kidney failure, stroke, and Alzheimer's disease. Moreover, the gene polymorphism (rs4344) is a widely recognized marker for hypertension and other cardiovascular diseases and is also linked to the risk of pre-eclampsia.

GENE AGT AND GENE NOS3 ANALYSIS

Gene AGT: The protein encoded by the AGT gene is preangiotensinogen, a precursor of angiotensinogen. This protein is expressed in the liver and is cleaved by the enzyme renin in response to decreased blood pressure and altered sodium transport in the kidney tubules. The resulting product, angiotensin-I, is then converted by the angiotensin-converting enzyme (ACE) into the physiologically active enzyme angiotensin-II. Preangiotensinogen is involved in maintaining blood pressure, fluid and electrolyte homeostasis, and the pathogenesis of essential hypertension. Mutations in the AGT gene are associated with susceptibility to essential hypertension and may cause renal tubular dysgenesis, a severe defect in kidney tubule development. Defects in this gene are also linked to non-familial atrial fibrillation and inflammatory bowel disease. AGT gene polymorphism is also closely involved in the pathogenesis of preeclampsia (PE). Pregnant individuals with PE have elevated levels of AGT protein in the urine. Available data suggest that the presence of the 217A, 235T, and 704C alleles of the AGT gene significantly increases the risk of severe PE by 2.23 times. Notably, the frequency of the CC genotype of the T704C polymorphism contributes to an increased risk of early PE compared to those with late preeclampsia.

Gene NOS3: This gene encodes the enzyme eNOS (endothelial nitric oxide synthase). Nitric oxide (NO) is a reactive free radical that acts as a biological mediator in several processes, including neurotransmission, antimicrobial, and antitumor activities. Nitric oxide is synthesized from the amino acid L-arginine by nitric oxide synthases. NO influences vascular tone, platelet aggregation, stabilization of blood rheological properties, neurotransmission, and is involved in mitigating metabolic syndrome and several immune responses. Alterations in the NOS3 gene are associated with susceptibility to various cardiovascular diseases, including PE. Allelic variants of the NOS3 gene lead to reduced NO levels, resulting in the development of hypertensive conditions, associated with changes in vascular tone due to adverse environmental and internal conditions. This may lead to PE, placental insufficiency, and fetal growth restriction. In the pathogenesis of preeclampsia, endothelial dysfunction is significant, characterized by increased sensitivity of the vascular wall to pressor influences of mediators and simultaneous reduced production of vasodilators such as NO. Studies conducted in Northern European populations indicate that individuals homozygous for rs1799983 in NOS3 are 1.6 times more likely to suffer from severe PE. Specifically, the genotype for SNP NOS3 rs1799983 is not only associated with PE risk but also with reduced effectiveness of antihypertensive medications in PE patients. Additionally, the 894G>T polymorphism in exon 7 of the NOS3 gene (marker rs1799983) results in the substitution of glutamine with asparagine at

position 298 of the protein, which leads to decreased nitric oxide concentration in the bloodstream and reduced vasodilation. Importantly, this variant is more common in Russian populations and is associated with PE. Diagnostic value is reported for rs2070744 in NOS3 in the early prediction of PE. The analysis of multilocus interactions of seven alleles, including rs2070744 in NOS3, associated with PE (rs4762 in AGT, rs1800896 in IL-10, rs1800629 and rs1799724 in TNF α , rs2070744 in NOS3, rs7412 in APOE, and rs2549782 in ERAP2) demonstrated high accuracy in predicting PE risk.

GENES REGULATING LIPID METABOLISM IN THE PATHOGENESIS OF PREECLAMPSIA

Disruptions in lipid metabolism, induced by lipid peroxidation due to oxidative stress, are considered significant pathogenetic mechanisms of preeclampsia (PE). It is known that the enzyme lipoprotein lipase (LPL) and the protein apolipoprotein E (ApoE), whose primary function is to transport lipids between various cells and tissues of the body, are important regulators of lipid metabolism and are abundantly expressed in the placenta. Specifically, LPL is involved in the hydrolysis of triglycerides in chylomicrons and very low-density lipoproteins. Thus, the genes encoding these proteins have been proposed as potential candidate genes. Additionally, it is known that PE is associated with alterations in the covalent modification of circulating ApoE isoforms. Specifically, the proportion of the deglycosylated main isoform of ApoE was increased in PE, while the acidic sialylated isoform of ApoE was decreased. Functionally, this may increase the risk of atherogenic changes in the placenta. The most promising genetic variant in this context is the missense mutation, Asn318Ser, in LPL, which correlates with reduced LPL enzyme activity and increased dyslipidemia. It has also been reported that the fetal genotype of these two genes contributes to the metabolism of maternal lipoproteins. Lipid transport in the placenta is not only a primary energy source for the fetus but also a key material that supports the development of the fetal nervous system. Disruption of lipid metabolism may cause inflammatory reactions and accumulation of oxidative stress products, affecting trophoblast invasion and placental development, as well as causing endothelial damage, leading to a range of pathological changes. Compared to pregnant individuals with a physiological pregnancy course, serum lipid levels and lipoprotein metabolism parameters are impaired in patients with PE. High lipid levels in the blood of patients with PE lead to atherosclerotic changes during pregnancy, which may persist for several years postpartum, increasing the risk of cardiovascular diseases.

Some authors highlight polymorphisms of the ADIPOQ gene as a risk factor for PE. This gene encodes the physiologically active protein adiponectin, which is exclusively expressed in adipose tissue. Obesity negatively influences the course of pregnancy and alters the quality of labor (premature or prolonged labor) and increases the need for oxytocin. Adipose tissue, the largest endocrine and paracrine organ, especially in obesity, is responsible for the production of adipokines. Low serum adiponectin concentrations are associated with diabetes, overweight, insulin resistance, and metabolic syndrome. Disrupted regulation of adiponectin during pregnancy contributes to gestational diabetes mellitus and preeclampsia.

It is important to note that lipid metabolism disorders caused by excess body weight have an extremely negative impact on pregnancy, altering tissue sensitivity to sex steroid hormones and creating additional risks during childbirth.

The LPL gene encodes lipoprotein lipase, expressed in the heart, muscles, and adipose tissue. The LPL enzyme functions as a homodimer and has dual roles: triglyceride hydrolase activity and acting as a ligand/bridging factor for receptor-mediated lipoprotein uptake. Severe mutations causing LPL deficiency lead to Type I hyperlipoproteinemia. Less severe mutations

in LPL are associated with various disorders of lipoprotein metabolism. Lipoprotein lipase (LPL) participates in triglyceride hydrolysis and plays a crucial role in lipid transport in the placenta. It has been identified that the maternal allele LPL N291S (rs268, c.953A>G (p.Asn318Ser)) is associated with an increased risk of preeclampsia and HELLP syndrome. Additionally, the presence of the LPL N291S allele in the fetal genotype is linked to higher triglyceride levels in the mother and an increased risk of PE.

APOE GENE

The protein encoded by the APOE gene is a major apolipoprotein of chylomicrons. It binds to specific receptors on liver and peripheral cells and is necessary for the normal catabolism of triglyceride-rich lipoprotein constituents. This gene is located on chromosome 19. Mutations in this gene lead to familial dysbetalipoproteinemia or type III hyperlipoproteinemia (HLP III), where elevated plasma cholesterol and triglycerides result from impaired clearance of chylomicron and VLDL remnants. It is known that the levels of triglycerides, low-density lipoprotein cholesterol, and Apo-E from weeks 4-16 to weeks 28-42 of pregnancy increased more among pregnant individuals with pre-eclampsia (PE) than those with chronic hypertension. Most authors suggest that increased levels of triglycerides and Apo-E elevate the risk of developing pre-eclampsia. Some data indicate that the ApoE rs440446 allele is associated with a higher risk of PE. At the same time, it is suggested that this gene variant may be involved in the pathogenesis of gallstone disease and is associated with the risk of bile duct cancer in males.

ADIPOQ GENE

The ADIPOQ gene consists of three exons and two introns, located on chromosome 3q27, where a diabetes susceptibility locus has been mapped. This gene encodes the physiologically active protein adiponectin and is expressed exclusively in adipose tissue. It encodes a protein similar to collagens X and VIII and complement factor C1q. The encoded protein circulates in plasma and participates in metabolic and hormonal processes. Adiponectin is the most abundant circulating protein synthesized exclusively in adipose tissue and acts as a hormone with anti-inflammatory, antidiabetic, and insulin-sensitizing properties. It plays a significant role in various metabolic processes, including glucose regulation and fatty acid catabolism. Decreased plasma adiponectin levels are closely associated with the pathogenesis of type 2 diabetes mellitus (T2DM), several cardiovascular diseases, obesity, and disruptions in normal pregnancy progression. Adiponectin is found in plasma in oligomeric complexes composed of trimers (low molecular weight), hexamers (medium molecular weight), and large multimers of 12-18 subunits (high molecular weight (HMW)). Moreover, the HMW isoform binds more effectively to its receptors and activates AMP-activated protein kinase, one of the primary enzymes controlling adiponectin's metabolic effects. The HMW form of adiponectin, mediating insulin sensitivity in peripheral tissues, is the most physiologically active form regarding glucose homeostasis. Mutations in this gene are often associated with adiponectin deficiency. It is known that the gene variant rs2241766 (+45T/G) closely correlates with the risk of pregnancy complications and PE induced by obesity and gestational diabetes against the background of lower plasma adiponectin levels.

GENES OF THE BLOOD COAGULATION SYSTEM REGULATION IN THE PATHOGENESIS OF PE

Thrombophilia is a pathological condition characterized by a blood coagulation system disorder, leading to an increased risk of thrombosis. However, adequate placental circulation is necessary for the physiological course of pregnancy. Meanwhile, thrombophilias may heighten the risk of placental insufficiency due to microvascular thrombosis of the placenta, macrovascular thrombosis, or both, and influence the growth and differentiation of the trophoblast. The role of coagulation mechanism abnormalities in the pathogenesis of PE is widely acknowledged. Endothelial damage in PE is closely associated with a shift from an anticoagulant to a procoagulant phenotype and reduced endothelium-mediated vasorelaxation. This phenotype may be present before PE during pregnancy or may develop as a result of damage initiated during placentation. Against the backdrop of coagulation mechanism disorders in PE, individuals develop marked thrombocytopenia, often linked to hemolysis, elevated liver enzyme levels in the blood plasma, and the low platelet count syndrome (HELLP). Currently, the role of three thrombophilic factor genes, Factor V Leiden (FVL), methylenetetrahydrofolate reductase (MTHFR), and prothrombin (F2), is well recognized.

The FV gene encodes Factor V, a crucial cofactor in the blood coagulation cascade. This factor circulates in plasma and is converted into its active form by the release of an activation peptide through thrombin during coagulation, forming a heavy chain and a light chain held together by calcium ions. The activated protein serves as a cofactor, working with the activated clotting Factor X to convert prothrombin into thrombin. Defects in this gene can lead to either autosomal recessive hemorrhagic diathesis or an autosomal dominant form of thrombophilia known as resistance to activated protein C. The anomaly of the FV coagulation factor gene involves a single nucleotide substitution from guanine to adenine at position 1601 (rs6025, c.1601G>A (p.Arg534Gln)), causing Factor V to resist degradation. Current research indicates that FV rs6025 is associated with an increased risk of preeclampsia in European and some Asian populations. Pregnant individuals with the FV mutation are more frequently found to have placental thrombi, thus significantly increasing the risk of preeclampsia (PE) and other obstetric pathologies, including antenatal fetal death. Additionally, the FV gene polymorphism, specifically rs6022 (c.319G>C (p.Asp107His)), in conjunction with smoking, increases the risk of pregnancy complications and premature birth.

The F2 gene encodes the protein prothrombin (also known as coagulation Factor II), which undergoes proteolytic cleavage in several stages to form the active serine protease thrombin. Activated thrombin plays a vital role in thrombogenesis and hemostasis by converting fibrinogen into fibrin during clot formation, promoting platelet aggregation, and activating additional clotting factors. Thrombin also has a role in cell proliferation, tissue repair, angiogenesis, and maintaining vessel integrity during embryonic development and postnatal life. Peptides derived from the C-terminus of this protein exhibit antimicrobial activity against *E. coli* and *P. aeruginosa*. Mutations in this gene can lead to various forms of thrombosis. A hereditary mutation in prothrombin (G20210A) exacerbates the existing hypercoagulable state during pregnancy. The nucleotide base substitution from guanine to adenine at position 20210 (G20210A) in the prothrombin gene (F2) leads to increased prothrombin synthesis and a higher risk of thrombosis. The prothrombin F2 gene mutation in combination with the FV gene mutation occurs frequently and significantly increases the risk of pregnancy complications. There are indications of a close association between the F2 gene mutation and adverse pregnancy outcomes, such as PE, spontaneous abortion, and fetal growth retardation.

IMMUNE SYSTEM FACTORS IN THE PATHOGENESIS OF PRE-ECLAMPSIA

Currently, there is substantial literature supporting the immunological causes of preeclampsia. Immunogenetic memory suggests an increased risk of PE if PE occurred in a

previous pregnancy. The risk of PE increases from 4.1% in the first pregnancy to 14.7% in the second pregnancy after the first pregnancy complicated by PE and to 31.9% in the third pregnancy in cases of two consecutive previous pregnancies complicated by PE. Unlike protective immunology, reproductive immunology is based not on self/non-self-recognition systems but on immunotolerance in maternal-fetal cell interactions. The adaptation of the maternal immune system is crucial for determining the outcome and physiological course of pregnancy. Increased cases of PE in first-time pregnancies indicate a strong connection between immune factors and the risk of PE. Additionally, advances in assisted reproductive technologies (widespread use of in vitro fertilization - IVF) present new challenges for the maternal immune system. The use of donor sperm or eggs increases the risk of PE threefold. It has been shown that placentas from IVF with donor oocytes have significantly increased lesions associated with chronic inflammation compared to placentas from IVF with non-donor oocytes.

The primary mechanisms by which fetal immunity evades maternal immune system control at the maternal-fetal interface include reduced expression of major histocompatibility complex molecules by trophoblast cells, the presence of complement regulators, increased production of indolamine 2,3-dioxygenase, activation of regulatory T cells, and enhancement of immune checkpoints. This immune protection is more akin to tumor biology immune responses than allograft biology.

The genes of the major histocompatibility complex (MHC), also known as human leukocyte antigens (HLA), are of particular interest in terms of the risk of pre-eclampsia (PE). The MHC is located on chromosome 6. The HLA molecules encoded by these genes are classified into Class I and Class II antigens. MHC Class I molecules (HLA-A, HLA-B, HLA-C) are expressed on the surface of all somatic cells, except for trophoblast cells and erythrocytes. Allelic variants of MHC Class II HLA molecules include HLA-DP, HLA-DQ, and HLA-DR. The primary function of MHC proteins is to present peptide antigens to T-lymphocytes. HLA-DR4, a Class II antigen, shows the strongest association with PE. Non-classical HLA antigens such as HLA-E and HLA-G are considered significant in the pathogenesis of PE because they are expressed by trophoblast cells. Among these, only HLA-G shows pronounced polymorphism.

The expression of HLA on the implanting cytotrophoblast is crucial as it interacts with killer-cell immunoglobulin-like receptors (KIR) expressed on maternal uterine natural killer (uNK) cells and cytotoxic T-lymphocytes, thereby reducing their cytolytic activity and stimulating the production of cytokines necessary for successful placentation. Natural killer (NK) cells identify and destroy cells that do not express an MHC molecule, a recognition mechanism vital for protecting the organism from foreign pathogens and cancer cells. Unlike circulating peripheral blood NK cells, uNK cells in contact with the trophoblast perform a cytokinic rather than cytotoxic function, being activated by mechanisms involving interactions between their receptors and HLA-C, G, and E ligands expressed by the trophoblast. Activation of uNK or decidual NK cells results in the production of angiogenic factors like VEGF, stimulating angiogenesis, as well as chemokines IL-8 and IP-10, which facilitate trophoblast invasion.

Disruptions in the interaction between uNK cells and HLA-G on the trophoblast alter the decidual immune response, leading to impaired trophoblast implantation and inadequate spiral artery remodeling, associated with PE. Thus, the placenta and solid tumors exhibit similarities regarding cell invasion processes and microenvironments. Indeed, tumor cells and trophoblast cells are supported by a favorable microenvironment that allows cell invasion, angiogenesis, and immunotolerance, thus facilitating immune evasion from the host. The decidual and tumor microenvironments share similar immune cells, including NK cells, T cells, regulatory T cells, macrophages, and non-immune cells like decidual stromal cells and fibroblasts. In this

microenvironment, both tumor and trophoblast cells can express HLA-G, which interacts with NK cells, reducing their toxicity. These cells can also secrete immunosuppressive factors that inhibit T cell activity, promote the local recruitment of regulatory T cells with their immunosuppressive function, and decrease the expression of polymorphic MHC Class I and II antigens to evade the host's immune system.

During pregnancy, syncytiotrophoblasts in contact with maternal tissue do not express any MHC antigens. The trophoblast interacting with decidual cells of the uterus expresses only the classical polymorphic HLA-C Class Ia antigens and non-classical non-polymorphic HLA-G and HLA-E antigens. HLA-G shows limited polymorphism; however, some gene variants are also associated with PE. These variants of HLA-G, recognized by the maternal organism as foreign, can lead to histoincompatibility between mother and child, also resulting in PE. HLA-G protects trophoblast cells from lysis and regulates tolerance levels in the maternal-fetal interaction system. One allelic variant of this gene, HLA-G, is an insertion/deletion of 14 base pairs in the 3'-untranslated region of the gene, which may become a significant marker for PE.

The most optimal predictive tool in the study of HLA genes concerning pregnancy pathology, specifically PE, involves genotyping maternal genes together with allelic variants of DNA from both the paternal genotype and the fetal genotype. Thus, the high polymorphism of the HLA-C molecule expressed by the trophoblast and its interaction with highly polymorphic KIR receptors expressed by maternal NK cells allows for numerous combinations and potential mismatches between the mother and fetus (highly diversified interaction). In contrast, the non-classical MHC Class Ib includes ancestral ligands HLA-E, HLA-F, and HLA-G, and demonstrates a low level of polymorphism.

When describing the role of immune system proteins, it is important to note that certain pro-inflammatory cytokines can also contribute to the pathogenesis of preeclampsia (PE). Specifically, it has been established that excessive secretion of tumor necrosis factor alpha (TNF α) may participate in endothelial activation, which, in turn, can contribute to the symptoms of PE. During pregnancy, TNF α influences hormone synthesis, placental architecture, and embryonic development. It is notable that TNF α levels in blood plasma are significantly higher in women with PE than in those with normal pregnancies. Inadequate stimulation of TNF α secretion promotes the production of reactive oxygen species, leading to tissue damage, including endothelial damage. Meanwhile, oxidative stress and inflammatory responses may play a common role in the progression of PE. Understanding this fact forms the basis for identifying the genetic mechanism of PE. Additionally, elevated TNF α levels in PE negatively affect trophoblast functions related to its migratory activity, syncytialization, and subsequent endocrine function. Increased TNF α levels adversely impact the relationship between the mother and fetus, altering the secretory profile of placental immunomodulatory factors, which in turn affects the state of the maternal immune system.

Interleukin-10 (IL-10) is also presumed to participate in the pathogenesis of preeclampsia, enhancing the inflammatory response to trophoblast cells and leading to reduced invasion and remodeling of spiral arteries.

The HLA-G gene encodes the human leukocyte antigen HLA-G protein, which is a paralog of the heavy chain of the major histocompatibility complex (MHC) class I. This class I molecule is a heterodimer consisting of a heavy chain and a light chain (beta-2 microglobulin). The heavy chain is anchored in the membrane. HLA-G is expressed on placental cells. The heavy chain is approximately 45 kDa, and its gene contains 8 exons. Exon 1 encodes the leader peptide; exons 2 and 3 encode the alpha1 and alpha2 domains that bind the peptide; exon 4 encodes the alpha3 domain; exon 5 encodes the transmembrane region; and exon 6 encodes the cytoplasmic tail.

Seven different HLA-G mRNA transcripts have been identified, attributed to alternative splicing of its seven exons. HLA-G2, G3, and G4 transcripts are translated into membrane-bound isoforms; HLA-G5, G6, G7 into soluble HLA-G isoforms (sHLA-G); and HLA-G1 into both types of isoforms. HLA-G is expressed on the surface of pre-implantation embryos and trophoblasts. The level of HLA-G in plasma in women with placental abruption was significantly reduced. Meanwhile, sHLA-G was detected in the culture medium of in vitro fertilized (IVF) embryos, maternal blood, amniotic fluid, and fetal blood. HLA-G plays an important role in spiral artery remodeling, fetal development, and immune tolerance. The presence of soluble HLA-G in embryo culture medium significantly correlates with increased pregnancy rates in assisted reproductive technologies (IVF). Additionally, the level of soluble HLA-G in patients' serum correlates with pregnancy outcomes after IVF.

There are several single nucleotide polymorphisms (SNPs) in the 3'-untranslated region (UTR) of HLA-G that affect mRNA stability, metabolism, and transport, including +3035C/T, +3187A/G, +3196C/G, +3142C/G, +3001C/T, +3003C/T, +3010C/G, +3027C/A, +3032C/G, +3052C/T, +3092G/T, +3111A/G, +3121C/T, +3177G/T, +3183A/G, and +3227A/G. These SNPs alter micro-RNA binding, leading to reduced HLA-G expression. In turn, abnormal expression and polymorphisms of HLA-G are associated with poor pregnancy outcomes, such as PE and recurrent spontaneous abortions. A significant association of the GC-ins haplotypes (-725G>C SNP in the HLA-G promoter region) with infertility has also been identified. The role of gene variants rs9380142, rs1610696, and rs1707 in the pathogenesis of PE in the context of viral diseases is discussed.

The HLA-DQB1 gene encodes the HLA-DQB1 protein, which is a paralog of the beta chain of HLA class II. This class II molecule is a heterodimer consisting of an alpha (DQA) and beta chain (DQB), both anchored in the membrane. It plays a central role in the immune system by presenting peptides derived from extracellular proteins. Class II molecules are expressed in antigen-presenting cells (APC: B-lymphocytes, dendritic cells, macrophages). The beta chain is approximately 26-28 kDa and contains six exons. Exon 1 encodes the leader peptide; exons 2 and 3 encode two extracellular domains; exon 4 encodes the transmembrane domain; and exon 5 encodes the cytoplasmic tail. In the DQ molecule, both the alpha and beta chains contain polymorphisms that determine peptide-binding specificity, leading to the formation of up to four different molecules. Genotyping these polymorphisms is commonly performed in bone marrow transplantation. Alternative splicing results in multiple transcript variants.

Compared to the group of women experiencing normal pregnancies, the frequency of the HLA-DQB1*0502 allele was significantly higher among women with pregnancy-induced hypertension. It is known that the HLA-DQB1*0503 allele may serve as a marker for high blood pressure during pregnancy, whereas women carrying the HLA-DQB1*04 allele in their genotype are at increased risk of developing preeclampsia (PE). Conversely, the absence of the HLA-DQB1*0602 allele is a significant risk factor for PE.

The TNF- α gene encodes a multifunctional pro-inflammatory cytokine that belongs to the tumor necrosis factor (TNF) superfamily. It is located in the human leukocyte antigen (HLA) region, class II, of the major histocompatibility complex (MHC) at chromosome location 6p21.33. This cytokine is primarily secreted by macrophages and can bind to receptors TNFRSF1A/TNFR1 and TNFRSF1B/TNFR2, thereby functioning through them. It is involved in regulating a wide range of biological processes, including cell proliferation, differentiation, apoptosis, lipid metabolism, and coagulation. This cytokine is implicated in the pathogenesis of various diseases, including autoimmune diseases, insulin resistance, psoriasis, rheumatoid arthritis, ankylosing spondylitis, tuberculosis, autosomal dominant polycystic kidney disease, and cancer. Mutations in this gene affect susceptibility to cerebral malaria, septic shock, and

Alzheimer's disease. TNF- α gene knockout studies in mice suggest a neuroprotective function for this cytokine.

During pregnancy, TNF- α may be involved in the etiology of preeclampsia. This hypothesis is supported by findings that patients with preeclampsia have elevated levels of TNF- α in their blood plasma, while levels of IL-10 and IL-4 are reduced. This cytokine imbalance leads to chronic peripheral and placental inflammation. Additionally, evidence points to certain SNPs acting as transcriptional regulators of this cytokine in the promoter region (-308G/A, -238G/A). Specifically, the G/A polymorphism at position -308 has been associated with an increased risk of preeclampsia. Furthermore, women with preeclampsia are significantly more likely to carry the mutation that increases TNF- α expression, -308 A/G, compared to women with normal blood pressure. The frequency of the AA allele at position -308 is notably higher among Slovakian, Finnish, Iranian, and Turkish populations with preeclampsia and Turkish women with eclampsia. Notably, the frequency of the AA allele is significantly higher among European populations with PE combined with intrauterine growth restriction (IUGR) compared to those without IUGR.

It should be noted that the TNF gene promoter polymorphism at position -308G/A (rs1800629) directly influences TNF- α cytokine expression, contributing to severe PE in populations from Europe, Asia, the Middle East, and North America. It is possible that TNF gene variants associated with PE possess population-specific characteristics, as the -238G/A allele (rs361525), used as a disease marker in European populations, has not proven relevant for predicting and diagnosing PE in Southeast Asian populations.

The IFNG gene encodes the soluble cytokine IFN- γ , which is a member of the type II interferon class. It is located in the chromosomal region 12q15. The encoded protein is secreted by cells of both the innate and adaptive immune systems. The active protein is a homodimer that binds to the gamma interferon receptor, triggering a cellular response to viral and microbial infections. Mutations in this gene are associated with increased susceptibility to viral, bacterial, and parasitic infections, leukemia, and certain autoimmune diseases. High levels of IFN- γ during pregnancy are associated with complications such as PE, premature birth, and IUGR. The IFN- γ cytokine, along with TNF- α , induces trophoblastic apoptosis and limited trophoblast differentiation, leading to incomplete trophoblastic invasion of spiral arteries and impaired placental implantation, resulting in PE. The SNP +874A/T (rs2430561) is associated with preeclampsia, where the T/T genotype at this position correlates with high levels of IFN- γ production and an increased risk of severe PE in Brazilian and Chinese women.

The IL10 gene encodes the cytokine interleukin-10 (IL-10), primarily produced by monocytes and to a lesser extent by lymphocytes. This cytokine plays a pleiotropic role in immune regulation and inflammation. It inhibits the expression of Th1 cytokines, MHC class II antigens, and co-stimulatory molecules on macrophages. It also enhances B-cell survival, proliferation, and antibody production. This cytokine can block NF-kappa B activity and is involved in regulating the JAK-STAT signaling pathway. Knockout studies in mice have shown that this cytokine is a critical immune regulator in the intestinal tract. Mutations in this gene are associated with increased susceptibility to HIV-1 infection and rheumatoid arthritis.

During a physiological pregnancy, IL-10 performs three main beneficial functions: it facilitates successful placentation, controls inflammation, and regulates vascular function. Consequently, serum IL-10 levels rise during normal pregnancy and remain elevated until childbirth. Additionally, higher IL-10 levels are observed in the first and second trimesters compared to the third trimester, with normal placental tissue, ensuring an important balance of inflammation in the interaction between fetal and maternal organisms.

It should be noted that certain Single Nucleotide Polymorphisms (SNPs) in the proximal and distal promoter regions (-1082A/G, -819T/C, and -592A/C) of the IL10 gene exert a

suppressive effect on the transcription rate. For example, the A/G polymorphism at position -1082 is associated with lower production of IL-10. It is linked to an increased risk of preeclampsia in Turkish, Brazilian, and Egyptian populations. Women with the IL-10 (-1082) AA genotype are found to have a 3.38 times higher risk of developing preeclampsia, alongside higher plasma glucose levels and lower HDL cholesterol levels.

The IL22 gene belongs to the IL10 cytokine family, which mediates cellular inflammatory responses. The encoded IL-22 protein plays a key role in protecting the host from extracellular pathogens on mucosal surfaces by strengthening epithelial barrier functions and is involved in tissue homeostasis, tissue repair, and wound healing. This protein also possesses pro-inflammatory properties and plays a role in the pathogenesis of several intestinal diseases. The encoded protein is a crucial cytokine that regulates host immunity in infectious diseases, including COVID-19.

It is known that IL-22 plasma concentrations are significantly higher in groups of patients with preeclampsia compared to control groups. Increased IL-22 production leads to the activation of immune responses in the maternal organism. The resulting immune imbalance can, in turn, lead to preeclampsia through placental ischemia. Furthermore, laboratory analysis of IL-22 levels and the IL2/IL-4 concentration ratio can be used to refine pregnancy pathologies to differentiate between preeclampsia, gestational hypertension, and physiological pregnancy.

The polymorphism of the IL17 gene may be considered one of the causes for elevated IL-22 levels in preeclampsia. Moreover, the single nucleotide polymorphism (SNP) (rs2227485) in the IL22 gene is associated with the risk of preeclampsia. It has been established that rs2227485 in IL22 showed a significant difference in allele frequency in the early-onset preeclampsia group and genotypic subgroups of late-onset preeclampsia, as well as control subgroups. Significant differences were found between preeclampsia patients and the control group for this SNP in terms of genotype frequencies. The pathogenic contribution of rs2227485 to the development of preeclampsia was assessed under a dominant model of the C allele (genotype CC/CT + TT) or a recessive model of the T allele (genotype TT/CC + CT).

RECOMMENDED LITERATURE FOR CHAPTER 4. PREECLAMPSIA

1. Genetic Passport: The Foundation of Individual and Predictive Medicine / Edited by V. S. Baranov. — Saint Petersburg: N-L Publishing, 2009. — 528 pages.
2. Glotov A. S., Vashukova E. S., Glotov O. S., et al. Study of Population Frequencies of Gene Polymorphisms Associated with Gestosis. *Ecological Genetics*. 2013; Volume XI, No. 1: 91-100.
3. Tsakhilova S. G., Akulenko L. V., Kuznetsov V. M., et al. Genetic Predictors of Preeclampsia. *Problems of Reproduction*. 2017; No. 1: 110-114. doi: 10.17116/repro2017231110-114
4. Kaptilnyy V. A., Reyshtat D. Y. Preeclampsia: Definition, Advances in Pathogenesis, Guidelines, Treatment, and Prevention. V. F. Snegirev *Archives of Obstetrics and Gynecology, Russian Journal*. 2020; 7(1): 19-30. (in Russian). DOI: <http://dx.doi.org/10.18821/2313-8726-2020-7-1-19-30>
5. Williams P. J., Broughton Pipkin F. The Genetics of Pre-eclampsia and other Hypertensive Disorders of Pregnancy. *Best Pract Res Clin Obstet Gynaecol*. 2011; 25(4): 405–417. doi: 10.1016/j.bpobgyn.2011.02.007
6. Kleinrouweler C. E., van Uitert M., Moerland P. D., et al. Differentially Expressed Genes in the Pre-Eclamptic Placenta: A Systematic Review and Meta-Analysis. *PLoS One*. 2013; 8(7): e68991. doi: 10.1371/journal.pone.0068991

7. Parada-Niño L., Castillo-León L. F., Morel A. Preeclampsia, Natural History, Genes, and miRNAs Associated with the Syndrome. *J Pregnancy*. 2022; 2022: 3851225. doi: 10.1155/2022/3851225
8. Duan W., Xia C., Wang K., et al. A Meta-Analysis of the Vascular Endothelial Growth Factor Polymorphisms Associated with the Risk of Pre-eclampsia. *Biosci Rep*. 2020; 40(5): BSR20190209. doi: 10.1042/BSR20190209
9. Rokoťanskaya E. A., Panova I. A., Malysheva A. I., et al. Technologies for Prediction of Preeclampsia. *Sovrem Tekhnologii Med*. 2020; 12(5): 78–84. doi: 10.17691/stm2020.12.5.09
10. Sandrim V. C., Rizzatti Luiz M., Pihan E. Interaction Between NOS3 and HMOX1 on Antihypertensive Drug Responsiveness in Preeclampsia. *Rev Bras Ginecol Obstet*. 2020; 42(8): 460–467. doi: 10.1055/s-0040-1712484
11. Lykke J. A., Bare L. A., Olsen J., et al. Vascular Associated Gene Variants in Patients with Preeclampsia: Results from the Danish National Birth Cohort. *Acta Obstet Gynecol Scand*. 2012; 91(9): 1053–1060. doi: 10.1111/j.1600-0412.2012.01479.x
12. Liu L., Zhang X., Qin K., et al. Characteristics of Serum Lipid Metabolism among Women Complicated with Hypertensive Disorders in Pregnancy: A Retrospective Cohort Study in Mainland China. *Obstet Gynecol Int*. 2024; 2024: 9070748. doi: 10.1155/2024/9070748
13. Hajagos-Tóth J., Ducza E., Samavati R., et al. Obesity in Pregnancy: A Novel Concept on the Roles of Adipokines in Uterine Contractility. *Croat Med J*. 2017; 58(2): 96–104. doi: 10.3325/cmj.2017.58.96
14. Howlader M., Sultana I., Akter F., Hossain M. Adiponectin Gene Polymorphisms Associated with Diabetes Mellitus: A Descriptive Review. *Heliyon*. 2021; 7(8): e07851. doi: 10.1016/j.heliyon.2021.e07851
15. Zhang G., Zhao J., Yi J. Association Between Gene Polymorphisms on Chromosome 1 and Susceptibility to Pre-Eclampsia: An Updated Meta-Analysis. *Med Sci Monit*. 2016; 22: 2202–2214. doi: 10.12659/MSM.896552
16. Boulanger H., Bounan S., Mahdhi A., et al. Immunologic Aspects of Preeclampsia. *AJOG Glob Rep*. 2024; 4(1): 100321. doi: 10.1016/j.xagr.2024.100321
17. Xu X., Zhou Y., Wei H. Roles of HLA-G in the Maternal-Fetal Immune Microenvironment. *Front Immunol*. 2020; 11: 592010. doi: 10.3389/fimmu.2020.592010
18. Zhuang B., Shang J., Yao Y. HLA-G: An Important Mediator of Maternal-Fetal Immune-Tolerance. *Front Immunol*. 2021; 12: 744324. doi: 10.3389/fimmu.2021.744324
19. Wang X., Zhu H., Lei L., et al. Integrated Analysis of Key Genes and Pathways Involved in Fetal Growth Restriction and Their Associations with the Dysregulation of the Maternal Immune System. *Front Genet*. 2020; 11: 581789. doi: 10.3389/fgene.2020.581789
20. Mora-Palazuelos C., Bermúdez M., Aguilar-Medina M., et al. Cytokine-Polymorphisms Associated with Preeclampsia: A Review. *Medicine (Baltimore)*. 2022; 101(39): e30870. doi: 10.1097/MD.00000000000030870
21. Chen A., Zhao H., Wang J., et al. Haplotype Analysis of Candidate Genes Involved in Inflammation and Oxidative Stress and the Susceptibility to Preeclampsia. *J Immunol Res*. 2020; 2020: 4683798. doi: 10.1155/2020/4683798

CHAPTER 5. POLYGENIC BASIS OF MULTIFACTORIAL ONCOLOGICAL DISEASES IN HUMANS: THE CASE OF BREAST CANCER

INTRODUCTION

Breast cancer is one of the most prevalent forms of cancer worldwide. Between 2015 and 2020, breast cancer was diagnosed in 7.8 million women globally. In 2020 alone, 2.3 million women were diagnosed with the disease, and 685,000 of them succumbed to it. Breast cancer occurs in every country, with the risk of developing the disease increasing with age. Early detection significantly enhances the chances of long-term survival.

The use of risk models integrating standard risk factors, such as family history, hormonal/reproductive factors, mammographic density, and DNA testing, now makes it possible to accurately assess the risk of developing the disease at a population level. Research in cellular and molecular oncology has led to the development of genetic criteria that determine personal risk of developing breast cancer. These advancements influence treatment plans and contribute to individualized recommendations for the prevention and diagnosis of the disease. Genetic profiling allows for the assessment of predisposition to breast cancer, especially in individuals with a positive family history, enabling earlier and more effective screening for carriers of risk-associated gene variants.

Current risk factors for developing breast cancer include:

- Age: According to the International Agency for Research on Cancer, the incidence of breast cancer ranges from 1.5 cases per 100,000 women aged 20-24 to 421.3 cases among those aged 75-79.
- Hormonal treatments: The use of hormonal medications during premenopause and hormone replacement therapy in postmenopausal individuals increases the risk of hormone-positive subtypes of breast cancer.
- Family history and genetic risk factors: Epidemiological studies report a 1.8-fold increase in breast cancer risk in families with one first-degree relative diagnosed with the disease, a 2.93-fold increase with two affected relatives, and a 3.9-fold increase with three or more relatives.
- Risk indicators can vary by ethnicity, with Ashkenazi Jewish women, for instance, having an elevated risk of developing the disease.
- Alcohol consumption and smoking: Moderate alcohol consumption increases breast cancer risk by 1.04–1.61 times. Smoking also heightens the risk, with higher mortality rates seen among active smokers.
- Early-age chest radiation therapy exposure increases breast cancer risk more than fourfold.

POLYGENIC NATURE OF ONCOLOGICAL DISEASES: THE ROLE OF GENE POLYMORPHISM IN THE DEVELOPMENT OF CANCER PATHOLOGIES

The significance of mutations in tumor suppressor genes and proto-oncogenes in the pathogenesis of oncological diseases is widely recognized. Alongside, gene polymorphism can also significantly contribute to susceptibility (increased risk) to cancer. Specific characteristics of an individual's genetic profile can play an important role in determining cancer risk. However, accurately identifying the role of possible genome allele variations in disease pathogenesis is a complex task. Identifying genes that confer a predisposition to this pathology requires consideration of the genetic structure of the specific population and the ethnic backgrounds of the subjects. Human genotype is composed of thousands of interacting polymorphic alleles, each possessing only a moderate individual effect. Currently, hundreds of candidate genes with

polymorphic alleles have been identified, potentially participating in the formation of cancer risk.

Cancer risk is closely associated with the nature of the cellular response to DNA damage. It is reasonable to consider the effectiveness of repair processes as a main factor in cancer resistance. A large number of DNA damages are corrected through the coordinated actions of more than 100 genes in the DNA repair control system. Additionally, a significant portion of kinase enzymes (about 500 genes and more than 1000 proteins) are involved in the cell's response to DNA damage, for instance, by activating p53, balancing cell survival and death, and controlling the mitotic cycle. The apoptosis of cells with severe DNA damage is another cancer resistance mechanism widely employed in the body.

Given these facts, it is logical to hypothesize that studying gene polymorphisms regulating DNA repair mechanisms is a priority in identifying genes associated with cancer predisposition, including breast cancer (BC). Indeed, most BC marker genes are represented by alleles of genes determining DNA repair effectiveness: TP53, BRCA1, BRCA2, BRIP1, ATM, CHEK2, and RAD51D. Furthermore, these genes are functionally interconnected, forming a unified genetic network aimed at preventing genetic mutations and chromosomal instability.

Moreover, the multifactorial nature of diseases implies not only a genetic predisposition to the disease but also an environmental factor that triggers the phenotypic manifestation of the pathological trait. For many oncological diseases, such an environmental factor is diet. The anti-cancer effect of strictly limiting dietary calories is well-known. It is also known that deficiencies in certain vitamins and microelements, such as folic acid, B6, B12, selenium, iron, zinc, and vitamins C and E, increase the risk of developing cancer. Additionally, cancer incidence is observed to be twice as high among populations consuming minimal fruits and vegetables compared to those adhering to a diet balanced with fruits and vegetables.

Non-genetic, most significant risk factors for developing BC include age (over 40 years); high estrogen levels in the blood; the use of hormonal medications for contraception or menstrual cycle regulation, and hormone replacement therapy during menopause.

GENETIC RISK FACTORS FOR BREAST CANCER

Breast cancer is considered a hereditary disease. Epidemiological studies indicate that approximately 4-5% of cases are due to high-penetrance autosomal-dominant predisposition. Results using the twin method suggest that about 27% of cases likely have a strong hereditary component. There are numerous high- and moderate-risk genes associated with the development of breast cancer.

Between 15% to 30% of breast cancer patients have a positive family history. On average, 5% to 10% of cases are considered hereditary, with pathogenic mutations in high-penetrance genes inherited in an autosomal-dominant manner present in 4-5% of these cases.

The contribution of genetic factors to breast cancer development is characterized by the penetrance of disease-associated genes. In medical genetics, penetrance refers to the proportion of individuals with a genetic mutation who manifest clinical symptoms of the disease, compared to those who have the gene but do not exhibit symptoms. Notably, mutations in high-penetrance genes are associated with a more than four-fold increased risk of breast cancer compared to the general population risk. In contrast, pathogenic variants in moderate-penetrance genes indicate a 2-4 times higher risk compared to the general population. Low-penetrance genes are associated with less than a two-fold risk or insufficient evidence of their contribution to disease development. As more clinical information accumulates and new research is published, the list of low-penetrance genes is continually updated.

Single nucleotide polymorphisms (SNPs) are the most common cause of new variants in DNA sequences. There are over 10 million SNPs in an individual's genome. SNPs account for normal variations among individuals and typically have minimal functional impact. However, when an SNP occurs in or near a gene's regulatory area, it can affect gene function and potentially lead to disease development.

Many SNPs individually contribute to a relatively small increase in breast cancer risk. For instance, 14% of women in the United Kingdom who are homozygous for the rs2981582 allele of the FGFR2 gene (a gene that encodes the fibroblast growth factor receptor protein, interacting with growth factors and influencing the mitotic cycle and cell differentiation) have an increased risk (1.38 versus 1.0) of developing breast cancer compared to the average population risk. Meanwhile, 38% of women without this allele have a reduced risk (0.83) of developing breast cancer. This pattern suggests that expression products of different gene alleles may result in either increased susceptibility or resistance to the disease. Notably, SNPs associated with breast cancer are far more common in the population than pathogenic variants found in high- and moderate-risk genes. Over 300 such SNPs have been identified to date, and the risks associated with each amplify the oncogenic effects of one another. Approximately 30% of familial breast cancer cases can be explained by known SNPs. Particular attention is given to SNPs in regulatory regions of genes, which serve as binding sites for transcription factors, in hereditary forms of breast cancer.

Most socially significant diseases, including breast cancer, are multifactorial, with genetic components arising from the combined influence of specific DNA mutations. Identifying new diagnostic markers for such diseases is a complex task. One approach to address this is the Genome-wide Association Study (GWAS).

GWAS is based on identifying SNPs that are more common in individuals with the disease and absent (or rare) in healthy individuals. A set of such mutations is identified using statistical methods or machine analysis by comparing genomes of two groups (affected and healthy individuals). This set of SNPs associated with a particular disease is termed a polygenic risk score (PRS).

Genetic variants leading to breast cancer are identified through GWAS, and based on these findings, polygenic risk scores (PRS) are determined. These scores facilitate the stratification of women into groups for tailored treatment and preventive measures most effective for each specific type of cancer.

According to research findings published in 2019, 313 SNPs associated with breast cancer (BC) were identified, taking into account both ER-positive and ER-negative tumor types (estrogen receptor ER-positive or ER-negative). Many gene variants create risks that vary depending on the subtype of breast cancer. This suggests that polygenic risk scores (PRS) specific to the tumor subtype may allow for better disease prediction specific to that subtype, including the more aggressive ER-negative breast cancer, and allow for the selection of women for preventive treatment. The studies were conducted using DNA samples from 94,075 BC patients and 75,017 control samples. This led to the development of the PRS313 genetic panel for prognosis and diagnosis of BC.

It should be noted that, firstly, PRS are a key component in accurately predicting breast cancer risk and hold potential for improving screening and prevention strategies. The PRS indicator provides an assessment of genetic predisposition to diseases and is typically calculated based on the cumulative effect of multiple disease-associated common genetic variants or single nucleotide polymorphisms (SNPs) obtained from genome-wide association studies (GWAS). In this context, it mainly refers to germline mutations. PRS indices combine the risks from inherited SNPs associated with breast cancer in GWAS and account for over 30% of the heritability of breast cancer. Personalized PRS risk assessment helps to identify women at

higher risk of developing breast cancer and allows for the implementation of individualized screening and prevention approaches.

Secondly, modern molecular biology methods not only enhance the quality of disease risk predictions but also contribute to understanding the molecular basis of tumor progression and predicting its responses to various anticancer therapies. The identification of molecular biomarkers, which may serve as prognostic and predictive markers, assists physicians in making therapeutic decisions, facilitating more personalized treatment approaches, thereby optimizing therapy and avoiding erroneous treatment strategies. For example, prognostic genetic markers can help doctors predict the aggressiveness and invasiveness of the tumor, enabling the use of more optimal treatment strategies.

GENETIC MARKERS FOR BREAST CANCER

The highest risk of developing breast cancer (BC) is associated with genes of high penetrance, characterized by a more than five-fold increase in the risk of the pathology compared to the general population. These include BRCA1, BRCA2, TP53, CDH1, PTEN, and STK11 (LKB1).

Genes of moderate penetrance associated with the development of BC include ATM and CHEK2, which are involved in the repair of double-strand DNA breaks, BRCA1 and BRCA2 modifiers—BRIP1 (BACH1) and PALB2 (FANCN). Mutations in moderately penetrant genes occur in 2–5% of BC cases and can increase the likelihood of developing BC by 2–4 times. However, some mutations in moderately penetrant genes, particularly PALB2 and CHEK2, can confer a risk of BC as high for individuals with a positive family history as mutations in high penetrance genes.

The current trends in determining individual cancer risk, and specifically BC, involve integrating polygenic components into existing mathematical models. Risk assessment that includes the determination of the sum of all potentially hazardous variants can provide the most accurate scheme for identifying individuals with an increased genetic risk of developing BC.

Modern methods of cancer risk assessment and diagnosis widely use multigene panels based on next-generation sequencing technology. Below are the characteristics of certain genes most relevant in the analysis of breast cancer.

The TP53 gene encodes the tumor suppressor protein p53, which contains domains for transcriptional activation, DNA binding, and oligomerization. The encoded protein responds to various cellular stresses by regulating the expression of target genes, thereby causing cell cycle arrest, apoptosis, senescence, DNA repair, or metabolic changes. Mutations in this gene are associated with various types of human cancers, including hereditary cancers such as Li-Fraumeni syndrome. Germline mutations in TP53 are associated with an increased risk of developing many cancer types, including BC. The frequency of somatic mutations in TP53 is higher among the basal subtype of BC. Tumors with TP53 mutations are characterized by a more aggressive disease course and resistance to chemo- and radiotherapy. Notably, the TP53 gene holds great potential as a molecular marker for BC risk, similar to BRCA1. Both tumor suppressor genes are potential biomarkers for monitoring, early risk assessment, and predisposition to BC, serving as therapeutic targets for chemoprevention and targeted therapy. In breast cancer, the TP53 gene is most frequently mutated, present in approximately 30–35% of cases. The intensity of TP53 mutations varies depending on the subtypes of breast cancer, reaching up to 80% in the most dangerous form of BC—triple-negative breast cancer (TNBC, characterized by the absence of three key receptors: estrogen, progesterone, and HER2) and 70% in HER2-positive breast cancer cases. TP53 mutations are significant biomarkers in clinical practice and represent potential therapeutic targets. The nature of TP53 mutations is

determined through DNA sequencing. Tumors with TP53 mutations have long been characterized by therapeutic resistance. However, new strategies in targeted therapy are being developed that can restore the properties of wild-type p53, providing new opportunities in cancer treatment. Specifically, derivatives of 3-quinuklidinone have the ability to reactivate mutant p53, inducing a therapeutic apoptotic response in TNBC tumor cells.

The BRCA1 and BRCA2 genes are tumor suppressor genes that play a fundamental role in DNA repair through homologous recombination. Germline mutations in these genes are associated with an increased risk of developing breast and ovarian cancer. The average cumulative risk of breast cancer by age 80 is 72% for BRCA1 and 69% for BRCA2. Loss of function in these genes leads to ineffective DNA repair, increasing mutation rates and contributing to tumor development. Patients with pathogenic or likely pathogenic variants in BRCA1 are predisposed to develop triple-negative breast cancer (TNBC), while pathogenic or likely pathogenic variants in BRCA2 are more commonly associated with ER-positive tumors. There is ongoing intensive research into the value of laboratory analysis of BRCA gene mutations in predicting patient survival and metastasis risks. Tumors with pathogenic BRCA1/BRCA2 mutations exhibit greater sensitivity to both chemotherapy and targeted therapy agents.

The PTEN gene, a homolog of phosphatase and tensin, is a tumor suppressor gene. The protein encoded by this gene is a phosphatidylinositol-3,4,5-triphosphate 3-phosphatase. It contains a tensin-like domain as well as a catalytic domain analogous to the dual-specificity tyrosine protein phosphatase domain. Unlike most tyrosine protein phosphatases, this protein primarily dephosphorylates phosphoinositide substrates rather than proteins. It negatively regulates intracellular levels of phosphatidylinositol-3,4,5-triphosphate in cells and functions as a tumor suppressor by inhibiting the PI3K/Akt signaling pathway. Pathogenic genetic variants in the PTEN phosphatase gene are associated with an increased risk of several cancers, including colorectal cancer, kidney cancer, melanoma, and breast cancer. Mutations in the promoter region of the PTEN gene are associated with a significant risk of breast cancer, endometrial cancer, and thyroid cancer. This gene is one of the most frequently mutated in cancers, including breast cancer, and its role is closely connected with cell cycle progression, cell growth, and survival. Tumor cells with loss of heterozygosity or mutations in PTEN exhibit significantly increased migratory and invasive activity, promoting proliferation, invasion, and metastasis. In metastatic breast cancer cells, PTEN levels are much lower than in localized cancer cells. Loss of PTEN function leads to excessive activation of the PI3K/Akt oncogenic signaling pathway, which promotes cell growth and survival. Therefore, PTEN loss of activity has been recorded in nearly half of all breast cancer cases. Inactivation of PTEN primarily occurs due to somatic mutations. PTEN loss can often indicate tumor resistance to various treatment methods.

The CHEK2 gene, also known as checkpoint kinase 2, encodes the serine/threonine kinase protein CHK2, which is involved in DNA repair mechanisms. It functions as an important tumor suppressor gene in regulating the cell cycle, inhibiting cell proliferation, activating DNA repair, and apoptosis. Abnormal expression of CHEK2 can result in breast cancer, among other things. It has been established that certain pathogenic variants of the CHEK2 gene, such as c.470T>C (p.Ile157Thr), lead to reduced binding of the CHEK2 protein with the Cdc25A, BRCA1, and p53 proteins, but do not affect the protein kinase activity of the CHEK2 protein. Germline mutations in CHEK2 are linked to susceptibility to several cancer types, with a mutation frequency of 1.08% among breast cancer patients. A relatively widespread pathogenic variant of the CHEK2 gene, c.349A>G (p.Arg117Gly), results in more than a twofold increase in breast cancer risk. This variant has been identified in many cancer cases, including familial cases of breast/ovarian cancer, prostate cancer, pancreatic cancer, colorectal cancer, as well as

other cancer types. This pathogenic CHEK2 variant leads to the loss of protein kinase activity, which increases the risk of developing breast cancer. Most patients with pathogenic or likely pathogenic CHEK2 variants develop luminal A breast cancer subtypes (characterized by the presence of estrogen receptors (ER) and progesterone receptors (PR), negative HER2 status, and low Ki67, being the most commonly diagnosed breast tumor subtype) or luminal B subtypes (where estrogen receptors are present, progesterone receptors may be absent, and the Ki67 level is high, occurring in 20% of breast cancer cases).

The gene ATM, responsible for ataxia-telangiectasia, encodes a protein that belongs to the PI3/PI4-kinase family. This protein is a crucial kinase at the cell cycle checkpoint, functioning as a regulator of various downstream proteins, including tumor suppressors p53 and BRCA1, checkpoint kinase CHK2, checkpoint proteins RAD17 and RAD9, and the DNA repair protein NBS1. The ATM gene product, along with the closely related ATR kinase, is considered a primary controller of cell cycle checkpoint signaling pathways essential for cellular responses to DNA damage and genomic stability. A common gene associated with breast cancer susceptibility is the mutated ATM gene. Pathogenic or likely pathogenic variants of this gene are linked to the repair mechanism of DNA double-strand breaks. The mutation frequency of this gene in breast cancer patients is 0.78%. The ATM gene expresses proteins involved in DNA repair and cell cycle regulation, critical in cellular stress situations and response to DNA damage. Truncated and missense variants of the ATM gene are associated with a higher risk of estrogen receptor-positive disease compared to estrogen receptor-negative disease. The ATM missense variant c.7271T>G (p.Val2424Gly) carries a significant risk for breast cancer, comparable to the risk of mutations in the BRCA2 gene. Many breast cancer patients with ATM mutations develop severe disease forms with a high likelihood of lymph node metastasis, more aggressive tumors, and poor prognosis. A reduction in ATM levels in cancer cells may enhance sensitivity to targeted therapy and sensitize cancer cells to platinum-based drugs. However, ATM mutations increase the risk of tumor recurrence following radiation therapy.

The PALB2 gene (Partner and Localizer of BRCA2) is a tumor suppressor involved in maintaining genomic integrity. This protein binds and colocalizes with the Breast Cancer Type 2 Susceptibility Protein (BRCA2) in the nucleoplasm, likely ensuring stable intranuclear localization and accumulation of BRCA2. PALB2 is critically important for resolving DNA double-strand breaks. Pathogenic variants lead to an increased risk of developing breast cancer. The cumulative risk of breast cancer in patients with a germline PALB2 mutation by age 70 is 35%, with a ten-year survival rate lower compared to patients without PALB2 mutations. The PALB2 gene is one of eight genes frequently mutated in metastatic breast cancer. The PALB2 missense variant c.2586 + 58C>T was associated with a 36% increased risk of breast cancer and is included in genetic panels for breast cancer diagnosis and screening. Patients with breast cancer who carry pathogenic PALB2 variants are more frequently observed to have a triple-negative phenotype and lower survival rates. Platinum-based therapeutic regimens associated with PARP inhibitors have significant potential in patients with germline PALB2 variants. Approximately 0.5% of breast cancer patients carry pathological mutations in the PALB2 gene. Although PALB2 was previously considered a moderate-penetrance breast cancer risk gene, recent data suggest that PALB2 should be classified as a high-risk gene. Large population studies allow the assessment of breast cancer risk in carriers of pathogenic PALB2 mutations as comparable to those with pathogenic mutations in the BRCA2 gene.

The BRIP1 gene (Breast Cancer 1 Interacting Protein 1 helicase) is necessary for DNA repair and genomic stability. The BRIP1 and BRCA1 protein complex is crucial for the normal function of DNA double-strand break repair. Mutated or overexpressed BRIP1 is associated with the onset of breast cancer and is recognized as a candidate for tumor progression. BRIP1 is located on the long arm of chromosome 17 and encodes a protein belonging to the RecQ DEAH

helicase family, which helps repair damaged DNA by interacting with BRCA1. Consequently, if BRIP1 is mutated, its protein does not interact with BRCA1 and cannot repair damaged DNA. Therefore, BRIP1 plays a vital role in preserving the genetic information of the cell and acts as a tumor suppressor. BRIP1 expression was significantly increased in 13 types of cancer, including urothelial carcinoma of the bladder (BLCA), invasive breast carcinoma (BRCA), cervical squamous cell carcinoma and endocervical adenocarcinoma (CESC), colon adenocarcinoma (COAD), diffuse large B-cell lymphoma (DLBC), esophageal carcinoma (ESCA), glioblastoma multiforme (GBM), lung squamous cell carcinoma (LUSC), rectal adenocarcinoma (READ), stomach adenocarcinoma (STAD), thymoma (THYM), uterine corpus endometrial carcinoma (UCEC), and uterine carcinosarcoma (UCS). The BRIP1 gene may be involved in breast cancer oncogenesis in families without BRCA1/BRCA2 mutations. A link has been established between rare missense mutations in BRIP1 and two BRIP1 single nucleotide polymorphisms with breast cancer, giving this gene prognostic significance. Overexpression of BRIP1 is associated with promoter methylation status, determining survival in breast cancer patients. Thus, BRIP1 may be not only a predictive molecular biomarker for breast cancer development and prognosis but also a potential therapeutic target.

The CDH1 gene encodes the E-cadherin-1 cell adhesion molecule, which inhibits the spread of tumor cells (metastasis). E-cadherin-1 is a calcium-dependent intercellular adhesion molecule and a tumor suppressor protein, recognized as the only germline molecular defect associated with hereditary diffuse gastric cancer and lobular breast cancer. The E-cadherin-1 protein plays a crucial role in establishing and maintaining a polarized and differentiated epithelium through intercellular adhesive complexes. This molecule is considered an invasion suppressor, and its regulatory disruption is often found in advanced cases of certain epithelial carcinomas. Mutations in this gene correlate with cancers of the stomach, breast, colon, thyroid, and ovaries. The loss of function of this gene is believed to contribute to cancer progression by increasing proliferation, invasion, and metastasis. The ectodomain of this protein mediates bacterial adhesion to mammalian cells, while the cytoplasmic domain is essential for internalization. Reduced E-cadherin expression is associated with cancer metastasis due to loss of cell adhesion, leading to increased cell motility, allowing cancer cells to cross the basement membrane and invade adjacent tissues. Patients with CDH1 promoter hypermethylation have a 5.83-fold increased risk of breast cancer. CDH1 promoter hypermethylation is commonly elevated in HER2-negative and ER-negative breast cancer, irrespective of PR status. It is suggested that CDH1 is the most frequently methylated gene (90%) in cases with lymph node metastases, confirming the relationship between CDH1 hypermethylation and metastasis. CDH1 hypermethylation may potentially be used as a biomarker for the metastatic potential of tumors. The CDH1 gene is a potential new target for drugs aimed at inhibiting DNA methylation.

The BARD1 gene (BRCA1-associated RING domain 1) represents a partner protein that binds to BRCA and is necessary for DNA damage repair associated with breast cancer susceptibility. The BARD1/BRCA1 interaction is disrupted by tumorigenic amino acid substitutions in BRCA1. Therefore, the formation of a stable complex between these proteins may be a critical factor in BRCA1 tumor suppression. This protein could be a target for oncogenic mutations in breast or ovarian cancer. Germline mutations in the BARD1 gene have been significantly associated with a high risk of TNBC – a subtype of breast cancer characterized by the absence of three key receptors: estrogen, progesterone, and HER2. Studies on laboratory animals have shown that partial repression of Bard1 using antisense RNA leads to the development of early-stage malignancy phenotypes in mammary epithelial cells. Conversely, breast cancer tumor cells may express elevated levels of BARD1 and BRCA1, leading to chemotherapy resistance. Moreover, higher expression of BARD1 and BRCA1 is associated

with a poorer prognosis for patients with early-stage breast cancer, particularly those who received radiation therapy, indicating the potential use of PI3K inhibitors to overcome chemoresistance and radio-resistance in ER-positive breast cancer patients. Thus, the BARD1 gene may play a significant role in the pathogenesis of breast cancer and mechanisms of chemoresistance. It is suggested that BARD1 expression levels in breast cancer could be closely related to TNBC.

The PIK3CA gene encodes phosphatidylinositol-3-kinase (PI3K), which regulates proliferation and apoptosis, with somatic mutations in PIK3CA potentially activating these processes. PI3K is involved in several cellular processes such as protein synthesis, cell proliferation, survival, glucose homeostasis, and DNA repair. PIK3CA mutations are among the most common types in breast cancer. Activating PIK3CA mutations are detected in approximately 30–40% of cancer patients, causing hyperactivation of the α -PI3K isoform. These mutations are associated with chemoresistance and poor prognosis, with reduced overall survival. Furthermore, PIK3CA mutations in breast cancer are oncogenic and play a vital role in tumor pathogenesis and progression.

The STK11 (LKB1) gene product is the LKB1 protein, which interacts with the p53 protein and regulates specific p53-dependent apoptotic pathways. The LKB1 protein is present in both the cytoplasm and nucleus of living cells and translocates to mitochondria during apoptosis. A close functional association between LKB1 and PTEN proteins is observed. Some point mutations in LKB1 disrupt interaction with PTEN, suggesting that loss of this interaction may contribute to increased cancer risk. There is a significant reduction in STK11 gene expression levels in cancer cells. Reduced gene expression is observed in many cancers: lung, stomach, intestinal, pancreatic, and breast cancer. A link has been established between estrogen-positive and HER2-negative breast cancer in families with STK11 gene mutations.

The RAD51D gene participates in DNA repair mechanisms through homologous recombination. It plays a crucial role in maintaining genomic stability, and mutations in this gene may be linked to carcinogenesis. Pathogenic variants of RAD51D have been identified in patients with breast cancer through gene panel testing. The presence of such pathogenic forms of the RAD51D gene in an individual's genotype accounts for more than 20% of the estimated lifetime risk of developing breast cancer. There is a correlation between pathogenic variants of the gene and triple-negative breast cancer (TNBC). It is known that the pathological variant RAD51D c.270_271insTA (in the presence of normal BRCA1/BRCA2 genes) causes highly malignant TNBC with a high risk of metastasis to lymph nodes. The pathological form RAD51D 270_271insTA has been observed in individuals with a history of breast and ovarian cancer and is often detected in familial forms of breast and ovarian cancer. Another oncogenic mutation of the gene (c.273del; p.Lys91fs) has been identified. Currently, the link between pathogenic variants of germline mutations in RAD51D and the risk of breast cancer is confirmed. It has also been shown that pathological variants of RAD51D have a strong association with ER-negative breast cancer and TNBC.

RECOMMENDED LITERATURE FOR CHAPTER 5. BREAST CANCER

1. Zolotykh M.A., Bilyalov A.I., Nesterova A.I., et al. Breast Cancer: Genetic Factors of Personal Risk. // *Modern Oncology*. 2023; 25(2): 190–198.
2. Smirnova N.V., Elmuradov A.U. Prospects for Using the Polygenic Risk Score (PRS) for Assessing Breast Cancer Risk. // *Bulletin of Surgut University. Medicine*. 2019, 42(4): 74–78.

3. Mavaddat N., Michailidou K., Dennis J., et al. Polygenic Risk Scores for Prediction of Breast Cancer and Breast Cancer Subtypes. *Am J Hum Genet.* 2019; 104(1): 21–34. doi: 10.1016/j.ajhg.2018.11.002
4. Graffeo R., Rana H.Q., Conforti F., et al. Moderate Penetrance Genes Complicate Genetic Testing for Breast Cancer Diagnosis: ATM, CHEK2, BARD1 and RAD51D. *Breast.* 2022; 65: 32–40. doi: 10.1016/j.breast.2022.06.003
5. Roberts E., Howell S., Evans D.G. Polygenic Risk Scores and Breast Cancer Risk Prediction. *Breast.* 2023; 67: 71–77. doi: 10.1016/j.breast.2023.01.003
6. Rebello Alves L.N., Meira D.D., Meriguetti L.P., et al. Biomarkers in Breast Cancer: An Old Story with a New End. *Genes (Basel).* 2023; 14(7): 1364. doi: 10.3390/genes14071364
7. Karami F., Mehdipour P. A Comprehensive Focus on Global Spectrum of BRCA1 and BRCA2 Mutations in Breast Cancer. *Biomed Res Int.* 2013; 2013: 928562. doi: 10.1155/2013/928562
8. Polyak K. Breast Cancer: Origins and Evolution. *J Clin Invest.* 2007; 117(11): 3155–3163. doi: 10.1172/JCI33295
9. Sun Y.-S., Zhao Z., Yang Z., et al. Risk Factors and Preventions of Breast Cancer. *Int J Biol Sci.* 2017; 13(11): 1387–1397. doi: 10.7150/ijbs.21635
10. Schon K., Tischkowitz M. Clinical Implications of Germline Mutations in Breast Cancer: TP53. *Breast Cancer Res Treat.* 2018; 167(2): 417–423. doi: 10.1007/s10549-017-4531-y
11. Waarts M.R., et al. Targeting Mutations in Cancer. *J Clin Invest.* 2022; 132(8): e154943. doi: 10.1172/JCI154943
12. Brett J.O., et al. ESR1 Mutation as an Emerging Clinical Biomarker in Metastatic Hormone Receptor-Positive Breast Cancer. *Breast Cancer Res.* 2021; 23: 85. doi: 10.1186/s13058-021-01462-3
13. Dustin D., Gu G., Fuqua S.A.W. ESR1 Mutations in Breast Cancer. *Cancer.* 2019; 125(21): 3714–3728. doi: 10.1002/cncr.32345
14. Silwal-Pandit L., Langerød A., Børresen-Dale A.-L. TP53 Mutations in Breast and Ovarian Cancer. *Cold Spring Harb Perspect Med.* 2017; 7(1): a026252. doi: 10.1101/cshperspect.a026252
15. Kuusisto K.M., Bebel A., Vihinen M., et al. Screening for BRCA1, BRCA2, CHEK2, PALB2, BRIP1, RAD50, and CDH1 Mutations in High-risk Finnish BRCA1/2-founder Mutation-Negative Breast and/or Ovarian Cancer Individuals. *Breast Cancer Res.* 2011; 13(1): R20. doi: 10.1186/bcr2832
16. Yoshimura A., Imoto I., Iwata H. Functions of Breast Cancer Predisposition Genes: Implications for Clinical Management. *Int J Mol Sci.* 2022; 23(13): 7481. doi: 10.3390/ijms23137481
17. Long G., Hu K., Zhang X., et al. Spectrum of BRCA1 Interacting Helicase 1 Aberrations and Potential Prognostic and Therapeutic Implications: A Pan Cancer Analysis. *Sci Rep.* 2023; 13: 4435. doi: 10.1038/s41598-023-31109-6
18. Mahdi K.M., Nassiri M.R., Nasiri K. Hereditary Genes and SNPs Associated with Breast Cancer. *Asian Pac J Cancer Prev.* 2013; 14(6): 3403-409. doi: 10.7314/apjcp.2013.14.6.3403.
19. Hasson S.P., Menes T., Sonnenblick A. Comparison of Patient Susceptibility Genes Across Breast Cancer: Implications for Prognosis and Therapeutic Outcomes. *Pharmgenomics Pers Med.* 2020; 13: 227–238. doi: 10.2147/PGPM.S233485
20. Corso G., Figueiredo J., De Angelis S.P., et al. E-cadherin Deregulation in Breast Cancer. *J Cell Mol Med.* 2020; 24(11): 5930–5936. doi: 10.1111/jcmm.15140

CHAPTER 6. NUTRITIONAL GENETICS

NUTRIGENETICS, NUTRIGENOMICS, AND DIETARY EPIGENETICS

Nutrigenetics and nutrigenomics are scientific fields focusing on the role of individual genetic variations in the body's response to nutrients and the role of diet and bioactive substances in gene expression, respectively. It is important to differentiate between the terms nutrigenomics and nutrigenetics, as although they are closely related, they are not interchangeable.

Nutrigenetics examines the modifying effects of germline (or somatic in cases of cancer pathology) mutations in genes related to nutrition effectiveness, microelement absorption, and the impact of diet on health. Nutrigenetics assesses the role of gene variants in the body's response to nutrient intake. While interconnected, nutrigenetics and nutrigenomics are distinct; the main difference is that nutrigenetics investigates the influence of hereditary genetic variants on micro-nutrient absorption and metabolism, whereas nutrigenomics explores the relationship between the genome and diet, highlighting the effect of dietary patterns on metabolic, proteomic, transcriptional, and translational changes.

Nutrigenetics research has identified genetic variants associated with disease susceptibility through interactions with specific dietary factors. For instance, gene variants involved in metabolic pathways influence the consumption and utilization of various micronutrients. In this context, nutrigenetic studies can be used to predict the risk of different polygenic diseases. Consequently, personalized nutrition can be utilized to prevent or more effectively manage these diseases. It is proposed that the study of all mentioned effects be conducted independently or comprehensively for diagnostic or prognostic purposes regarding disease progression.

As previously mentioned, nutrigenetics encompasses the impact of genetic variations on dietary responses and nutrient functionality. However, among the aforementioned genetic markers in nutrigenetics, only structural DNA molecule alterations can be considered true pathology markers. Such DNA structural disruptions include: (1) damage to individual bases (e.g., hydroxyl radical attachment to guanine due to oxidative stress); (2) nucleotide loss of nitrogenous base (apurinic/apyrimidinic sites); (3) DNA strand breaks; (4) telomere shortening; (5) deletions or chromosomal losses; (6) mitochondrial DNA damage (deletions or base damage in circular mitochondrial DNA). It is important to note that not only variations in the primary DNA structure within a gene in the form of SNPs are significant, but also dynamic mutations and gene copy number variants. Notably, a connection between copy number variations (CNV) for small genomic regions and metabolic disease risk has been identified. Examples of such mutations include: 1) leptin receptor gene copy number variations linked to metabolic traits and type 2 diabetes risk; 2) a lower copy number of the salivary alpha-amylase gene associated with obesity predisposition; 3) a pentanucleotide in/del (CTTTA) in the 3'-untranslated region of the leptin receptor gene connected with type 2 diabetes risk.

The field of nutrigenomics encompasses several disciplines and includes the effects of dietary patterns on genome stability (DNA damage at the molecular and chromosomal levels), epigenomic changes (DNA methylation), mRNA and microRNA transcription (transcriptomics), protein expression (proteomics), and metabolite alterations (metabolomics). Nutrigenomics is a new area of molecular biology where advanced genomic tools are used to analyze the effects of nutrients on the genome and gene expression, as well as the influence of genetic variants on nutrient intake. The term "nutrigenomics" was coined to describe the interaction between nutrients and genes. The fundamental issues that nutrigenomics seeks to address stem from the rich individual and interethnic diversity affecting nutrient bioavailability and metabolism, significant individual differences determining food availability and variety due to cultural,

economic, geographical, and taste differences, and imbalanced nutrition (nutrient deficiency or excess) leading to adverse events for humans that result in gene expression disruption, genome instability, intensified mutational processes, which ultimately may lead to the manifestation of pathological phenotypes at various life stages.

The Fundamental Principles of Nutrigenomics.

1. Nutrition can influence health by directly affecting gene expression through metabolic processes in the body, thereby impacting the intensity of genetic and chromosomal mutations. Such events can alter gene dosage and expression.

2. The effect of consumed foods on human health is significantly determined by genetic variants, which alter the absorption and metabolism of nutrients. This factor is regulated on a molecular level through the interaction of enzymes in metabolic processes and nutrient cofactors.

3. Sustainable health indicators can be ensured if the diet is tailored individually, taking into account inherited and acquired genetic characteristics, depending on the stage of ontogenesis, dietary preferences, and health status. Contemporary medical daily recommended nutrient intakes are generally based on various metabolic studies and are not optimized for genetic subgroups. These approaches in developing recommendations can vary significantly.

The ultimate goal is to: (1) align the nutrionome (the set of consumed foods) with the current genome status to ensure the genome maintenance, gene expression, metabolism, and cell function occur in a homeostatically balanced manner; and (2) create conditions for more favorable health trends in the population by forming personalized nutrition recommendations.

Nutritional Epigenetics.

Nutritional epigenetics is a field studying the role of foods in epigenetic changes. Epigenetics encompasses both reversible and heritable changes in gene expression not associated with changes in the DNA coding sequence. In this context, epigenetic changes can substantially influence the pathogenesis, methods of correction, and treatment of a range of polygenic diseases. For instance, the anti-inflammatory effects of a diet rich in polyunsaturated fatty acids are due to inhibitory hypermethylation of pro-inflammatory gene promoters by DNA methyltransferase enzymes. Furthermore, high intake of polyunsaturated fatty acids positively regulates the expression of specific microRNAs, which inhibit lipogenic and oncogenic genes. The plant-derived compound curcumin is not only a potent antioxidant but also directly involved in epigenetic mechanisms. Specifically, as an epigenetic regulator, curcumin exerts protective effects against heart failure and liver damage through the DNA methylation and histone modification enzyme systems. These data indicate that incorporating certain dietary compounds into the human diet that modulate epigenetic processes may be an effective strategy aimed at reducing the prevalence of chronic diseases. It should be emphasized that disruption of epigenetic mechanisms may be the cause of various chronic diseases and their progression. Notably, complex interactions between nutrients and epigenetic mechanisms may participate in the pathogenesis of obesity, diabetes, dyslipidemia, cardiovascular diseases, non-alcoholic fatty liver disease, and oncological conditions. A diet based on excessive intake of animal fats and readily available carbohydrates may be associated with abnormal methylation processes of neuropeptide genes that control the behavioral responses determining food intake and may be involved in the development of obesity and type 2 diabetes. Additionally, a deficiency in certain trace elements and vitamins, such as vitamin A, B vitamins, selenium, potassium, and iron, is associated with hypermethylation of tumor suppressor genes, which play a crucial role in cancer development.

GENETIC BASIS OF PERSONALIZED NUTRITION

Personalized nutrition is a fundamentally new approach in dietetics, where individual genetic profiles dictate personalized dietary recommendations. This approach supports active longevity and facilitates the effective correction of chronic diseases. The phenomenon of human gene polymorphism is explained by differences at the primary DNA structure level, which can contribute to the diversity of phenotypes observed in human populations. Types of genetic variants include SNPs, insertions and deletions (indels), copy number variations (CNVs), and more complex structural changes.

Gene polymorphisms significantly impact nutrient metabolism. For instance, variants of the methylenetetrahydrofolate reductase (MTHFR) gene can affect folate metabolism, which can be linked to cardiovascular diseases and diabetes. Individuals with specific SNPs in this gene may have different responses to folate intake, which can be managed with personalized folate supplementation. Homozygotes with the MTHFR mutation (TT) (MTHFR Ala222Val polymorphism, C>T) have an increased folate requirement due to disruptions in its metabolism, necessitating higher folate intake than current dietary guidelines suggest to mitigate disease susceptibility related to folate deficiency. Additionally, those with Down syndrome (having an extra copy of chromosome 21) show increased needs for zinc and folate due to the enhanced expression of cystathionine- β -synthase, a key enzyme in folate metabolism. Moreover, common gene variants regulating homocysteine metabolism, such as MTHFR and methionine synthase (MTR), are associated with an increased risk of breast cancer in individuals with inadequate intake of vitamin B6, vitamin B12, and folate.

Genetic variations can also influence selenium metabolism. Some SNPs affect protein expression or selenoprotein activity in response to selenium supplementation. Variations in the beta-carotene oxygenase 1 (BCMO1) gene, essential for beta-carotene metabolism, may lead to variability in plasma carotenoid levels and possess significant clinical implications, such as in the development of liver steatosis independent of dietary vitamin A. Supplementing vitamin A is a primary method for correcting its deficiency. Understanding gene variants related to vitamin A metabolism allows for more effective and safer treatment of vitamin A deficiency across populations. Gene variants involved in lipid metabolism, such as cholesterol ester transfer protein, lipoprotein lipase, low-density lipoprotein receptor, and apolipoprotein E, may contribute to the risk of coronary artery disease and type 2 diabetes. SNPs in genes associated with choline and folic acid metabolism can determine dietary choline requirements, critical for liver function and fetal development. The examples mentioned emphasize the importance of considering genetic factors in dietary planning to optimize health outcomes based on individual genetic profiles.

The role of polymorphisms in the CYP1A2 gene (cytochrome P450 1A2) is established. This gene encodes an enzyme central to caffeine metabolism, playing a potential role in the pathogenesis of cardiovascular diseases. High caffeine intake has been linked to increased cardiovascular disease risk in individuals whose gene variants encode a "slow" caffeine metabolism enzyme. In contrast, individuals with genetic variants associated with fast caffeine metabolism are protected from the effects of moderate caffeine consumption.

The presence of a G>A substitution in the APOA1 gene (the main protein component of high-density lipoproteins (HDL) in blood plasma) results in increased HDL cholesterol levels after consuming long-chain omega-3 polyunsaturated fatty acids (PUFAs) in individuals with the A allele, whereas this effect is not observed in those with the GG genotype.

The SNP risk factor in the serum- and glucocorticoid-induced kinase 1 (SGK1) gene is due to the kinase activating potassium, sodium, and chloride channels, thus controlling renal sodium excretion. High expression levels of this gene can contribute to conditions like hypertension and diabetic nephropathy. Consequently, polymorphism in this gene is one of the reasons for elevated blood pressure when consuming a high-salt diet. On the other hand, the development

of personalized diets according to identified ACE (angiotensin-converting enzyme) genotypes may include recommendations for sodium chloride intake compared to standard dietary guidelines for the population.

Single nucleotide polymorphisms (SNPs) in the VDR gene (vitamin D receptor) affect vitamin D absorption and may be associated with osteoporosis predisposition in postmenopausal individuals with insufficient calcium intake.

PRACTICAL APPLICATION OF GENETIC MARKERS IN FOOD CONSUMPTION RECOMMENDATIONS

Diet plays a direct role in influencing human health. Specifically, a dietary pattern characterized by a high intake of processed grains, meats, sweets, and desserts (the Western diet) can lead to a genetic profile that promotes carcinogenesis and inflammatory responses. Furthermore, high consumption of animal protein increases the risk of colorectal cancer. However, diets with insufficient protein content may enhance the expression of liver gluconeogenesis genes, leading to impaired glucose tolerance. High levels of animal fat consumption can result in a gene expression profile associated with glucose intolerance, lipid deposition in the liver, inflammation induction, and increased expression of neuropeptides, leading to obesity. Conversely, consuming balanced amounts of protein, unprocessed grains, mono- and polyunsaturated fatty acids, fruits, and vegetables can counteract these adverse effects. High intake of monounsaturated fatty acids (such as olive oil) is associated with reduced expression of genes involved in inflammation and lipid deposition. Diets rich in polyunsaturated fatty acids positively regulate the expression of neuropeptide genes that control energy homeostasis.

Genetic abnormalities in metabolic processes arise from mutations in the relevant genes encoding these enzymes. Such pathogenic gene variants qualitatively alter the body's nutritional needs. Classic examples of these hereditary anomalies are lactose intolerance and phenylketonuria. The T>C-13910 variant in the lactase gene (LCT) leads to instability or absence of the lactase enzyme in adulthood, rendering individuals with this genetic variant unable to digest lactose. In contrast, phenylketonuria is an autosomal recessive disorder caused by mutations in the phenylalanine hydroxylase gene (PAH), a crucial liver enzyme responsible for converting phenylalanine to tyrosine.

Dietary models that recommend grains, vegetables, and fruits while limiting saturated fat consumption have proven effective in reducing obesity risk in individuals with high genetic predisposition to obesity, especially those with risk alleles in the FTO gene, such as rs9939609, rs1121980, and rs1421085. These FTO gene variants necessitate a diet rich in grains, vegetables, and fruits with minimal animal fat content. Moreover, the FTO gene variant (rs1421085) predisposes individuals to increased protein consumption, independent of body mass index.

Therefore, FTO gene polymorphisms pose risks for individuals suffering from obesity. For such individuals, a diet with reduced animal fat intake and increased physical activity is recommended. The presence of the (A) allele in the CAB39 gene variant rs6722579 is associated with enhanced susceptibility to abdominal obesity in those exceeding recommended fat intake. Conversely, individuals with the (T) allele of the CPQ gene rs59465035 exhibit reduced susceptibility to abdominal obesity, particularly with increased vitamin C intake.

For individuals with folic acid deficiency (MTHFR gene polymorphism), a diet high in folic acid-rich foods (citrus fruits, bananas, legumes, greens, eggs, liver) is recommended, leading to a marked reduction in cardiovascular disease risk. The APOE gene polymorphism is linked to cardiovascular disease risk, recommending reduced animal fat consumption and

increased omega-3 fatty acid intake. TCF7L2 and PPARG gene polymorphisms are associated with diabetes risk, necessitating the limitation of high-carbohydrate food intake.

GENETIC FOUNDATIONS OF BODY WEIGHT CONTROL

Obesity is a global public health issue, serving as a predictor for dangerous chronic diseases worldwide, including type 2 diabetes, cardiovascular diseases, metabolic syndrome, neurodegenerative diseases, and certain types of cancers. Statistics from 2017–2018 illustrate the gravity of this issue in the United States, indicating that 42.4% and 9.2% of adult citizens suffer from obesity and severe obesity, respectively.

At first glance, the solution to obesity seems straightforward—maintaining a balance between calories consumed and expended. However, this is not the case. A key feature of obesity is the disruption of the physiological mechanisms regulating energy metabolism. The central nervous system (CNS) and the endocrine system play pivotal roles in regulating appetite, satiety, and energy expenditure, with the hypothalamus serving as the principal regulator of these processes. The arcuate nucleus of the hypothalamus (ARC), or the lateral hypothalamus, is the central hub for controlling satiety processes. The ARC contains neurons that express functionally antagonistic peptides: orexigenic peptides, which stimulate appetite, such as neuropeptide Y (encoded by the NPY gene) and agouti-related peptide (an antagonist of melanocortin-3 and melanocortin-4 receptors encoded by the AGRP gene), and anorexigenic peptides, which suppress appetite, such as proopiomelanocortin (encoded by the POMC gene) and cocaine and amphetamine-regulated transcript (encoded by the CARTPT gene). Notably, polymorphism of these genes in humans may lead to metabolic disorders, including obesity. Peripheral signals involved in food intake regulation include hormones such as leptin, insulin, and gut hormones (cholecystokinin, encoded by the CCK gene; ghrelin, encoded by the GHRL gene; peptide YY, encoded by the PYY gene; glucagon-like peptide-1, GLP1, synthesized from the proglucagon peptide during post-translational processing in intestinal endocrine L-cells, and other physiologically active peptides). The integration of hormonal and nutritional signals from the circulatory system and peripheral and central neuronal signals ultimately occurs in the ARC, forming a response to the degree of food satiety.

The state of components at both levels of energy metabolism control, consumption, and expenditure can be influenced by gene polymorphism, epigenetic modifications, the state of the microbiome, lifestyle, and social and psychological factors, ultimately leading to obesity. Therefore, it is important to consider that the genetic component is only one of many factors to be taken into account when determining dietary plans. In addition to genetic factors, individual reactions to food products can also be influenced by epigenetic, microbiomic (e.g., gut microbiota), and environmental factors.

Nevertheless, the individual variability in body mass has a significant genetic basis. Studies using twin methods and familial heritage analyses estimate the heritability of body mass to be between 40% and 70%. Approximately 500 genetic loci associated with key traits of obesity, such as waist-to-hip ratio or body mass index, have been identified. The FTO genetic locus, associated with fat mass and obesity, is considered the most relevant genetic marker affecting body weight.

Key genetic variants linked to the risk of obesity and related disorders are located in the following genes: ADIPOQ, FTO, LEPR, LEP, MC4R, INSIG2, PPARG, PCSK1, ADRB3, ADRB2, PPAR γ , APOA1, GHRL, APOA5, FABP2, LIPC, MTNR1B, TCF7L2, CETP, GIPR, NPY, IRS1, and PCSK1.

For example, the FABP2 gene (fatty acid binding protein 2), expressed in the epithelium of the small intestine, participates in the absorption of fats. Variants of this gene may contribute to

higher fat absorption and lead to obesity. The PPARG gene (peroxisome proliferator-activated receptor gamma), expressed in adipocytes, plays a critical role in their differentiation. The polymorphism of the PPARG gene is associated with insulin sensitivity and body mass index. Additionally, it is known that the TMEM18 gene (transmembrane protein 18) can regulate appetite, body weight, and contribute to obesity. Furthermore, it has been established that the monogenic form of obesity may be caused by reduced expression of the MC4R gene (melanocortin-4 receptor).

INTERACTION OF GENE VARIANTS WITH DIET LEADING TO OBESITY

Interest in studying the interaction of genetic variants with diet and other environmental factors, such as physical activity, is driven by the development of methods for preventing obesity. Therefore, various populations are analyzed to evaluate the influence of gene marker polymorphisms and dietary patterns (including diets low or high in simple sugars, rich in animal fats, or deficient in protein) on the risk of developing obesity. This assessment considers not only the chemical composition of food and the enrichment of the diet with carbohydrates or fats but also the daily rhythm of food intake.

Results in this field, for instance, have shown that reducing saturated fats in the diet is more effective among individuals with overweight and obesity carrying the genotype IRS1 rs2943641 CC. It has also been established that overweight and obese individuals who carry the T allele of PPM1K rs1440581 may achieve more effective weight loss when adhering to a low-carbohydrate diet.

Moreover, individual food preferences for certain products are also linked to obesity risks. Literature suggests a list of genes influencing food preferences, which includes: FTO (rs9939609), MC4R (rs17782313), DRD2 (rs1800497). In this context, food preferences refer to the tendency to overeat, which is determined by genetic polymorphisms.

The FTO gene encodes a protein involved in energy metabolism, oxidative reactions, and fatty acid metabolism. It is possible that the protein plays a role in DNA demethylation. Population studies have established a strong association between this gene's polymorphisms with body mass index, obesity risk, and type 2 diabetes. One SNP of this gene (rs9939609) is associated with increased body mass index and obesity. High expression of FTO is associated with higher food intake and subsequent increases in body mass and fat, provided that energy expenditure and physical activity levels remain unchanged.

The MC4R gene encodes a protein that is a membrane-bound receptor and a member of the melanocortin receptor family. The encoded protein interacts with adrenocorticotrophic hormones and MSH (melanocyte-stimulating hormones), and its effects are mediated by G-proteins. This protein plays an important role in energy homeostasis and the regulation of dietary preferences. Together with melanotropin, it controls the sensation of satiety. Furthermore, the MC4R gene plays a key role in glucose homeostasis regulation. The polymorphism of this gene (rs17782313) is a cause of monogenic obesity inherited in an autosomal dominant manner.

RECOMMENDED LITERATURE FOR CHAPTER 6. GENETIC NUTRITION

1. Fenech M. et al. Nutrigenetics and Nutrigenomics: Viewpoints on the Current Status and Applications in Nutrition Research and Practice. *J Nutrigenet Nutrigenomics*. 2011; 4(2): 69–89. doi: 10.1159/000327772

2. Singar S. et al. Personalized Nutrition: Tailoring Dietary Recommendations through Genetic Insights. *Nutrients*. 2024; 16(16): 2673. doi: 10.3390/nu16162673

3. Suzuki M., Tomita M. Genetic Variations of Vitamin A-Absorption and Storage-Related Genes, and Their Potential Contribution to Vitamin A Deficiency Risks Among Different Ethnic Groups. *Front Nutr.* 2022; 9: 861619. doi: 10.3389/fnut.2022.861619
4. Kiani A.K. et al. Polymorphisms, diet, and nutrigenomics. *J Prev Med Hyg* 2022;63(suppl.3):E125-E141. <https://doi.org/10.15167/2421-4248/jpmh2022.63.2S3.2754>
5. Voruganti V.S. Precision Nutrition: Recent Advances in Obesity. *Physiology (Bethesda)*. 2023; 38(1): 42–50. doi: 10.1152/physiol.00014.2022
6. Softic S. et al. Divergent effects of glucose and fructose on hepatic lipogenesis and insulin signaling. *J Clin Invest.* 2017; 127(11): 4059–4074. doi: 10.1172/JCI94585
7. Vesnina A. et al. Genes and Eating Preferences, Their Roles in Personalized Nutrition. *Genes (Basel)*. 2020; 11(4): 357. doi: 10.3390/genes11040357

CHAPTER 7. ATHLETE'S GENETIC MAP

INTRODUCTION

The heritability of an athlete's status is estimated to be approximately 66%, implying that genetic factors account for 66% of the variance in sports performance, while the remaining 34% is attributed to environmental factors. Specifically, the heritability of certain phenotypes related to athletic achievements is highest for the composition of skeletal muscle fibers, ranging from 45–99.5%, 49–56% for muscular strength, and 44–68% for cardiorespiratory capacity. Additionally, the heritability of the risk of sports injuries is estimated to be between 60–80%. These data suggest that yet unidentified genetic factors may play a crucial role in the selection of elite athletes and the prediction of sports outcomes.

Mapping and studying the dynamic molecular networks underlying human physical exercise generally involve approaches focused on genes, proteins, peptides, post-translational modifications, lipids, and metabolites across various tissue and body fluid samples. These metabolic aspects should not be considered entirely separate as integration across these fields will be essential for delineating molecular networks and connections. Furthermore, each field can be subdivided into distinct areas involving various technological approaches. Collectively, genomic research directions have enhanced our understanding of the interaction between genetic background and the environment in human sports performance and response to physical exercise. Technological advancements in genomics now provide platforms for cheaper and faster deep sequencing of the human genome. Previous genetic research, primarily based on smaller, heterogeneous cohorts, evaluated candidate genes and variants. Genome-Wide Association Studies (GWAS) have identified gene variants (such as ACTN3 and ACE) and single nucleotide polymorphisms (SNPs) that potentially influence human physical performance, health benefits from exercise, and predictors of response to physical exercise. The molecular basis of gene interactions underlying human response to physical exercise is complex. However, it is well-established that genetic variability affects gene interactions and sports success. Efforts to map entire genomes in larger cohorts will help unravel this complexity. Recently, multi-center research initiatives, like the Athlome Project Consortium, have been launched to help implement this global strategy. These endeavors will significantly expand the catalog of gene-exercise interactions and lead to future research on how these genes are related to changes in the external environment to determine the functional foundations of training adaptation.

GENETIC FOUNDATIONS OF SPORTING ACHIEVEMENTS

Human athletic performance is a complex phenotype influenced by both environmental and genetic factors, with most traits associated with physical exercise being polygenic in nature. Physical performance traits (PPT) and training adaptations in humans result from the interaction of numerous internal and external factors, including those classified as environmental, genetic, anatomical, physiological, and psychological. Sports genomics is a scientific discipline that studies the organization and functions of the genome in elite athletes and aims to develop molecular methods that can be utilized in sports medical practice, personalized training, nutrition recommendations, and the prevention of injuries and exercise-related diseases. Recently, a new discipline known as kinesiology-genomics has emerged in this field, combining the scientific study of human movement (kinesiology) with genomics.

In terms of athletic training and exercise, metrics determined by common DNA sequence variations include muscle preparedness, exercise behavior, and the capacity for cardiorespiratory, cardiovascular, and metabolic adaptation to acute stress, as well as

susceptibility to consistent training. Athlete status is itself a highly heritable trait, with 66% to 70% of the variance in athlete status explained by additive genetic factors, while the remaining variance is attributable to environmental factors including training, nutrition, and athlete motivation. Heredity plays a crucial role in an athlete's future. Evidence suggests that 60% of the variance in baseline aerobic capacity and cardiac function (30-40% of the cardiac parameter), 70-90% of the variance in anaerobic power, and 50-70% of the variance in muscle strength can be explained by genetic factors.

It is well-known among sports geneticists that many of the key biological processes underlying athletic performance are under genetic influence, including skeletal muscle energy metabolism, mitochondrial biogenesis, muscle, bone, and cartilage formation, tissue oxygenation, erythropoiesis, and angiogenesis. Recently, there has been a significant increase in research aimed at identifying specific genes and their variants that contribute to outstanding performance, which could explain physiological capability differences among elite athletes. Based on the data obtained, the potential use of genetic information to maximize sports performance and prevent injuries has been proposed, alongside practical developments aiming for specific dietary interventions that could be individually tailored according to an athlete's microbiome.

GENES DETERMINING SUCCESS IN SPORTS

The era of sports genomics began in the early 2000s following the decoding of human DNA and the discovery of the first DNA polymorphisms associated with athletic performance (gene variations ACE, ACTN3, AMPD1, PPARG, and PPARGC1A). The status of an elite athlete is partially a heritable trait, much like numerous foundational physiological, anthropometric, and psychological traits that contribute to elite performance. In recent years, the scope of genetic analysis of specific genetic variants that contribute to elite athlete traits has significantly increased. This process encourages attempts to use genetic information as a tool for predicting the future status of elite athletes. Accumulated results evaluate the extent of genetic heritability contributing to athletic performance. Genetic traits predisposing to endurance demonstrate a range of 44%–68% heritable factors, while traits associated with strength display a range of 48%–56%. An example of illustrating the role of genes in physical exercise efficiency is the ACTN3 gene. This gene encodes the protein alpha-actinin-3, responsible for strong muscle contractions. The gene variant with a stop codon (rs1815739; R577X) results in the suppression of the synthesis of a functionally active protein (genotype XX). Individuals carrying this genotype lack the ACTN3 protein. Conversely, allele R (genotype RX or RR) is responsible for the synthesis of functional ACTN3 protein. It is known that athletes in power sports more frequently possess the functional protein compared to those in endurance sports. Thus, allele R has been associated with power sports, and genotype XX with endurance sports.

It is essential to emphasize that sports genetics not only presents opportunities for talented generations but also clarifies numerous health issues concerning the general population. Sports genetics is closely related to core medical concepts. For instance, studies on genetic predisposition to severe sarcopenia (age-related decline in skeletal muscle mass, strength, and gait speed) have shown that gene alleles associated with a higher risk of severe sarcopenia are closely correlated with parameters such as fatigue, alcohol consumption, smoking, time spent watching television, and higher consumption of salt and processed meat. In contrast, alleles associated with a lower risk of severe sarcopenia positively correlate with levels of serum testosterone, IGF1, and 25-hydroxyvitamin D; height; physical activity; and healthier dietary habits (consumption of grains, cheese, fatty fish, protein, fruits, and vegetables).

Research into the genetic basis of human physical performance (PPT) was initially conducted using twin and genealogical methods. One of the early achievements in human sports genetics was determining the heritability indices of individual structural and functional traits. Twin studies and family analyses demonstrated that traits such as anaerobic power, maximal oxygen consumption, performance capacity, and muscle fiber composition are partially inherited. These approaches are currently undergoing evolutionary shifts, transitioning from genetic epidemiology studies to biomarker analysis of DNA sequences, leading to the identification of genetic markers that determine athletic performance and physiological responses to training.

The identification of genetic markers has enabled targeting genetic variants associated with sports phenotypes to suggest their functional role in human physical performance based on each variant's influence on specific physiological processes. For example, variants of the angiotensin-converting enzyme (ACE) gene, a known regulator of blood pressure, have been shown in numerous studies to be determinants of endurance potential in athletes. Other examples of candidate genes identified and subsequently associated with phenotypes underlying sprinting, strength, or endurance performance include the skeletal muscle-specific isoform of α -actinin-3 (ACTN3), adenosine monophosphate deaminase (AMPD1), muscle creatine kinase (CKM), insulin-like growth factor (IGF1), and the family of peroxisome proliferator-activated receptor (PPAR) genes.

Among the total number of DNA polymorphisms related to sports, discovered from 1998 to 2020, the most promising genetic markers include HFE rs1799945, MCT1 rs1049434, MYBPC3 rs1052373, NFIA-AS2 rs1572312, PPARA rs4253778, PPARGC1A rs8192678, and VEGFR2 rs1870377 for endurance; ACTN3 rs1815739, AMPD1 rs17602729, CPNE5 rs3213537, CKM rs8111989, and NOS3 rs2070744 for power; LRPPRC rs10186876, MMS22L rs9320823, PHACTR1 rs6905419, and PPARG rs1801282 for strength; COL1A1 rs1800012, COL5A1 rs12722, COL12A1 rs970547, MMP1 rs1799750, MMP3 rs679620, and TIMP2 rs4789932 for soft tissue injury.

Discoveries in the genetic profiles of elite athletes have revealed DNA sequence variations associated with an individual's propensity for success in various sports, intrinsic vulnerability to sports-related injuries, and unique nutritional needs.

The search for candidate genes, which constitutes a significant portion of current sports genomics research, frequently investigates variants originally identified in animal studies using models such as gene knockout, gene knockdown, and transgenic mice. These approaches have identified numerous candidate genes subsequently analyzed for their potential contributions to human performance phenotypes. For instance, mutations in the myostatin gene (MSTN) result in extreme skeletal muscle growth, indicating that myostatin is a key regulator of muscle mass.

IDENTIFICATION OF CANDIDATE GENES IN SPORTS GENOMICS

Research in the field of sports genomics is based on the identification of candidate genes. This involves selecting genes and their variants that are associated with parameters of physical performance (PPT) in humans. To achieve this, case-control association studies are conducted to determine if a genotype or allele is more prevalent in a cohort of athletes compared to a control group of practically healthy non-athletes from the general population. The genome-wide association study (GWAS) approach is considered the most effective research design for identifying genetic markers associated with athletic performance. Indeed, the application of this approach has led to the discovery of hundreds of single nucleotide polymorphisms (SNPs) directly or indirectly associated with physical exercise and sports, such as height (12,111 SNPs), lean limb mass (1,059 SNPs), testosterone levels (855 SNPs), handgrip strength (170 SNPs),

sarcopenia (78 SNPs), and brisk walking (70 SNPs). It is important to emphasize that the discovery of 12,111 independent genome-wide significant SNPs associated with height required the study of 5.4 million individuals. This also suggests that very large samples of athletes and non-athletic control subjects are necessary to discover a significant number of genetic markers associated with athlete status. Some propose that these markers could be used as part of sports talent identification strategies.

The main challenge in sports genetics is that national samples of elite athletes are relatively small. To overcome this limitation, it has been proposed to use genome-wide significant markers of phenotypes associated with physical exercise (correlating with sports traits) discovered in large cohorts of untrained subjects. This approach can significantly reduce the risk of false-positive results when identifying markers associated with athlete status, which is acknowledged as a key limitation of sports genomics research conducted with limited sample sizes.

Another approach in sports genomics that has proven effective in eliminating the possibility of false-positive results involves conducting studies in two or more independent sports cohorts (even with small or medium sample sizes) followed by meta-analysis to quantitatively assess the overall effect of polymorphism on athlete status or individual traits related to sports and physical exercise. However, in some cases, extrapolation of results is not possible due to the uniqueness of a polymorphism to certain populations based on their geographical origin. For example, the rs671 G/A polymorphism of the aldehyde dehydrogenase 2 (ALDH2) gene has been associated with strength in athletes and non-athletes from the Japanese population. Notably, the adverse allele rs671 A (associated with reduced strength) is absent in Europeans (frequency = 0%) but is common in the Chinese, Japanese, and Vietnamese populations (15–25%).

Nonetheless, case-control studies remain the most prevalent research design in sports genomics. Typically, such studies involve determining whether a specific allele of a DNA sequence (gene or non-coding DNA region) is more common in a group of elite athletes than in the general population, thereby implying that the allele enhances athletic performance.

The identified DNA polymorphisms can be generally classified as genetic markers associated with endurance (such as marathon runners and swimmers over distances of 800–1500 meters), strength (such as weightlifters and powerlifters), power (such as sprinters, throwers, jumpers, wrestlers, and gymnasts), or with the status of an athlete in team sports. The association of a genetic marker with a specific athletic discipline can be explained by various mechanisms. For example, sprinters are not only faster but also taller than the general population and endurance athletes. Therefore, genes associated with both speed and height can also be considered sprinter status genes. Notably, the gene polymorphisms ACTN3 and CPNE5 are significant genetic markers associated with sprinter status. The ACTN3 gene encodes a protein from the alpha-actinin-binding protein family, primarily expressed in skeletal muscles, functioning as a structural component of the sarcomere Z-line and involved in cross-linking actin-containing thin filaments. Allelic polymorphism in this gene results in both coding and non-coding variants; the reference genome represents a coding allele. Interestingly, the non-functional allele of this gene is associated with elite athlete status. The CPNE5 gene encodes a protein from the calcium-dependent membrane-binding protein family, which may regulate molecular events at the cell membrane-cytoplasm interface. It is one of several genes encoding a calcium-dependent protein with two N-terminal C2-type II domains and an integrin A-domain-like sequence at the C-terminal. Several alternative splice variants have been identified for this gene, encoding different isoforms.

Numerous exercise-related traits are genetically determined and crucial for athletic performance, including intelligence, reaction time, personality traits, muscle mass, flexibility,

testosterone and lactate levels, immunological status, fat metabolism efficiency, cardiovascular and metabolic responses, and more. For instance, Mixed Martial Arts (MMA) fighters, compared to non-athletic controls, possess the 'warrior genotype' (COMT rs4680 GG; Val/Val), indicative of high catechol-O-methyltransferase enzyme activity against low dopamine levels. Such genetic traits provide an advantage related to stress resistance.

Candidate genes are typically identified based on known metabolic pathways or gene expression experiments, with subsequent case-control analysis performed on a single genetic marker. However, physical performance (PPT) is a polygenic trait, influenced by numerous genes, many of which are poorly characterized or unknown. To overcome this limitation, analyzing entire genomes rather than individual gene variants has been suggested. Consequently, genome-wide association studies (GWAS) have become a high-throughput alternative, bypassing several limitations of the candidate gene approach.

Despite the apparent advantages of GWAS, a significant limitation is that certain countries and sports disciplines have a limited number of elite athlete cohorts, lacking homogeneity. To address this issue, the ATHLOME project consortium was established in 2015, comprising 15 research centers worldwide, collectively studying genotype and phenotype data of elite athletes to advance modern understanding of the molecular basis of sports performance. ATHLOME focuses on understanding how athletes adapt to training, analyzing the root causes of exercise-related injuries and epigenetic changes impacting performance and adaptation. Ultimately, the consortium aims to develop guidelines for personalized training strategies, injury prevention, and anti-doping measures.

One of the 15 centers initially focused on analyzing genetic data from athletes and control groups in Eastern Europe (Belarus, Lithuania, Russia, and Ukraine; including 8,228 athletes and 4,121 control individuals). Athletes were grouped into international (including Olympic and World Championship participants), national, regional, or local categories. They included biathletes, long-distance runners, cyclists, triathletes, kayakers, rowers, canoeists, modern pentathletes, orienteers, skiers, speed skaters, short-track skaters, weightlifters, bodybuilders, powerlifters, strongmen, short-distance runners (≤ 400 m), short-distance swimmers (50–100 m), decathletes, heptathletes, martial artists, general athletes, bobsledders, artistic and rhythmic gymnasts, figure skaters, fencers, and team sports players.

It is noteworthy that athletes' polygenic profiles can be impacted by a phenomenon known as epistasis (or SNP–SNP interactions). These interactions occur when the combination of two or more SNPs influences the phenotype to a greater degree or differently than the effect observed with a single gene. For example, the epistatic effect of two SNPs (SNAP25 rs362584 and ACTN3 rs1815739) has been described in gymnasts.

GENETICS AND INDIVIDUAL DIFFERENCES IN RESPONSE TO TRAINING

The analysis of athletes who are either "high responders" or "low responders" to standardized training loads can provide valuable insights into the mechanisms of adaptation to training and help in developing individualized training schedules. Indeed, an accurate understanding of an athlete's genetics can offer a clearer comprehension of the body's individual response to training. Such responses include changes in body mass index, body composition, optimization of insulin activity for glucose absorption during exercise, favorable responses to endurance training for diastolic blood pressure and mean arterial pressure, and improvements in VO₂max following endurance training. The recovery rate between training sessions among endurance athletes, potentially determined by heredity, is also important. For example, it is suggested that polymorphisms in the cytochrome P450 2D6 gene (CYP2D6) and genes encoding the glutathione transferase family likely indicate the body's recovery ability by

eliminating free radicals or neutralizing pH. It has been established that the frequency of "favorable" variants in the cytochrome P450 gene (i.e., c.506-1G>A polymorphism in CYP2D6; formerly 1846G>A: rs3892097) and the glutathione S-transferase family (GSTP p.Ile105Val polymorphism, rs1695) was different among athletes with high endurance compared to a control population. Maladaptation of athletes due to the imbalance between overall stress and recovery can lead to chronic fatigue. Consequently, it is necessary to consider that the response to training can be modulated by timing and dietary composition, and therefore, nutritional factors, as environmental components, can also potentially contribute to these individual variations.

SPORTS NUTRIGENOMICS

Nutrients are one of the most critical environmental factors determining success in sports. A fundamental aspect that can influence the relationship between athletic performance and genetics is the field of nutrigenomics and nutrigenetics. These are experimental approaches utilizing genetic testing technologies to study the role of individual genetic differences in athletes' adaptive responses to food. Genetic differences affect an athlete's body's individual ability to absorb, metabolize, utilize, and excrete substances obtained from food. Nutrigenomics and nutrigenetics are experimental approaches that use genetic information and genetic testing technologies to study the role of individual genetic differences in changing athletes' responses to nutrients and other food components. Individuals react differently to the same consumed foods, nutrients, and supplements, which holds true across different ages, ethnic groups, and levels of athletic expertise. Personalized nutrition based on knowledge of a person's genotype is not a fundamentally new approach to addressing the genetic characteristics of individuals. There are several examples of rare (phenylketonuria) and common (lactose intolerance) inherited disorders that require dietary adjustments. Although genetic tests are well established in clinical settings, there is a growing number of opportunities to enhance the quality of a healthy lifestyle and improve athletic performance through genome testing focused on dietary habits.

Another illustration of this approach is the sensitivity to caffeine. Caffeine, present in products such as coffee, tea, cocoa, etc., is widely used in sports as a performance-enhancing or ergogenic aid, often in the form of caffeine-containing pills or other forms. It has been found that the single nucleotide polymorphism (SNP) in the CYP1A2 gene -163A>C (rs762551) alters the activity of the CYP1A2 enzyme and is used to identify individuals as "fast" or "slow" caffeine metabolizers. For instance, in the SNP CYP1A2 rs726551, individuals with the AA genotype (fast metabolizers) are predisposed to a positive response (increased performance) to caffeine. In contrast, individuals with the CYP1A2 AC or CC genotype experience no effect or a decrease in physical performance upon caffeine consumption.

Vitamin A is a fat-soluble vitamin that exists in various biologically active forms, such as retinal and retinoic acid, in the body. It plays a crucial role both in the visual system and in the body's immune responses. Vitamin A performs a variety of immunomodulatory functions; thus, a deficiency in Vitamin A is associated with immune dysfunctions in the intestine and several systemic immune disorders. Furthermore, Vitamin A is regarded as a powerful antioxidant that provides protection to the eyes against diseases and helps maintain vision. Elite athletes are likely to possess excellent visual capabilities, based on their various visual skills such as contrast sensitivity, dynamic acuity, and stereo acuity, all of which are required in many sports. Slow visual-motor reaction times have been linked to an increased risk of musculoskeletal injuries in sports situations when there are significant issues with the efficacy of visual stimulus and motor response execution. These visual-motor skills are key factors in achieving sports performance and, accordingly, require exceptional eye health.

After absorption, provitamin A carotenoids are easily converted into Vitamin A by the BCMO1 enzyme expressed in intestinal enterocytes. A variant of the BCMO1 gene, rs11645428, affects circulating carotenoid levels in the plasma, influencing the conversion of dietary provitamin A carotenoids into active forms of Vitamin A in the small intestine. For this reason, individuals with the GG genotype may be inefficient in this conversion process and, consequently, may be at a higher risk of Vitamin A deficiency. It is suggested that such individuals do not efficiently respond to dietary β -carotene; therefore, consuming the necessary amounts of preformed Vitamin A in the diet can help ensure sufficient levels of circulating active Vitamin A to maintain vision and immunity.

The prevalence of anemia due to iron metabolism disorders, as well as low levels of folic acid and Vitamin B12, appears to be higher in elite athletes compared to the general population. A Vitamin B12 deficiency can have significant negative consequences for sports performance. The most common symptoms of this disorder include fatigue, weakness, breathlessness, or palpitations. Iron is part of hemoglobin and myoglobin, proteins responsible for oxygen transport and storage. Vitamin B12 is associated with erythropoiesis and aerobic processes in the body. Megaloblastic anemia arises from a Vitamin B12 deficiency and is associated with elevated homocysteine levels, leading to a general sense of fatigue and weakness. Hyperhomocysteinemia may result from low intake of folic acid and Vitamin B12. These circumstances may contribute to an increased risk of impaired skeletal muscle function, including muscle weakness and regeneration.

Methylenetetrahydrofolate reductase (MTHFR) is an enzyme that limits the rate within the methylation cycle and is encoded by the MTHFR gene. The C677T polymorphism (rs1801133) in the MTHFR gene has been associated with low levels of folate in serum and erythrocytes, as well as elevated homocysteine levels in plasma, which is an independent risk factor for cardiovascular diseases. Several studies among athletic and non-athletic populations have shown that individuals with the CT or TT genotype are at an increased risk of low circulating folate levels. Compared to individuals with the CC rs1801133 genotype, those with the TT and possibly CT genotype may be at greater risk of hyperhomocysteinemia, although this may not necessarily result in decreased physical performance. On the other hand, football players and sedentary individuals with the CC genotype have a more favorable prognosis and performance metrics compared to T allele carriers. Additionally, the FUT2 gene allele (rs602662) significantly influences plasma B12 levels, with individuals possessing GG or GA genotypes at the highest risk of low plasma B12 levels.

Iron deficiency anemia is the most common type of anemia among athletes, who have a higher demand for iron due to increased needs for erythropoietin resulting from higher intensity and volume of training. Female athletes are at particular risk of iron deficiency due to menstruation and generally lower overall energy or food intake compared to males.

At the genetic level, the iron content in blood plasma is regulated by several genes, such as HFE, Tmprss6, TFR2, and TF. For instance, the HFE gene is involved in the regulation of intestinal iron absorption. It is known that polymorphisms in this gene can be associated with an increased risk of hemochromatosis (iron overload disorder). Excess iron can be toxic to tissues and cells, as "free" iron reacts with reactive oxygen species (ROS), such as superoxide and hydrogen peroxide, or lipid peroxides, forming free radicals. These free radicals can, in turn, cause damage to cells and tissues (including muscles) and ultimately lead to cell death. Iron biomarkers, such as ferritin and transferrin, are more commonly found in individuals genetically predisposed to iron overload based on the HFE gene variant. Notably, athletes with the rare HFE genotype (rs1800562) AA, associated with increased risk of hemochromatosis, may have a genetic advantage for success in sports if their iron levels are at the upper normal limit but not excessive enough to cause tissue damage. It is important to note that the risk of iron overload is

more common in elite athletes compared to the general population. Therefore, two SNPs in the HFE gene (rs1800562 and rs1799945) can be utilized in practical medicine both to identify the risk of hereditary hemochromatosis and, conversely, to assess genetic risks that may positively impact endurance in athletes who nonetheless need to avoid iron supplements, as they could pose health risks. Three other significant SNP polymorphisms: TMPRSS6 (rs4820268), TFR2 (rs7385804), TF (rs3811647), can be used to assess the genetic risk of insufficient iron levels in the body. In this context, these gene variants are involved in the regulation of hepcidin expression, a peptide hormone that controls iron absorption. Iron-deficiency anemia impairs performance by reducing hem-containing proteins' oxygen-carrying capacity.

Vitamin C functions as a water-soluble antioxidant and coenzyme in the synthesis of certain proteins. In sports medicine, however, dietary intake of ascorbic acid is primarily valued for its antioxidant properties, which help reduce free radical formation caused by intense physical exercise. A total dietary intake of vitamin C up to 250 mg per day is likely sufficient to reduce oxidative stress without negatively impacting performance. Additionally, vitamin C is necessary for collagen biosynthesis, suggesting it may play a role in muscle growth and recovery. Consequently, the presence of vitamin C in the diet during a training cycle improves collagen synthesis and may play a beneficial role in preventing injuries and accelerating the recovery of the musculoskeletal system, ligaments, and/or tendons. The relationship between dietary vitamin C and circulating ascorbic acid levels depends on the genotypes of glutathione S-transferase GSTT1 and vitamin C transporters SVCT1 and SVCT2, which are encoded by SLC23A1 and SLC23A2, respectively. Individuals with the GSTT1 Del/Del genotype have a significantly higher risk of developing vitamin C deficiency compared to those with the Ins allele.

Literature suggests the role of vitamin D in bone health, immunity, and recovery after training. Genetic determinants of circulating 25-hydroxyvitamin D (25(OH)D) may influence each of these factors, thereby affecting an athlete's performance. Only two genes, CYP2R1 (rs10741657) and GC (rs2282679), have been significantly associated with vitamin D status. Specifically, athletes with the GG or GA genotype of CYP2R1 (rs10741657) have nearly four times the tendency towards insufficient vitamin D levels. Those with the GG genotype of the GC gene (rs2282679) are significantly more prone to low vitamin D levels compared to those with the TT genotype.

This brief overview underscores that understanding individual genetic variability is essential for determining athletes' consumption of vitamin supplements and trace elements.

GENETIC VARIANTS RESPONSIBLE FOR ATHLETE ENDURANCE

Human endurance is influenced by numerous factors, including muscle fiber typology, hemoglobin mass, mitochondrial biogenesis, maximum cardiac output, and maximal oxygen consumption rate (VO₂max), among others. The ability to perform endurance exercises is contingent upon aerobic metabolism and the efficiency of oxygen utilization for energy production. These factors are significantly supported by enhanced mitochondrial function, gene expression, and enzyme activity among elite athletes. Endurance in elite athletes is bolstered by sustained muscle contraction over prolonged periods, utilizing high VO₂max and other endurance reserves linked with the mitochondrial respiratory chain.

The ability to perform endurance exercises depends on several central factors related to muscle and cardiovascular function, such as the proportion of slow-twitch fibers in skeletal muscles and maximum cardiac output, which underlies VO₂max. There is evidence suggesting that genetic factors significantly impact the phenotypic expression of these traits, potentially explaining up to 70% of endurance-related variability. Genetic markers associated with athletic

endurance are typically identified by comparing allele frequencies between endurance athletes (e.g., biathletes, road cyclists) and a control group of non-professional athletes. Furthermore, recovery levels between training sessions among endurance athletes are associated with genes influencing metabolic processes in the liver, such as cytochrome P450 2D6 (CYP2D6) and genes encoding the glutathione S-transferase family.

Regular endurance training promotes skeletal muscle adaptation, ensuring metabolic adaptation to endurance exercises. Examples include a slower rate of muscle glycogen and blood glucose utilization, a higher contribution of fat oxidation to muscle energy exchange during physical activity, and reduced lactate production at high exercise intensities, leading to increased myocardial volume and circulating blood volume. This optimizes ventricular filling, resulting in greater stroke volume and cardiac output, which significantly enhances the ability to perform prolonged, intense endurance exercises. However, not all physiological parameters mentioned can be attributed solely to genetics.

To confirm the hereditary nature of case-control study findings aimed at identifying the genetic basis of endurance, additional functional laboratory studies are conducted to establish links between genotypes and physiological indicators. These supplementary genomic studies include, but are not limited to, measurements of VO₂max, forced expiratory volume in one second (FEV₁), the proportion of slow-twitch muscle fibers, recovery speed, long-distance running performance, running economy, lactate threshold, erythropoietin and hemoglobin levels, erythrocyte count, capillary density, mitochondrial density, fatty acid metabolism, and fatigue resistance.

It is widely acknowledged that most endurance-related phenotypes are under strong genetic influence. Some 40-50% of the variance in the proportion of slow-twitch fibers in human muscles seems to be genetically determined. Furthermore, VO₂max and aerobic power exhibit high heritability, with up to 50% of the variance in VO₂max due to innate factors. The genetic component of muscle fiber type variability is approximately 40-50%, indicating that muscle fiber composition is equally influenced by genotype and environment.

Highlighting the connection between sports medicine and practical health care, it's noteworthy that VO₂max serves as an effective prognostic factor for morbidity and mortality due to various causes, particularly cardiovascular diseases. Numerous genetic association studies for other cardiovascular risk factors, such as poor diet, high blood pressure and cholesterol, stress, smoking, and obesity, have revealed new therapeutic targets. However, while VO₂max has a substantial genetic component, association studies with human health are still lacking.

The endurance athlete status is possibly the most researched trait, with genetic predisposition currently linked to approximately 100 genetic variants. Based on these principles, research has shown that at least 41 out of 114 recorded genetic markers may be associated with an athlete's endurance status according to specified criteria. Notably, elite athletes with high endurance levels possessing the TT genotype in the ACTN3 gene or DD genotype in the ACE gene appear more frequently in cohorts of elite athletes than in control groups of untrained individuals. Prominent genetic markers for determining endurance status in athletes include alleles AMPD1 rs17602729 C, CDKN1A rs236448 A, HFE rs1799945 G, MYBPC3 rs1052373 G, NFIA-AS2 rs1572312 C, PPARA rs4253778 G, and PPARGC1A rs8192678 G.

GENE VARIANTS RESPONSIBLE FOR ATHLETE STATUS IN STRENGTH SPORTS

The strength metrics of athletes correlate with various physiological parameters such as testosterone levels, the proportion of fast-twitch muscle fibers, muscle mass, body and heel

bone height, muscle bundle length, reaction time, among other traits. It is estimated that the heritability of power-related phenotypes ranges from 49% to 86% across different phenotypes. Genetic factors significantly influence the capability of skeletal muscles to generate force rapidly, which is crucial for success in both strength and sprint sports. Achievements in strength sports are based on a multitude of factors that contribute substantially to the strength phenotype. The identified genetic factors are selected based on a range of physiological parameters that determine the success of a power athlete: skeletal muscle hypertrophy (muscle fiber size), the prevalence of fast-twitch muscle fibers, a greater angle of muscle bundle pennation, enhanced neurological adaptation, high glycolytic capacity, and elevated circulating testosterone levels. There is evidence that strength athletes exhibit quite specific transcriptomic, biochemical, anthropometric, physiological, and biomechanical parameters compared to endurance-oriented athletes as well as a control group. The noted intergroup differences can be explained both by environmental factors (training, nutrition, etc.) and the effects of genetic factors. Research findings indicate a strong heritability of traits related to power and strength, with genetic factors accounting for up to 85% of the variation in maximal isometric, isotonic, and isokinetic strength. The heritability of strength-related phenotypes ranges approximately from 49% to 86%. Furthermore, several strength characteristics are positively associated with indicators such as circulating testosterone levels, the percentage and cross-sectional area of fast-twitch muscle fibers, muscle mass and strength, body and heel bone height, muscle bundle length, and reaction time, among others.

Typically, marker genes associated with athlete status are determined by comparing the frequency of allele occurrences between athletes (e.g., 100-meter runners, shot putters, arm wrestlers, etc.) and non-trained individuals. To verify case-control study results, a genotype-phenotype analysis is conducted based on measurements such as sprint time, jump performance, muscle fiber size, muscle fiber typology, maximal strength, rate of force development, and levels of anabolic hormones in the blood, such as testosterone. At least 45 out of 95 genetic markers associated with strength athlete status met true criteria. Among the most relevant potential genetic markers associated with athlete status are alleles ACTN3 rs1815739 C, AMPD1 rs17602729 C, CDKN1A rs236448 C, CPNE5 rs3213537 G, GALNTL6 rs558129 T, IGF2 rs680 G, IGSF3 rs699785 A, NOS3 rs2070744 T, and TRHR rs7832552 T.

It is noteworthy that muscle strength phenotypes in the general population may be less pronounced than in strength athletes. In this context, strength athletes may be considered an ideal population for identifying genomic variants that determine skeletal muscle volume. On the other hand, insights gained from this athlete sample expand our understanding of pathological phenomena such as sarcopenia, aiding in developing strategies to mitigate the adverse effects of age-related muscle mass decline. Generally, power athlete marker genes can be identified by comparing allele occurrence frequencies between strength athletes and the general population control group. To verify these results, genotyping and phenotyping studies are conducted, measuring grip strength and isokinetic strength, performance in powerlifting/weightlifting, and assessing acute and chronic responses to strength training.

Presently, 23 DNA polymorphisms have been verified that reflect hand grip strength and are associated with strength athlete status. Among these SNPs, the polymorphisms LRPPRC rs10186876, MMS22L rs9320823, and PHACTR1 rs6905419 also correlate with high performance in weightlifting competitions. It is possible that 42 genetic markers can objectively reflect strength athlete status. The most promising genetic markers for determining strength athlete status are alleles ACTN3 rs1815739 C, LRPPRC rs10186876 A, MMS22L rs9320823 T, PHACTR1 rs6905419 C, and PPARG rs1801282 G, including AR ≥ 21 CAG repeats.

GENETICS OF SPORTS INJURIES

In modern sports, elite athletes face significant physical and emotional demands due to an intense schedule of competitions and increasing training volumes. It is evident that as the intensity of the competitive calendar and training load rises, so does the risk of injury. Notably, the likelihood of injury is higher during competitions (72.2%) than during training sessions (21.8%). Typically, about 80% of sports injuries affect the musculoskeletal system. The type of injury may be influenced by the specific sport in which the athlete participates. For example, the annual risk of deltoid muscle injury ranges from 12% to 19% for baseball players and from 23% to 38% for swimmers. Among middle-distance runners, the incidence of lower limb muscle injuries related to training processes is estimated to be between 19% and 58%. In tennis, joint issues are commonly reported (29.5%), including the ankle (20%) and wrist (15.8%).

The idea that genetic factors partly explain susceptibility to sports injuries was proposed a decade ago (Collins M., 2010). It was suggested that genetic polymorphisms might serve as biomarkers for sports injuries. Recent studies have shown that certain genetic polymorphisms are associated with sports injuries, such as IGF2 (rs3213221), CCL2 (rs2857656), ELN (rs2289360), COL1A1 (rs1107946), COL5A1 (rs12722), COL22A1 (rs11784270 and rs6577958), VDR (rs7975232), MCT1 (rs1049434), ACTN3 (rs1815739), HGF (rs5745697 and rs1011694), SOX15 (rs4227), ESR1 (rs2234693), CYP19A1 (rs936306), ACE (rs4341), MMP3 (rs679620), TNC (rs2104772), IL6 (rs1800795), NOS3 (rs1799983), and HIF1A (rs11549465), which are considered markers for potential muscle injuries; ELN (rs2289360), COL1A1 (rs1800012), COL5A1 (rs12722, rs71746744, rs16399, rs4919510), and MIR608 (rs4919510) for potential ligament or tendon injuries; and RANK (rs3018362), RANKL (rs1021188), OPG (rs4355801), P2X7R (rs1718119 and rs3751143), VDR (rs10735810, rs731236), SOST (rs1877632), CYP19A1 (rs936306), and COL1A1 (rs1107946) for potential stress fractures.

Consequently, these genetic polymorphisms may contribute to susceptibility to sports injuries. However, the quality of evidence may be insufficient due to the small sample size of elite athletes. Meanwhile, the application of molecular biology methods has enabled the identification of genetic markers associated with the risk of specific sports injuries. The findings of research in this area offer hope that identified genetic markers could be used by athletes to tailor their training and dietary regimes.

By the early 2020s, approximately 124 single nucleotide polymorphisms (SNPs) had been identified using next-generation sequencing technology, allowing for the prediction of potential sports injury risk related to an athlete's anatomical or physiological characteristics. The most commonly predicted risks are associated with a possible anterior cruciate ligament (ACL) tear, Achilles tendon injury, low bone mineral density and stress fracture, osteoarthritis, vitamin/mineral deficiencies, and sickle cell trait. It is particularly noteworthy that 91% of the identified SNPs had not previously attracted the attention of sports genetics specialists.

The first category of injuries, the anterior cruciate ligament (ACL) tear, is one of the most severe injuries incurred during exercises involving repeated landings and abrupt turns. Candidate gene studies conducted on athletes have identified single nucleotide polymorphisms in the COL1A1, COL5A1, COL12A1, and MMP12 genes, which are associated with ACL tears. On the other hand, it is known that dynamic warm-up programs aimed at correct landing and sudden movement techniques, as well as core and hamstring strength regulation, are effective in reducing the risk of ACL injuries if the athlete is informed of the existing risk.

The second category of injuries, Achilles tendon injury, includes degenerative and painful conditions that affect athletes across a wide range of sports, impacting up to 18.5% of runners. Candidate gene studies have identified variants in the MMP3, COL5A1, COL1A1, CASP-8, and GDF5 genes, which are associated with Achilles tendon pathologies. The rs1800012 gene variant is located in the promoter region of the COL1A1 gene, which encodes type I collagen,

the main protein component of tendons and ligaments. The rare T allele increases gene expression, whereas the G allele reduces expression and is associated with an increased risk of Achilles tendon rupture. Effective preventive measures for this type of injury include stretching to enhance flexibility and strengthening the calf and soleus muscles.

The third category of injuries, stress fractures, is a common overuse injury affecting up to 20% of female athletes and novice athletes. Unlike fractures caused by a single strong impact, stress fractures result from accumulated microtrauma due to repetitive loading, such as running or jumping. Bone mineral density (BMD) is a leading factor in determining a bone's ability to withstand load, and low BMD is a significant risk factor for stress fractures in athletes. Heredity accounts for 50–85% of BMD variance. Active young individuals, with a family history of osteoporosis and low BMD, are almost twice as likely to develop stress fractures.

The fourth risk category involves osteoarthritis, which is a leading cause of pain and disability in older adults and is the most common form of arthritis worldwide. High-intensity sports training and competitions can lead to cartilage damage, eventually causing early osteoarthritis. Currently, five significant loci across the genome have been identified as associated with a high risk of osteoarthritis, and three loci are just below this threshold. The strongest association was found on chromosome 3 for rs6976, which is in linkage disequilibrium with rs11177. This SNP encodes a missense polymorphism in the GNL3 gene, which encodes nucleostemin. Nucleostemin levels were elevated in chondrocytes in patients with osteoarthritis. Other significant loci were found on chromosome 9 near ASTN2, chromosome 6 between FILIP1 and SENP6, chromosome 12 near KLHDC5 and PTHLH, and another region of chromosome 12 near CHST11. One of the signals near genome-wide value was located in the FTO gene, which is involved in body mass regulation—a strong risk factor for osteoarthritis. These polymorphisms have allelic odds ratios between 1.1 and 1.2; however, all seven loci have an additive allelic effect.

The fifth category includes SNPs associated with deficiencies in athletes' vitamins and minerals. Intense physical activity leads to biochemical adaptations in muscles, subjecting metabolic pathways to stress and increasing the demand for micronutrients. Regular training increases the loss of micronutrients, thus requiring more intensive replenishment to maintain and restore the body's chemical composition. Optimal levels of calcium and vitamin D are crucial for bone health, while iron, zinc, and the B vitamin complex are essential for maintaining hematological stability.

The concluding sixth category, sickle cell anemia (hemoglobin S, rs334), involves athletes who are heterozygous for the mutant allele of the β -globin chain. In this case, under several adverse conditions such as high altitude, extreme heat, dehydration, and exhaustion, erythrocytes may adopt a pathological sickle phenotype. In turn, with increased physical exertion, this can lead to dangerous consequences such as macrohematuria, splenic infarction, and rhabdomyolysis, which can be fatal. Current guidelines advise athletes heterozygous for the sickle cell gene to follow slow and gradual training regimens, rest and fully recover during training cycles, immediately cease physical activity if symptoms occur, maintain optimal hydration, use supplemental oxygen as needed, and promptly seek medical attention if feeling unwell.

Additionally, some authors distinguish muscle tissue damage as a separate category. Muscle injuries pose a serious problem for elite athletes, accounting for at least 30% of cases and resulting in temporary loss of physical activity. The majority of muscle injuries occur during sports competitions, comprising 10–55% of all injuries. Muscle damage generally occurs via two main mechanisms: sudden direct compressive force leading to contusion, or excessive stretching force causing muscle fiber damage and possible tearing. The muscles most frequently injured in sports are the quadriceps and calf muscles. These muscles span two joints and are

most susceptible to acceleration and deceleration forces. Muscle damage is accompanied by temporary ultrastructural disruption of myofibrils, muscle strength loss, delayed muscle soreness, restricted movement of the affected limb, and swelling. It is important to acknowledge that there is high individual variability in athletes' responses to muscle damage induced by physical exercise. Increasing reports suggest that certain gene variants may be linked to exercise-induced muscle damage. Polymorphisms in the following genes have been detected in individuals suffering from exercise-induced muscle damage: ACTN3 (R577X, rs1815739), TNF (-308 G>A, rs1800629), IL6 (-174 G>C, rs180795), and IGF2 (G>A, rs680). Additionally, epigenetic mechanisms may also be involved in these processes. Specifically, data suggest that acute physical exertion can influence DNA methylation processes, leading to hypomethylation of the promoter regions of the TFAM, PGC-1 α , PDK4, and MEF2A genes.

MITOCHONDRIAL DNA AND ATHLETIC PERFORMANCE

Mitochondrial DNA (mtDNA), a double-stranded circular DNA, consists of 16,569 base pairs and contains thirteen protein-coding genes, two ribosomal RNAs (rRNAs), and twenty-two transfer RNAs (tRNAs). One of the primary functions of mitochondria is energy production. Adenosine triphosphate (ATP), essential for skeletal muscle contraction during exercise, is most efficiently produced through the oxidative phosphorylation (OXPHOS) system in mitochondria. Due to their role in cellular energy production, mitochondria have been thoroughly studied by sports medicine specialists. It is known that endurance training increases mitochondrial content in skeletal muscles, leading to enhanced exercise endurance by improving ATP production.

Mitochondrial DNA (mtDNA) was discovered in 1963. It is located in the mitochondrial matrix as a double-stranded circular molecule. Typically, each mitochondrion contains 2 to 10 copies of mtDNA. In mtDNA, 37 genes have been identified, encoding 13 protein subunits of the respiratory chain, 22 tRNAs, and two rRNAs (12S rRNA and 16S rRNA). An important feature of the mitochondrial genome is its susceptibility to various damages, resulting in a mutation risk for mtDNA that is almost ten times higher than that of nuclear genes. The high mutation risk of mtDNA is due to several reasons. Firstly, mtDNA does not contain protective histone proteins; secondly, the efficiency of repair processes is somewhat reduced compared to nuclear DNA; thirdly, mtDNA is more susceptible to mutagenic effects of free radicals generated in mitochondria during oxidative phosphorylation; fourthly, mtDNA lacks introns and contains very little redundancy. Consequently, deletions or point mutations in mtDNA can easily and rapidly provoke biochemical effects.

As noted earlier, the thirteen proteins encoded by mtDNA are directly components of OXPHOS, essential for its proper functioning. Therefore, it is plausible that mitochondrial single nucleotide polymorphisms (mtSNPs) could impact athletic performance. Indeed, recent data suggest that mtSNPs, m.8701A>G, m.10873T>C, m.12705C>T, and m.15043G>A, are significantly associated with the body's response to training, as evaluated by changes in the lactate threshold. Additionally, the frequency of mtSNPs m.185G>A, m.228G>A, m.295C>T, m.462C>T, m.489T>C, m.4215T>C, m.10397A>G, m.11250A>G, m.13707G>A, m.14797T>C, m.15451C>A, m.16068C>T, and m.16125T>C is higher in subjects with the lowest VO₂ max levels compared to those with the highest response.

Notably, mtSNPs m.152T>C, m.4343A>G, m.11215C>T, m.15518C>T, and m.15874A>G have been associated with elite athlete status in endurance sports. In contrast, mtSNPs m.151C>T, m.204T>C, m.4833A>G, m.5108T>C, m.5601C>T, m.7600G>A, m.9377A>G, m.13563A>G, m.14200T>C, m.14569G>A, m.15314G>A, and m.16278C>T have been linked to elite sprint athlete status.

EPIGENETICS IN SPORTS MEDICINE

Epigenetics is a rapidly evolving field in the molecular physiology of physical exercise, studying changes in gene expression without alterations in the DNA sequence. Epigenetic modifications (such as DNA methylation, histone acetylation, and non-coding RNAs) regulate gene expression by activating or silencing genes and often occur in response to environmental or external stimuli, such as specific training regimens or diets, leading to particular changes in transcriptional responses. Epigenetics serves as a bridge between an individual's genetic status and their external environment. In sports medicine, epigenetic biomarkers are primarily understood as predictors of the body's adaptive response to exercise and training. For example, it has been shown that training can lead to DNA hypomethylation in the promoter regions of key metabolic genes regulated by physical exercise (e.g., PPARGC1A, PPARG), accompanied by the overexpression of these genes. Thus, the variability in phenotypic adaptation induced by physical activity can be partially explained by epigenetic factors. However, specific epigenetic changes induced by exercise have not been sufficiently studied. It has also been shown that epigenetic markers of training adaptation persist in skeletal muscles even after periods of detraining, suggesting that epigenetic mechanisms underlie the intrinsic ability of the human body to "remember" training adaptations. Recently, a new class of epigenetic markers known as chromosome conformation signatures has shown promising results in medical research and may also be used to study whether changes in genome conformation (the organization of the genome in three-dimensional nuclear space) participate in regulating the response to exercise.

A critical task for scientists remains determining the type, intensity, and frequency of exercise necessary to maintain epigenetic memory (such as DNA methylation) of exercise adaptation, which may persist over a prolonged period. If such advances are achieved, this information could potentially be used to optimize sports training programs or improve recovery processes after injuries. Physical activity significantly impacts DNA methylation patterns at the level of brain structures such as the hypothalamus, hippocampus, and cortex, highlighting the role of exercise and physical activity as epigenetic regulators of brain plasticity and cognitive processes that counteract aging and neurodegeneration.

Another important component of sports epigenetics is the interplay between nutrigenomics, nutrigenetics, sports performance, and genetics. Nutrigenomics and nutrigenetics are experimental approaches using genetic testing technologies to explore the role of individual genetic differences in an athlete's adaptation to dietary factors. Considering these factors is undoubtedly essential when evaluating the influence of genetics on sports performance, as an individual's dietary and nutritional strategies, being environmental factors, can affect athletic outcomes. Personalized nutrition for elite athletes based on genetic profiles aims to optimize health, metabolic processes, and sports performance. Sports nutritionists successfully apply extensive verification of dietary recommendations across all general nutrition guidelines for the population to account for different athletic groups.

The variability in phenotypic adaptation of an athlete induced by physical activity can be partially explained by epigenetic factors, although specific exercise-induced changes remain under-researched. A critical task for scientists is determining the type, intensity, and frequency of exercise necessary to maintain the epigenetic memory (e.g., at the DNA methylation level) of exercise adaptations that may persist over a long time. If such advances are achieved, this information could be used to optimize sports training programs or enhance recovery processes post-injury.

RECOMMENDED LITERATURE FOR CHAPTER 7. ATHLETE'S GENETIC MAP

1. Genetic Passport: The Foundation of Individual and Predictive Medicine, edited by V. S. Baranov. St. Petersburg: N-L Publishing, 2009. — 528 pages.
2. Ahmetov, I.I., Fedotovskaya, O.N. Current Progress in Sports Genomics. *Advances in Clinical Chemistry*. 2015;70:247-314. doi: 10.1016/bs.acc.2015.03.003
3. Ahmetov, I. et al. Chapter Five - Advances in Sports Genomics. *Advances in Clinical Chemistry*. 2022;107:215-263
4. Ginevičienė, V., Utkus, A., Pranckevičienė, E., et al. Perspectives in Sports Genomics. *Biomedicines*. 2022;10(2):298. doi: 10.3390/biomedicines10020298
5. Tanisawa, K., Wang, G., Seto, J., et al. Sport and Exercise Genomics: The FIMS 2019 Consensus Statement Update. *British Journal of Sports Medicine*. 2020;54(16):969–975. doi: 10.1136/bjsports-2019-101532
6. Hoffman, N.J. Omics and Exercise: Global Approaches for Mapping Exercise Biological Networks. *Cold Spring Harbor Perspectives in Medicine*. 2017;7(10):a029884. doi: 10.1101/cshperspect.a029884
7. Konopka, M.J. et al. Genetics and Athletic Performance: A Systematic SWOT Analysis of Non-Systematic Reviews. *Frontiers in Genetics*. 2023;14:1232987. doi: 10.3389/fgene.2023.1232987
8. Pitsiladis, Y.P., Tanaka, M., Eynon, N., et al. Athlome Project Consortium: A Concerted Effort to Discover Genomic and Other 'Omic' Markers of Athletic Performance. *Physiological Genomics*. 2016;48(3):183–190. doi: 10.1152/physiolgenomics.00105.2015
9. Semenova, E.A., Hall, E.C.R., Ahmetov, I.I. Genes and Athletic Performance: The 2023 Update. *Genes (Basel)*. 2023;14(6):1235. doi: 10.3390/genes14061235
10. Varillas-Delgado, D., Del Coso, J., Gutiérrez-Hellín, J., et al. Genetics and Sports Performance: The Present and Future in the Identification of Talent for Sports Based on DNA Testing. *European Journal of Applied Physiology*. 2022;122(8):1811–1830. doi: 10.1007/s00421-022-04945-z
11. Kraemer, W.J., Ratamess, N.A., Hymer, W.C., et al. Growth Hormone(s), Testosterone, Insulin-Like Growth Factors, and Cortisol: Roles and Integration for Cellular Development and Growth With Exercise. *Frontiers in Endocrinology (Lausanne)*. 2020;11:33. doi: 10.3389/fendo.2020.00033
12. Guest, N.S. et al. Sport Nutrigenomics: Personalized Nutrition for Athletic Performance. *Frontiers in Nutrition*. 2019;6:8. doi: 10.3389/fnut.2019.00008
13. Martín-Rodríguez, A. et al. Advances in Understanding the Interplay between Dietary Practices, Body Composition, and Sports Performance in Athletes. *Nutrients*. 2024;16(4):571. doi: 10.3390/nu16040571
14. Clemente-Suárez, V.J. et al. Personalizing Nutrition Strategies: Bridging Research and Public Health. *Journal of Personalized Medicine*. 2024;14(3):305. doi: 10.3390/jpm14030305
15. Pickering, C., Kiely, J., Grgic, J., et al. Can Genetic Testing Identify Talent for Sport? *Genes (Basel)*. 2019;10(12):972. doi: 10.3390/genes10120972
16. Rajesh Kumar et al. The Role of Mitochondrial Genes in Neurodegenerative Disorders. *Current Neuropharmacology*. 2022;20(5):824–835. doi: 10.2174/1570159X19666210908163839
17. Sellam, M. et al. Molecular Big Data in Sports Sciences: State-of-Art and Future Prospects of OMICS-Based Sports Sciences. *Frontiers in Molecular Biosciences*. 2021;8:815410. doi: 10.3389/fmolb.2021.815410
18. Giraldo-Vallejo, J.E. et al. Nutritional Strategies in the Rehabilitation of Musculoskeletal Injuries in Athletes: A Systematic Integrative Review. *Nutrients*. 2023;15(4):819. doi: 10.3390/nu15040819

19. Goodlin, G.T., Roos, A.K., Roos, T.R., et al. Applying Personal Genetic Data to Injury Risk Assessment in Athletes. *PLoS One*. 2015;10(4):e0122676. doi: 10.1371/journal.pone.0122676
20. arcOGEN Consortium and arcOGEN Collaborators. Identification of New Susceptibility Loci for Osteoarthritis (arcOGEN): A Genome-Wide Association Study. *Lancet*. 2012;380(9844):815–823. doi: 10.1016/S0140-6736(12)60681-3
21. Maffulli, N. et al. The Genetics of Sports Injuries and Athletic Performance. *Muscles, Ligaments and Tendons Journal*. 2013;3(3):173–189.
22. Tarnowski, M. et al. Epigenetic Alterations in Sports-Related Injuries. *Genes (Basel)*. 2022;13(8):1471. doi: 10.3390/genes13081471
23. Kumagai, H. et al. Novel Insights into Mitochondrial DNA: Mitochondrial Microproteins and mtDNA Variants Modulate Athletic Performance and Age-Related Diseases. *Genes (Basel)*. 2023;14(2):286. doi: 10.3390/genes14020286

ПОЛИГЕННЫЕ ЗАБОЛЕВАНИЯ

Авторы:

Доломатов Сергей Игоревич, Крымский Федеральный университет им. В.И. Вернадского, к.б.н., доцент

Жуков Валерий, Университет Николая Коперника, к.м.н., доцент

Жукова Анна Александровна, Крымский Федеральный университет им. В.И. Вернадского, к.б.н., доцент

Смирнова Светлана Николаевна, Крымский Федеральный университет им. В.И. Вернадского, к.б.н., доцент

РАЗДЕЛ 1. РОЛЬ НАРУШЕНИЙ НАСЛЕДСТВЕННОГО АППАРАТА КЛЕТКИ В ПАТОГЕНЕЗЕ НАСЛЕДСТВЕННЫХ ЗАБОЛЕВАНИЙ

ВВЕДЕНИЕ

В практической медицине всё чаще востребованы методы **генетического тестирования** в рамках предиктивной медицины. При этом следует отличать понятия «**генетический тест — ГТ**» (**genetic test**) и «**генетический (молекулярный) анализ — ГА**» (**genetic assay**). Первое понятие относится к уже выявленным и подтвержденным диагностическим генам-маркерам. В отличие от ГТ, **результаты ГА имеют прогностический характер и требуют дальнейших подтверждений, прежде чем обретут статус генетического теста**. Полученные на достаточно большой выборке пациентов, подкрепленные убедительной по численности группой контроля молекулярные анализы полиморфизма отдельных генов даже в условиях одной популяции или населения одного региона приближаются к критериям, предъявляемым к генетическому тесту.

Генетическое тестирование в досимптоматический период дает возможность выявить существующие пока только в геноме (не выраженные ещё фенотипически) наследственные тенденции к развитию будущих болезней. Следовательно, генетический тест позволяет исходя из современного врачебного опыта, наметить пути мониторинга этих болезней и разработать методы их ранней профилактики. На основе генетического тестирования может быть получена информация о той или иной степени риска развития указанных заболеваний. Благодаря этому врач, принимая во внимание результаты молекулярно-генетического анализа, может разработать тактику патогенетически обоснованной упреждающей терапии, то есть внести необходимую коррекцию врожденного метаболического дефекта.

Очевидно, что в результате этих исследований происходит накопление данных, как о геноме отдельных индивидов, так и о целых семьях, то есть постепенно формируются индивидуальные и семейные базы ДНК-данных. Эта **база ДНК-данных и является «генетическим паспортом»**.

Таким образом, **генетический паспорт представляет собой индивидуальную базу ДНК-данных, отражающую уникальные генетические особенности каждого человека, его предрасположенность к тем или иным наследственным заболеваниям**.

В настоящее время уже на основе имеющихся данных вполне оправданно тестирование многих генов-маркеров, ассоциацию которых с тяжелыми заболеваниями можно считать доказанной.

Информация, содержащаяся в генетическом паспорте, помогает избежать потенциально возможных нарушений здоровья человека, связанных с игнорированием индивидуальных особенностей генома, специфических характеристик наследственности. С другой стороны, такие данные позволяют человеку более полно реализовать свои генетические способности и представляют несомненную ценность для потомков.

По мнению В. С. Баранова и О.С. Глотова, индивидуальная генетическая карта (генетический паспорт) предполагает следующие основные разделы:

- (1) собственно паспортные данные (Ф И О, год рождения, национальность);
- (2) кариотип пациента;
- (3) собственный идентификационный генетический номер, полученный методом геномной дактилоскопии и представляющий собой число из 12 цифр;
- (4) результаты тестирования на бессимптомное носительство (гетерозиготность) мутаций наиболее частых моногенных болезней;

- (5) досимптоматическую диагностику болезней с поздней манифестацией;
- (6) результаты тестирования генов-маркеров частых мультифакториальных болезней;
- (7) медико-генетическое заключение по результатам генетического консультирования до и после генетического тестирования;
- (8) рекомендации по результатам генетического тестирования для консультируемого и для лечащего врача.

Генетическая карта (генетический паспорт) в полном варианте должна включать в себя результаты исследования не только генов предрасположенности, но и сведения о бессимптомном носительстве мутаций генов наиболее частых наследственных болезней (гетерозиготность по генам гемофилии, муковисцидоза, фенилкетонурии и др.). Практически востребованным является тестирование наследственной предрасположенности по какой-то одной или нескольким нозологиям, в отношении которых, согласно результатам предварительного медико-генетического консультирования, у пациента имеются наследственные риски.

Подтвердила свою эффективность **генетическая карта (генетический паспорт) репродуктивного здоровья**, составленная НИИ акушерства и гинекологии им. Д. О. Отта СЗО РАМН (Санкт-Петербург). Помимо анализа кариотипа и тестирования на носительство мутаций тяжелых наследственных заболеваний (гетерозиготность) у супругов, планирующих ребенка, важное прогностическое значение имеет исследование женщины по генным панелям заболеваний, осложняющих беременность, развитие плода, роды и послеродовый период (гестозы, привычное невынашивание, варикозная болезнь, фетоплацентарная недостаточность).

Принимая во внимание значение данных генетического паспорта для здоровья ребенка, гармоничного формирования его личности, рационального питания, эффективного образования, спортивных занятий, оптимальной профориентации и возможности упредить развитие ряда болезней с поздней манифестацией, составление такого генетического теста в раннем возрасте представляется вполне оправданным.

Кроме того, в семьях высокого риска по диабету 1-го и 2-го типа, бронхиальной астме, синдрому внезапной смерти, нарушениям сердечной проводимости и ритма, метаболическому синдрому и ожирению, а также при ряде других нозологий вполне оправданным представляется упредительное генетическое тестирование уже в раннем возрасте. Разумеется, что проводится такое исследование может только с согласия родителей, по направлению врача педиатра и после консультации семьи врачом-генетиком. Накапливается информация и о генах-маркерах, тестирование аллельных вариантов которых позволяет оценить пригодность подростка к тому или иному виду спорта.

Технология секвенирования экзома - WES позволяет изучать структуру, генетические полиморфизмы и функции различных вариантов генов в геноме в пределах определенных популяций. Это способствует пониманию наследственных механизмов конкретного моногенного заболевания или мультифакторного (полигенного) заболевания и делает возможным их точную диагностику, повышает эффективность профилактики и лечения. Данное направление стимулирует быстрое развитие молекулярной предиктивной медицины и структур генетического паспорта как ступеньки к **клиническому генетическому паспорту**, полученному в результате данных секвенирования экзома (технология WES).

Например, в репродуктивной медицине, в частности, при планировании беременности необходим комплексный подход, основанный на целом ряде **молекулярных, генетических, цитогенетических и эмбриологических методов**.

Клиническо-генетический паспорт, генетическое картирование и секвенирование нового поколения присутствуют во всех областях медицинской науки, позволяя клинической медицине решать важные репродуктивные проблемы. В частности, в области таких разделов медицины, как неинвазивное пренатальное тестирование (НИПТ) или скрининг моногенных и олигогенных заболеваний путем обнаружения патогенных вариантов у пробандов и семей с высоким риском; планирование беременности с использованием предимплантационной генетической диагностики (ПГД) и лечения; и подтверждение диагноза, перспективно распространяющееся на оценку риска мультифакторных болезней и инфекционных заболеваний в дополнение к идентификации фенотипических признаков у людей. Следует отметить критически важные обновления в генетической терминологии относительно недавно появившихся методов. **Вместо общепринятых терминов «мутация» и «полиморфизм»** в 2015 и 2017 годах Американский колледж медицинской генетики и геномики (ACMG) и Российское общество медицинской генетики, соответственно, **рекомендовали использовать термин «вариант нуклеотидной последовательности» со следующими модификаторами: (1) Р - патогенный, (2) LP – возможно патогенный, (3) неопределенная значимость, (4) вероятно доброкачественный или (5) доброкачественный.**

Технология WES позволяет быстро и эффективно идентифицировать сразу несколько наследственных заболеваний у любого человека. В частности, это касается **случаев совместно наследуемых X-сцепленных и аутосомно-доминантных форм патологий**: ихтиоза, болезни Вильсона и гемохроматоза. **Зная молекулярные дефекты, которые приводят к развитию заболевания, пациенты могут получить пользу от наиболее адекватного последующего наблюдения.**

Большинство заболеваний не являются моногенными, поэтому прежде, чем оценивать степень риска заболевания необходимо правильно понять природу заболевания (т. е. моногенное, олиогенное или многофакторное состояние), что не всегда легко. Открывается возможность выявлять лиц с высоким генетическим риском наиболее частых патологий (диабет, сердечно-сосудистые заболевания, онкопатология и т. д.) в любом возрасте. Ситуация с многофакторными заболеваниями несколько сложнее, поскольку изменения в геноме влияют на этиологию заболевания, при этом большой набор генов предрасполагает к заболеванию (феномен аддитивности). Кроме того, помимо наследственного фактора, **предрасположенность к заболеванию формируется большим количеством факторов окружающей среды, в то время как наследование не подлежит объяснению только законами Менделя.** Действительно, идентичные диагнозы могут быть вызваны разными факторами риска и этиологиями у разных людей. В этих случаях технология WES предлагает рациональные алгоритмы, с помощью которых можно идентифицировать генетические маркеры сложных заболеваний, даже в ограниченных выборках.

Подводя итог можно сказать, что концепция предиктивной медицины — использование клиническо-генетического паспорта для решения проблем предзачаточного скрининга, рождения здорового потомства, диагностики и профилактики мультифакторных заболеваний, а также инфекционных заболеваний — должна опираться на технологии NGS, как на фундаментальный инструмент с использованием специализированных баз данных, алгоритмов, биоинформатики и генетических пользовательских концепций экспрессии и пенетрантности.

МОЛЕКУЛЯРНАЯ ДИАГНОСТИКА МОНОГЕННЫХ И ПОЛИГЕННЫХ ЗАБОЛЕВАНИЙ

Применение NGS в медицинских исследованиях многогранно. По своему назначению методы секвенирования можно разделить на следующие группы: (1) анализ всего генома (секвенирование всего генома, WGS); (2) анализ генов, кодирующих белок в геноме (секвенирование всего экзона, WES); (3) анализ последовательностей генов, вызывающих определенные заболевания (от экзотов, охватывающих около 4000–5000 клинически значимых генов, до наборов для небольших целевых областей из одного–трех генов); (4) секвенирование транскриптома (РНК-секвенирование); и (5) анализ биологического разнообразия бактериального микробиома.

Учитывая, что целью является ограничение исследования последовательностями нуклеотидов, кодирующих белок, для выявления редких патологических SNP, вставок, делеций, которые могут лежать в основе заболевания, или выявления новых генетических маркеров олиогенных заболеваний и мультифакторных заболеваний, WES является более экономически эффективным методом, чем WGS. WES занимает особое место среди тестов NGS. Учитывая его относительно низкую стоимость, WES в настоящее время более привлекателен в клинических условиях. Кроме того, WES позволяет существенно минимизировать размер анализируемой базы данных до 6 Гб, в отличие от WGS (90 Гб).

Благодаря указанным качествам WES становится первоочередным методом обследования для сложных исследовательских и практических целей. WES/WGS позволяет идентифицировать новые неизвестные или редкие варианты, которые связаны с заболеваниями или очень редкими патологиями. В эту группу также следует отнести гены предрасположенности, которые присутствуют в некоторых популяциях, но до сих пор не учитывались в дифференциальной диагностике и не считались причиной заболевания.

Секвенирование всего экзона (исследование, которое позволяет расшифровать последовательность генов, кодирующих белки, и выявить патогенные (опасные) генетические варианты, которые были унаследованы от родителей или возникли случайно, WES) используется для проведения массового скрининга структурных и регуляторных генов с целью определения частот аллелей полиморфизмов, связанных с заболеваниями, в различных популяциях и, таким образом, обнаружения патогенных генетических изменений (мутаций или полиморфизмов), способствующих возникновению нарушений последовательностей белков. Благодаря своим обширным возможностям, секвенирование экзона сегодня позволяет, как диагностировать моногенные заболевания (МЗ), так и проводить обследование, казалось бы, здоровых популяций для выявления широкого спектра потенциальных рисков до проявления заболевания. Генетический паспорт человека определяется как новая клиническая концепция на основе WES для идентификации новых генов-кандидатов, вариантов генов и

молекулярных механизмов в диагностике, прогнозировании и лечении моногенных, олигогенных и многофакторных заболеваний.

Также необходимо отметить, что цитогенетический метод исследований также не утратил своей актуальности при формировании генетического паспорта. Некоторые виды делеций (микроделеций) и транслокаций (сбалансированные транслокации) встречаются довольно часто и могут рассматриваться в качестве серьезной медицинской проблемы, например в области репродуктивной медицины или в онкологии (потеря гетерозиготности).

Следовательно, данные цитогенетического анализа являются важной составляющей генетического паспорта.

МОНОГЕННЫЕ ЗАБОЛЕВАНИЯ ЧЕЛОВЕКА

База данных OMIM (по состоянию на 1 июня 2023 г.) включает записи для 7377 наследственных заболеваний и синдромов, а также их молекулярных ассоциаций. К ним относятся 6305 фенотипов, связанных с одним геном, т. е. показывающих моногенную природу генетического признака или синдрома. Это во многом было достигнуто благодаря активному внедрению WES.

Технология WES обеспечивает надежную методику диагностики моногенных болезней у людей. В частности, применение WES в качестве обследования первой линии для генетической диагностики и при идиопатических расстройствах у пробандов с «размытым» фенотипом. В этом случае, у пациентов со смазанной клинической картиной дифференциальная диагностика с WES необходима для выявления первопричины заболевания.

Например, в КНР было проведено исследование с применением WES для изучения 169 детей с врожденными аномалиями развития (медианой возраста 10,5 месяцев) и моногенных болезней. Моногенные расстройства были диагностированы у 43 (25%) пациентов. Патологии с самой высокой частотой включали метаболические (33%) и нервно-мышечные (19%) заболевания, а также множественные деформации (14%). Эффективность диагностики у детей с метаболическими расстройствами, нарушением роста или аномалиями зрительного анализатора существенно повысилась после того, как стали доступны подробные клинические данные. Данные WES позволили внести коррективы в 30 (70%) случаях, включая начало мониторинга заболевания в 41,9% (18 случаев), реабилитацию и паллиативную помощь в 27,9% (12 случаев), изменение текущего лечения в 25,6% (11 случаев), другие комплексные процедуры оценки в 7% (3 случая). Еще одним примером верификации с помощью WES сложной формы моногенного заболевания является случай анализа области горячей точки в промоторе гена эндорибонуклеазы процессинга РНК (*RMRP*) у пробанда с крайне редкой аутосомно-рецессивной скелетной хондродисплазией (анауксетическая дисплазия, AD). Гетерозиготные варианты rs387906533 (n.91_92delinsGC) нуклеотидной последовательности (chr9:35657924-35657925delCTinsGC) были обнаружены в экзоне 1 гена *RMRP* и неизвестная замена n.-6_-5insTCTCAGCTTCAC (chr9:g.35658020_35658021insTCTCAGCTTCAC) в области промотора гена. Вариант представляет собой вставку из 12 нуклеотидов между TATA-боксом и точкой старта транскрипции. При этом, было обнаружено, что мутация n.-6_-5insTCTCAGCTTCAC имела отцовское происхождение, а мутация n.91_92delinsGC – материнское. Ранее не сообщалось о каких-либо доказательствах относительно вставки в область промотора гена *RMRP* как причины болезни Альцгеймера без внескелетных проявлений (типичных для носителей подобных мутаций).

Таким образом, технология WES позволяет быстро и эффективно идентифицировать сразу несколько наследственных заболеваний у любого человека. В частности, это касается **случаев совместно наследуемых X-сцепленных и аутосомно-доминантных форм патологий**: ихтиоза, болезни Вильсона и гемохроматоза. **Зная молекулярные дефекты, которые приводят к развитию заболевания, пациенты могут получить пользу от наиболее адекватного последующего наблюдения.**

ГЕНЫ-МОДИФИКАТОРЫ МОНОГЕННЫХ ЗАБОЛЕВАНИЙ

Тем не менее, даже знание патологических генов и характера мутаций, вызывающих заболевания, не всегда позволяет с точностью составить прогноз в отношении ключевых аспектов фенотипического проявления заболевания, таких как возраст начала болезни, скорость прогрессирования патологии, тяжесть и проявление других сопутствующих заболеваний. В литературе имеются многочисленные сообщения, в которых пациенты с одинаковой первичной мутацией демонстрируют отчетливые различия в проявлении патологического фенотипа, скорее всего, из-за генетических детерминант, которые действуют как модификаторы. Накопленные клинические и генетические данные в настоящее время свидетельствуют о том, что точка зрения о том, что экспрессивность фенотипа является исключительно

свойством доминантных или рецессивных мутаций в одном локусе даже в отношении классических моногенных расстройств, может быть чрезмерным упрощением. В этом смысле, явления, описываемые такими терминами, как эпистаз или генетическое взаимодействие/модификация, являются эффективными синонимами одного и того же процесса, а именно, влияния одного гена на фенотипический результат второго гена. Различие, которое может помочь лучше определить эти генетические явления, состоит в установлении того, является ли первичный локус одновременно необходимым и достаточным для возникновения заболевания. Если да, то присутствие аллеля(ей) во втором локусе/гене оказывает чисто модифицирующую роль на тяжесть фенотипа, определяемую скоростью прогрессирования заболевания; тяжесть фенотипа в отношении его плейотропии; или на проявление эндофенотипов, которые требуют генетических взаимодействий. Напротив, если нет, то требование аллеля во втором гене для проявления патологии определяет случай дигенного или олигогенного наследования (простейшей формы полигенного заболевания).

Генетический модификатор может оказывать свое влияние на фенотип, установленный мутацией(ями) в первичном локусе, несколькими способами. В зависимости от природы модификатора это взаимодействие может привести к снижению пенетрантности определенных эндофенотипов и/или расстройств; к переменной экспрессивности с модифицирующим локусом, что приводит либо к усугубленным фенотипам (т. е. более раннему возрасту начала и более быстрому прогрессированию заболевания), либо, напротив, к более мягкой клинической картине (т. е. более позднему возрасту начала, ослабленному прогрессированию заболевания и отсутствию/защите от определенных эндофенотипов).

Модификаторы, которые усиливают проявление определенных эндофенотипов могут приводить к тому, что данного моногенное заболевание может быть описано, как новая форма заболевания.

ОЛИГОГЕННЫЕ И МНОГОФАКТОРНЫЕ ЗАБОЛЕВАНИЯ

Наследственные заболевания, вызванные патогенными вариантами в нескольких генах, встречаются гораздо чаще, т. е. так называемые олигогенные наследственные заболевания. Олигогенные заболевания являются промежуточным состоянием между моногенными заболеваниями, связанными с одним конкретным дефектным геном, и полигенными заболеваниями, вызванными несколькими генами и факторами внешней среды.

Олигогенные модели, даже те, которые включают только два гена (т. е. дигенные или билokusные), по-прежнему трудно обнаружить и проверить из-за того, что вовлеченные варианты генов могут быть довольно распространены в общей популяции, могут иметь меньший индивидуальный эффект на функцию гена и могут быть расположены в генах, которые не всегда известны, как биологически взаимодействующие. Данные, связанные с дигенными заболеваниями, простейшей формой олигогенности, были впервые собраны в базе данных дигенных заболеваний (DIDA). Дигенное наследование (DI) касается патологий с простейшей формой мультигенной этиологии, вовлекающей более 1 гена (и, возможно, факторы окружающей среды). Истинное DI имеет место, когда биаллельные мутации в 2 отдельных генах, в цис- или транс-положении, необходимы и достаточны для того, чтобы вызвать патологию с определенным диагнозом. База данных DIDA позволила разработать новое поколение прогностических инструментов, нацеленных на комбинации вариантов, связанных с дигенным заболеванием. База данных также предоставляет исходную оценку достоверности для каждой олигогенной комбинации на основе доказательств, подтверждающих их связь с заболеванием.

Для истинного понимания механизмов, приводящих к заболеванию, также важна правильная функциональная оценка того, как варианты и гены действуют синергически, вызывая патологический фенотип. «Истинный олигогенный» сценарий предполагаемой олигогенной патологии соответствует случаю, когда **все варианты должны присутствовать одновременно, чтобы проявить какие-либо симптомы заболевания**, в то время как в сценарии «моногенный плюс модификатор» вариант в первичном гене может по-прежнему вызывать более легкие симптомы или другой субфенотип. Кроме того, особенно важно не только присутствие индивидуальных вариантов, расположенных в разных генах, но и случаи, когда гены содержат гомозиготные или гетерозиготные составные варианты, поскольку надлежащие доказательства должны показывать, что эти биаллельные события в одном гене недостаточны для того, чтобы вызвать ожидаемый патологический фенотип. Другими словами, требуемые доказательства олигогенной природы заболевания должны продемонстрировать, что биаллельные события (гетерозиготность) в одном гене недостаточны для того, чтобы вызвать фенотип заболевания, при этом лица, несущие мутации только в одном гене, являются здоровыми или имеют менее серьезные симптомы. **Это означает, что комбинация дигенных вариантов связана с фенотипом заболевания в рамках истинной дигенной модели, где одновременное присутствие обоих вариантов у индивидуума необходимо для развития фенотипа заболевания.**

Еще одним направлением идентификации олигогенных заболеваний является предоставление функциональных доказательств взаимодействий генов, указывающих на олигогенную комбинацию в заболевании, и разработка биологических механизмов, в которые вовлечены гены и их варианты. Для получения функциональных олигогенных доказательств основное внимание следует уделять причинно-следственной связи между множественными вариантами генов и изучаемым фенотипом заболевания, а также демонстрации синергического или аддитивного эффекта (т. е. доказательства того, что эффект комбинации отличается от индивидуальных эффектов вариантов) с использованием экспериментов *in vivo* и/или *in vitro*. Т.е., рассмотрение патогенности олигогенной комбинации на уровне гена очень важно для понимания того, как они действуют синергически, вызывая заболевание.

ОЛИГОГЕННАЯ ЭТИОЛОГИЯ НА ПРИМЕРЕ КАРДИОМИОПАТИЙ

Понимание роли наследственного фактора в патогенезе заболевания имеет решающее значение. Является ли расстройство моно- или олигогенным, или многофакторным, ответ часто далек от очевидности. Технология WES помогает ответить на этот вопрос. Применение WES для выявления шести новых вариантов Р (патогенных) или LP (вероятно патогенных) генов у 40 пациентов с гипертрофической (HCM, n=14) и дилатационной кардиомиопатией (DCM, n=26). Гипертрофическая кардиомиопатия, вызванная вариантами генов, кодирующих саркомерные белки — тяжелую цепь β -миозина (MYH7) и миозин-связывающий белок С (MYBPC3), — **составляет до 50% всех клинических случаев**: легкая цепь миозина 2 (MYL2), легкая цепь миозина 3 (MYL3) и сердечный тропонин Т (TNNT2) в 5–10%, сердечный тропонин I (TNNI3) в 5%, сердечный тропонин С (TNNC1) в < 1%, сердечный α -актин (ACTC1) в < 1%, α -тропомиозин (TPM1) в 1,5% и богатый цистеином и глицином белок 3 (CSRP3).

В результате было установлено, что мутации со сдвигом рамки считывания (11:47372858, с.А224insG+) в гене MYBPC3 и миссенс-мутации (rs193922390, с.5135 G>A, p.R1712Q) в гене MYH7 являются патогенными вариантами для HCM. Миссенс-варианты, такие как rs138049878 (с.2608 C>T, p.R870C), rs27503260 (с.2302 G>C, p.G768R) и rs397516088 (с.1063 G>A, p.A355T) в гене MYH7 и rs199476306 (с.188 C>T, p.A63V) в гене TPM1, были оценены как вероятно патогенные для HCM. **Диагностический выход WES для HCM составил 43%** (шесть вариантов у четырнадцати пациентов). Это достаточно высокая эффективность диагностики, подтверждающая участие небольшого количества генов в развитии патологии.

МНОГОФАКТОРНЫЕ (ПОЛИГЕННЫЕ) ЗАБОЛЕВАНИЯ

К многофакторным заболеваниям (полигенным заболеваниям) относятся практически все наиболее распространенные хронические заболевания человека, включая атеросклероз, диабет, ожирение, бронхиальная астма, остеопороз, эндометриоз, злокачественные опухоли, а также нейropsychиатрические и сердечно-сосудистые заболевания, возникающие в результате взаимодействия многих генов с неблагоприятными факторами окружающей среды. В настоящее время Международная классификация болезней (МКБ) включает более 55 000 нозологических единиц. Подавляющее большинство из них относится к мультифакторным заболеваниям. По состоянию на 11 марта 2023 года зарегистрировано более 12 000 заболеваний человека, 220 322 SNP и 493 105 геномных ассоциаций. На сегодняшний день существует три основных подхода к идентификации генов-кандидатов: метод функционального картирования (анализ генов-кандидатов), генетическое сцепление в семьях высокого риска и GWAS, включая секвенирование генома. GWAS активно используются для анализа и тестирования образцов из разных национальных биобанков. Согласно современному уровню знаний, влияние генетических факторов на экспрессию и пенетрантность фенотипических признаков обусловлено наличием генетических полиморфизмов точечных мутаций с сильными эффектами, с одной стороны, или частых SNP со слабыми эффектами, с другой стороны. Для понимания корреляции генотип-фенотип предпочтительным подходом является исследование генов-кандидатов, связанных с МФЗ.

МЕДИЦИНСКИЕ АСПЕКТЫ ГЕННЫХ МУТАЦИЙ И ХРОМОСОМНЫХ АБЕРРАЦИЙ НА ПРИМЕРЕ ДИАГНОСТИКИ НАСЛЕДСТВЕННЫХ ПРИЧИН БЕСПЛОДИЯ

Бесплодие определяется, как неспособность зачать ребенка у пары, которая пытается завести ребенка в течение двух лет без зачатия. Примерно 15% пар бесплодны, и среди этих пар мужской фактор бесплодия составляет около 50% причин. Зачастую, но не всегда, мужское бесплодие является многофакторным синдромом, охватывающим широкий спектр расстройств. У более, чем 50% бесплодных мужчин причина их бесплодия неизвестна (идиопатическая) и может быть врожденной или

приобретенной. Бесплодие у мужчин может быть первоначально диагностировано с помощью анализа спермы. Семенные спермограммы бесплодных мужчин могут выявить множество аномальных состояний, которые включают азооспермию, олигозооспермию, тератозооспермию, астенозооспермию и т.д.. По текущим оценкам, около 30% мужчин, обращающихся в медицинские учреждения в связи с проблемой бесплодия, имеют олигозооспермию или азооспермию неизвестной этиологии. Причины бесплодия известны менее чем в 50% случаев, из которых основными факторами являются генетические или наследственные заболевания и специфические аномалии в хромосоме Y. Около 10-20% мужчин, у которых отсутствует сперма в эякуляте, имеют делецию хромосомы Y. Эта удаленная область включает локус фактора азооспермии (AZF), расположенный в области Yq11, который разделен на четыре повторяющихся удаленных неперекрывающихся подобластей, обозначенные как AZFa, AZFb, AZFc и AZFd. Каждая из этих подобластей может быть связана с определенной гистологией семенников, и в этих областях было обнаружено несколько генов-кандидатов. Сообщается, что семейство генов Deleted in Azoospermia (DAZ) является наиболее часто удаляемым геном-кандидатом AZF и расположенного в области AZFc. Кроме того, ген DAZ имеет аутосомный гомолог DAZL (DAZ-Like) на коротком плече хромосомы 3 (3p24), и возможно, что мутировавший аутосомный DAZL может быть ответственен за сперматогенный дефект. Случайные атаки на обнаженную митохондриальную ДНК (мтДНК) сперматозоидов активными формами кислорода или свободными радикалами неизбежно вызовут окислительное повреждение или мутацию митохондриального генома с патологическими последствиями и приведут к бесплодию у мужчин. Ключевым ядерным ферментом, участвующим в удлинении и восстановлении цепей мтДНК, является ДНК-полимераза гамма, картированная на длинном плече хромосомы 15 (15q25), и включающая в себя область повтора CAG. Ее мутация влияет на выработку аденозинтрифосфата. Внедрение молекулярных методов дало большое понимание генетики бесплодия. Тем не менее, наше понимание генетических причин мужского бесплодия остается ограниченным.

НАРУШЕНИЕ СТРУКТУРЫ И ЧИСЛА ХРОМОСОМ – ОДНА ИЗ ПРИЧИН БЕСПЛОДИЯ

Широко известно, что наличие цитогенетических дефектов является важным фактором нарушения сперматогенеза. **Впервые связь наследственного фактора с мужским бесплодием была установлена в 1959** году, когда было показано, что у мужчин с **синдромом Клайнфельтера (СК)** имеется дополнительная X-хромосома, что приводит к кариотипу 47,XXY. СК является распространенным заболеванием (~1–2 мужчины на 1000) и одной из ведущих наследственных причин мужского бесплодия (3–4% случаев). Пациенты с СК обычно характеризуются маленькими семенниками, необструктивной азооспермией (НОА) и высоким уровнем гормонов ФСГ и ЛГ, что называется гипергонадотропным гипогонадизмом.

Хромосомные аномалии подтверждаются у 5% пациентов с тяжелой олигозооспермией и у 10%-15% пациентов с азооспермией. Обычно для анализа хромосом проводится культивирование лимфоцитов (72 часа) – цитогенетический метод. При рутинном анализе анализируется 20 клеток. В случаях хромосомного мозаицизма или хромосомных аномалий анализируется 30 клеток.

Помимо СК, еще до 1,7% пациентов с мужским бесплодием имеют аутосомные аномалии, такие как транслокации или инверсии. Лица с аномалиями кариотипа, затрагивающими, как число хромосом, так и их структуру, имеют очевидную предрасположенность к бесплодию из-за неудач в достижении успешной беременности, часто из-за повторных спонтанных аборт. Основные типы аномалий кариотипа, которые часто могут приводить к снижению фертильности, а именно хромосомные транслокации или инверсии, анеуплоидия и микроделеции Y-хромосомы. У пациентов со структурными аномалиями (в основном, по отношению к транслокациям) нарушается образование нормальных бивалентов во время мейоза, что приводит к ожидаемым нарушениям мейоза и остановке созревания сперматогенеза. У большинства пациентов со структурными изменениями хромосом наблюдается олигозооспермия. Поэтому частота Робертсоновских транслокаций, реципрокных транслокаций и инверсий выше у мужчин с олигозооспермией по сравнению с мужчинами с азооспермией и мужчинами в общей популяции.

Характер лечения бесплодия зависит от знания точной хромосомной перестройки и степени нарушения сперматогенеза. Для предотвращения передачи генетического дефекта потомству могут использоваться варианты преимплантационного и пренатального генетического тестирования. Например, **носители Робертсоновских транслокаций имеют высокий риск зачать потомство с анеуплоидией**, бессимптомные носители микроделеций или дупликаций, связанных с клиническими синдромами, могут стать отцами ребенка, у которого разовьется тяжелое нарушение развития.

Хромосомные транслокации являются наиболее распространенными структурными нарушениями у мужчин с частотой 1,23 на 1000 и их распространенность в 10 раз выше в бесплодной

популяции. Хромосомные транслокации делятся на сбалансированные и несбалансированные транслокации. Сбалансированная реципрокная транслокация представляет собой обмен генетическим материалом между двумя или более хромосомами. Существуют аутомсомные и половые хромосомные транслокации в сбалансированных реципрокных транслокациях. В зависимости от точек разрыва, приблизительно 60% носителей аутомсомных транслокаций имеют по крайней мере один аномальный параметр в анализе спермы. Хотя частота встречаемости транслокаций половых хромосом невелика, некоторые отчеты показали связь между транслокациями Y-хромосомы и азооспермией.

Робертсоновская транслокация является наиболее распространенной формой несбалансированной хромосомной транслокации у людей, а также частой причиной мужского бесплодия. Робертсоновские транслокации встречаются у 0,9–3,4% бесплодных мужчин с тяжелой сперматогенной дисфункцией. Они могут возникать в пяти парах акроцентрических хромосом (это хромосомы 13, 14, 15, 21, 22) и приводить к их разрыву в области центромера, в результате чего два длинных плеча сливаются вместе, образуя одну большую хромосому. Таким образом, у людей с Робертсоновской транслокацией 45 хромосом. Остатки коротких плеч двух слитых хромосом обычно утрачиваются. Несмотря на эту генетическую аномалию, носители Робертсоновской транслокации фенотипически нормальны, поскольку короткие плечи двух акроцентрических хромосом не содержат важных генов. Однако носители подвержены повышенному риску анеуплоидии сперматозоидов, что может привести к выкидышу или рождению детей с грубой наследственной аномалией. Теоретически, одна шестая часть сперматозоидов носителей имеет нормальный кариотип, другая одна шестая несет робертсоновскую транслокацию, а оставшиеся две трети находятся в несбалансированном состоянии, либо в нуллисомии, либо в дисомии хромосом, вовлеченных в транслокацию. Известно, что синдром Прадера-Вилли и синдром Ангельмана связаны с материнской и отцовской однородительской дисомией хромосомы 15.

Одна из широко распространенных причин мужского бесплодия – микроделеции длинного плеча Y-хромосомы (Yq). Длинное плечо Y-хромосомы (Yq) содержит множество ампликонных и палиндромных последовательностей, что делает его предрасположенным к саморекомбинации во время сперматогенеза и, следовательно, восприимчивым к внутривхромосомным делециям. Такие делеции приводят к изменению числа копий генов Y-хромосомы, что приводит к мужскому бесплодию. Три распространенные делеции Yq, которые повторяются у бесплодных мужчин, называются микроделециями AZF (фактор азооспермии), а именно AZFa, AZFb и AZFc. По оценкам, основанным на данных исследования почти 40 000 Y-хромосом, глобальная распространенность микроделеций Yq составляет 7,5% у бесплодных мужчин; однако европейские бесплодные мужчины менее восприимчивы к микроделециям Yq, самая высокая распространенность наблюдается у американцев и восточноазиатских бесплодных мужчин. Кроме того, частичные делеции локуса AZFc были связаны с бесплодием, но эффект, по-видимому, зависит от этнической принадлежности.

МОНОГЕННЫЕ ПРИЧИНЫ БЕСПЛОДИЯ

Проблема моногенных причин бесплодия носит общий медицинский характер, она не является частной, поскольку более 50% бесплодных мужчин страдают от одного или нескольких хронических заболеваний или серьезных общих проблем со здоровьем не связанных с репродуктивной сферой. Недавний метаанализ четко показал, что бесплодные мужчины имеют более высокий риск ранней смерти. Более того, риск ранней смертности увеличивается с тяжестью нарушений сперматогенеза и снижением качества спермы. Точные механизмы этой корреляции, наблюдаемые в эпидемиологических анализах, пока не известны.

Большая часть зарегистрированных моногенных форм сперматогенных дефектов представляет собой гомозиготные **аутомсомно-рецессивные патогенные варианты**, картированные в кровнородственных семьях. Бесплодие является широко распространенной клинической проблемой, затрагивающей 8-12% пар по всему миру. Из них около **75% диагностируется как идиопатическое бесплодие, поскольку причинный фактор не обнаружен**. Фактически, 40–60% случаев нарушения сперматогенеза остаются необъяснимыми, а среди случаев умеренной олигозооспермии (количество сперматозоидов в эякуляте 10–39 миллионов) эта доля близка к 80%. В последние годы было приложено много усилий для исследования новых генов-кандидатов, ответственных за мужское бесплодие, и обусловленных мутациями одного гена.

Количественные или структурные хромосомные аномалии давно связывают с мужским бесплодием. **Моногенные мутации только недавно стали рассматриваться в патогенезе этого состояния**. Мутации определенных генов, участвующих в мейозе, митозе или спермиогенезе, приводят к сперматогенной недостаточности, что приводит к следующим аномалиям: недостаточное (олигозооспермия) или отсутствие (азооспермия) производство спермы, ограниченная прогрессивная

и/или общая подвижность сперматозоидов (астенозооспермия), измененная морфология сперматозоидов (тератозооспермия) или их комбинации.

Необходимо отметить, что моногенные причины бесплодия могут быть обусловлены генами, расположенными, как в Y хромосоме, так и в аутосомах. Например, как уже отмечалось, микроделеции Yq являются следствием внутривхромосомных неаллельных гомологичных рекомбинационных событий (из-за высокоповторяющейся структуры хромосомы Y), и в целом они присутствуют у 5%–10% пациентов с азооспермией. В этом случае, патологическое влияние на мужскую фертильность вызвано потерей нескольких ключевых регуляторных генов сперматогенеза.

В настоящее время ведутся активные разработки генетических панелей, нацеленных на обнаружение моногенных причин мужского бесплодия обусловленных недостаточностью сперматогенеза и нечувствительностью к андрогенам и основанных на тестировании генов, расположенных, как в аутосомах, так и в Y хромосоме.

В целях диагностики моногенных причин бесплодия у мужчин обычно используется полигенная панель NGS для обнаружения нуклеотидных вариаций в кодирующих экзонах и фланкирующих интронах вышеупомянутых генов. Патогенные варианты могут быть миссенс, нонсенс, сплайсинг или небольшие индели. Крупные делеции/дупликации были зарегистрированы в генах, определяющих развитие мужского бесплодия *USP9Y*, *DPY19L2*, *SPATA16*, *SUN5* и *CFAP43*.

МОНОГЕННЫЕ ПРИЧИНЫ НЕДОСТАТОЧНОСТИ СПЕРМАТОГЕНЕЗА

Дефекты количества сперматозоидов включают азооспермию и олигозооспермию. Азооспермия — это отсутствие сперматозоидов в семенной жидкости. Она поражает 1% мужского населения и составляет 20% всех случаев мужского бесплодия. Точечные мутации, вызывающие азооспермию, были недавно обнаружены в следующих генах: *NR5A1*, *SYCP3*, *ZMYND15*, *TAF4B*, *TEX11*, *NANOS1*, *PLK4*, *MEIOB*, *SYCE1*, *USP9Y*, *SOHLH1*, *TEX15*, *HSF2* и *KLHL10*. В частности, мутации сдвига рамки считывания в *ZMYND15* вызывают фенотип SPGF14. Белок, кодируемый этим геном, участвует во временно нормальной гаплоидной экспрессии генов во время сперматогенеза.

Гомозиготная мутация в *SYCE1* связана с фенотипом SPGF15. Этот ген кодирует компонент синаптомембранного комплекса, структуры, которая физически связывает гомологичные хромосомы во время мейоза I.

Мутации в *TEX11* связаны с остановкой мейоза и азооспермией с частотой 1-15% у мужчин с азооспермией. Белок, кодируемый *TEX11*, регулирует сцепление гомологичных хромосом при репарации двухцепочечной ДНК посредством образования синаптомембранного комплекса и хиазмы во время процесса кроссинговера. Похожую роль выполняет *SYCP3*, который также был обнаружен мутировавшим у бесплодных мужчин.

Ген *SOHLH1* мутирует в некоторых случаях азооспермии и кодирует тестикулярный фактор транскрипции, необходимый для сперматогенеза.

Мутация в гене *NR5A1*, кодирующем стероидогенный фактор 1, была зарегистрирована у пациента с остановкой мейоза и нормальными уровнями фолликулостимулирующего и лютеинизирующего гормонов, а также тестостерона.

В завершение, у пациента с заблокированным на стадии сперматоцитов сперматогенезом наблюдалась доминантная негативная мутация в гене *HSF2*, кодирующем фактор транскрипции белка теплового шока 2.

МОНОГЕННЫЕ ПРИЧИНЫ ДЕФЕКТОВ МОРФОЛОГИИ СПЕРМАТОЗОИДОВ

Тератозооспермия — это гетерогенная группа расстройств. Морфологическая оценка сперматозоидов учитывает основные функциональные области (головку, тело и хвост), которые могут иметь аномалии формы и размера. Фенотип может включать один тип порока развития у одного пациента или разные типы пороков развития у одного и того же пациента. Недавние исследования семейных случаев патологии и в небольших фенотипически однородных когортах позволили выявить аутосомно-рецессивные случаи тератозооспермии.

Гены, которые в настоящее время связаны с морфологическими дефектами сперматозоидов: *AURKC*, *ZPBP*, *DPY19L2*, *SPATA16*, *PICK1*, *BRDT* и *SUN5*.

МОНОГЕННЫЕ ПРИЧИНЫ ДЕФЕКТОВ ПОДВИЖНОСТИ СПЕРМАТОЗОИДОВ

Астенозооспермия — это состояние, приводящее к снижению подвижности сперматозоидов из-за дефектов жгутика. Могут быть затронуты аксонема, внешние плотные волокна, митохондрии или

фиброзная оболочка жгутика. Ультраструктурные дефекты в структуре аксономы 9+2 могут затрагивать внешние или внутренние динеиновые плечи, центральные микротрубочки и радиальные спицы.

Гены, связанные с дефектами подвижности сперматозоидов: *SLC26A8*, *CATSPER1*, *SEPT12*, *CFAP43*, *CFAP44*, *DNAH1* и *PLCZ1*.

ПОЛИГЕННЫЕ ПРИЧИНЫ БЕСПЛОДИЯ У МУЖЧИН

Несмотря на большой прогресс, достигнутый благодаря цитогенетическим исследованиям и технологиям NGS в выяснении генетических причин мужского бесплодия, остается большая доля неучтенной наследственности, которую все еще предстоит изучить. Действительно, у большой доли пациентов с нарушениями сперматогенеза этиология остается неизвестной. В то же время, все больше доказательств указывают, в частности, на идиопатическую необструктивную азоосперию как на сложное и высокополигенное состояние.

С этой целью в течение последних трех десятилетий были предприняты большие усилия для выявления полиморфизмов генов-кандидатов, связанных со сложной формой идиопатическими видами бесплодия. По этому направлению SNP были наиболее проанализированными вариациями, по сравнению с исследованиями VNTR и CNV.

В качестве кандидатов генов предрасположенности были выбраны гены которые участвуют в выработке гормонов, регуляции клеточного цикла и сперматогенеза. Однако, существенного прогресса в этих исследованиях достичь не удалось. Основными причинами такого состояния дел являются (1) чрезвычайно малые размеры проанализированных когорт случай/контроль, что, скорее всего, привело к многочисленным ошибкам при проверке статистических гипотез, и (2) неудовлетворительная клиническая характеристика включенных пациентов.

В ряде предпринятых популяционных исследований было предложено для продуктивного анализа около 2000 генов, которые напрямую участвуют в сперматогенном процессе, причем более 600 из них демонстрируют специфическую экспрессию в мужских половых клетках. Такое количество генов-кандидатов чрезвычайно затрудняет визуализацию общей молекулярной картины азооспермии, заставляя некоторых авторов скептически относиться к возможности перевода молекулярно-генетических идей в клиническую практику здравоохранения. Ряд исследовательских групп начали регулярно анализировать последовательность генов *NR5A1*, *DMRT1* и *TEX11* у пациентов с идиопатическими вариантами необструктивной азооспермии, выявляя причинно-патогенные мутации у 5% из них, повышая эффективность диагностики до 25%. При том, что в повседневной клинической практике около 75% случаев с тяжелыми нарушениями сперматогенеза имеют идиопатическое происхождение. Предлагается список из 60 генов-кандидатов для диагностирования бесплодия мужчин, вызванного нарушениями сперматогенеза. Предполагается, что исследование с помощью Next Generation Sequencing в больших когортах пациентов с явно идиопатическим бесплодием предоставит новые интересные данные об их расовой и этнической распространенности у бесплодных пациентов, что, вероятно, повысит диагностическую эффективность.

РЕКОМЕНДОВАННАЯ ЛИТЕРАТУРА К РАЗДЕЛУ 1. РОЛЬ НАРУШЕНИЙ НАСЛЕДСТВЕННОГО АППАРАТА КЛЕТКИ В ПАТОГЕНЕЗЕ НАСЛЕДСТВЕННЫХ ЗАБОЛЕВАНИЙ

Генетический паспорт — основа индивидуальной и предиктивной медицины / Под ред. В. С. Баранова. — СПб.: Изд-во Н-Л, 2009. — 528 с.

Glotov O.S., Chernov A.N., Glotov A.S. Human Exome Sequencing and Prospects for Predictive Medicine: Analysis of International Data and Own Experience. J Pers Med. 2023; 13(8): 1236.
doi: [10.3390/jpm13081236](https://doi.org/10.3390/jpm13081236)

Guerri G., Maniscalchi T., Barati S. и соавт. Non-syndromic monogenic male infertility. Acta Biomed. 2019; 90(Suppl 10): 62–67.
doi: [10.23750/abm.v90i10-S.8762](https://doi.org/10.23750/abm.v90i10-S.8762)

Laan M., Kasak L., Punab M. Translational aspects of novel findings in genetics of male infertility—status quo 2021. Br Med Bull. 2021; 140(1): 5–22.
doi: [10.1093/bmb/ldab025](https://doi.org/10.1093/bmb/ldab025)

Colaco S., Modi D. Genetics of the human Y chromosome and its association with male infertility. *Reprod Biol Endocrinol.* 2018; 16: 14.
doi: [10.1186/s12958-018-0330-5](https://doi.org/10.1186/s12958-018-0330-5)

Harton G.L., Tempest H.G. Chromosomal disorders and male infertility. *Asian J Androl.* 2012; 14(1): 32–39.
doi: [10.1038/aja.2011.66](https://doi.org/10.1038/aja.2011.66)

Kuroda S., Usui K., Sanjo H. и соавт. Genetic disorders and male infertility. *Reprod Med Biol.* 2020; 19(4): 314–322.
doi: [10.1002/rmb2.12336](https://doi.org/10.1002/rmb2.12336)

Massart A. и соавт. Genetic causes of spermatogenic failure. *Asian J Androl.* 2012; 14(1): 40–48.
doi: [10.1038/aja.2011.67](https://doi.org/10.1038/aja.2011.67)

Cerván-Martín M. и соавт. Genetic Landscape of Nonobstructive Azoospermia and New Perspectives for the Clinic. *J Clin Med.* 2020; 9(2): 300.
doi: [10.3390/jcm9020300](https://doi.org/10.3390/jcm9020300)

Papadimitriou S., Gravel B., Nachtegaal C. и соавт. Toward reporting standards for the pathogenicity of variant combinations involved in multilocus/oligogenic diseases. *HGG Adv.* 2023; 4(1): 100165.
doi: [10.1016/j.xhgg.2022.100165](https://doi.org/10.1016/j.xhgg.2022.100165)

Kousi M., Katsanis N. Genetic Modifiers and Oligogenic Inheritance. *Cold Spring Harb Perspect Med.* 2015; 5(6): a017145.
doi: [10.1101/cshperspect.a017145](https://doi.org/10.1101/cshperspect.a017145)

РАЗДЕЛ 2. ПОЛИМОРФИЗМ ГЕНОВ

ВВЕДЕНИЕ

Геномная карта строится на основе результатов секвенирования ДНК генома. По результатам определения последовательностей азотистых оснований мы можем идентифицировать гены и расстояния между ними. Анализируя нуклеотидную последовательность ДНК, принимающую участие в кодировании данного белка, мы можем оценить функции, выполняемые данной нуклеотидной последовательности. Существует предположение о том, что естественный отбор предотвращает накопление вредных мутаций в последовательностях, кодирующих функциональные продукты (белок). Следовательно, мы можем предположить, что **неповрежденная** кодирующая последовательность приведет к образованию функционально активного полипептида. Сравнивая последовательность ДНК дикого типа (не мутировавший, исходный ген) с последовательностью мутантного гена, можно определить природу мутации и ее точное местоположение в последовательности. Первоначальный менделевский взгляд на геном классифицировал аллели либо, как дикий тип, либо, как мутантный. Впоследствии была признана возможность существования множества аллелей для одного гена в популяции, каждая из которых может по-разному влиять на фенотип. В некоторых случаях может быть даже нецелесообразно определять какой-либо один аллель в качестве дикого типа.

Сосуществование нескольких аллелей в одном локусе гена в данной популяции называется генетическим полиморфизмом. Любой участок, на котором множественные аллели существуют как стабильные компоненты популяции, по определению полиморфен. Локус обычно определяется как полиморфный, если два или более аллеля присутствуют с частотой встречаемости более 1% в популяции. Цвет человеческих глаз является хорошим примером фенотипического полиморфизма, возникающего в результате лежащего в его основе генетического полиморфизма. Не существует единого “нормального” цвета глаз; у разных людей встречается много разных цветов, при этом различия в зрительных функциях между ними практически отсутствуют. Что лежит в основе полиморфизма между различными аллелями? Они обладают различными мутациями, которые могут изменить функцию их продукта, тем самым вызывая изменения в фенотипе. Популяционная динамика этих различных аллелей частично определяется их избирательным воздействием на фенотип. Дикий тип сам по себе может быть полиморфным. Множественные версии аллеля дикого типа можно отличить по различиям в последовательности, которые не влияют на их функцию и, следовательно, не приводят к фенотипическим вариантам.

С точки зрения молекулярного механизма, тремя основными типами вариаций на уровне ДНК (маркерами полиморфизма ДНК) являются, во-первых, однонуклеотидные изменения - SNP; во-вторых, вставки или делеции (Indels) различной длины в диапазоне от 1 до нескольких сотен пар оснований и, в-третьих VNTR - вариации числа tandemных повторов нуклеотидов. Механизмы мутации приводят либо к **переходам**: пури-пуриновый ($A \leftrightarrow G$) или пиримидин-пиримидиновый ($C \leftrightarrow T$) обмены, либо к **трансверсиям**: пури-пиримидиновый или пиримидин-пуриновый ($A \leftrightarrow C$, $A \leftrightarrow T$, $G \leftrightarrow C$, $G \leftrightarrow T$) замены. **При этом наблюдается явное преобладание именно переходов.** В чем причина растущей популярности SNP с точки зрения предоставляемой исследовательской или диагностической генетической информации? Дело в том, что более современная концепция SNP в основном возникла из-за потребности в очень высокой плотности генетических маркеров для изучения многофакторных заболеваний и недавнего прогресса в методах обнаружения полиморфизма и генотипирования (Vignal A. et al., 2002).

МУТАЦИИ – ПРИЧИНА ПОЛИМОРФИЗМА ГЕНОВ

Таким образом, мы понимаем, что **генетическая варибельность (отличия первичной структуры ДНК, обусловленные SNP, Indels или VNTR) внутри одного биологического вида называется генетическим полиморфизмом.**

Первичные структуры молекул ДНК (последовательность нуклеотидов) у всех людей (за исключением монозиготных близнецов) отличаются. Существенные популяционные, этнические и, главное, индивидуальные отличия первичных структур ДНК, включая как их **кодирующие участки (экзоны)**, так и **некодирующие последовательности (межгенные последовательности, интроны, регуляторные участки и прочее)** вызваны разнообразными мутациями, которые являются **причиной полиморфизма генов.** Генетический полиморфизм, как правило рассматривают в качестве менделевского признака, который встречается в популяции по крайней мере в 2 вариантах гена с частотой не менее 1% для каждого аллеля. Собственно, аллельным геном считается вариант гена,

частота встречаемости которого в популяции составляет не менее 1%. Если частота встречаемости гена менее 1% его называют редкой формой гена. Редкие формы генов также могут иметь важное медицинское значение, если они являются патологическими или предрасполагают к какому-либо заболеванию.

Мутации, ставшие причиной полиморфизма генов, можно обозначить, как, во-первых, **качественные**. Т.е. когда количество нуклеотидов в цепи ДНК сохраняется без изменений, однако происходят замены нуклеотидов. И, во-вторых, **количественным**, когда в нити ДНК изменяется число нуклеотидов. Например, может варьировать кратность повторов нуклеотидных последовательностей различной протяженности. Обе названные причины полиморфизма генов могут наблюдаться в кодирующих участках гена, интронах, промоторах и во внегенных последовательностях молекулы ДНК.

Качественный генетический полиморфизм — представлен преимущественно однонуклеотидными заменами, так называемыми single nucleotide polymorphism (SNP – читается, как «СНИП»). Данный вид полиморфизма генов встречается наиболее часто. На ранних этапах сравнительного изучения генома человека на примерах представителей разных рас и этнических групп было установлено глубокое генетическое родство всех людей на планете. Помимо этого, была получена ценная информация о процессах антропогенеза, расселения человека и его предков по планете о возможных путях формирования этнических групп. При этом около 70% SNP в популяции оказываются нейтральными, т.е. не оказывают существенного влияния на функцию кодируемого белка, не снижают жизнеспособность особи, ее плодовитость и жизнеспособность потомства. Тем не менее, нейтральные мутации могут составлять основу индивидуальных особенностей организма, формируя индивидуальное разнообразие особей. В тоже время, остальные 30% SNP вовлечены в патогенез моногенных или полигенных заболеваний. Таким образом, полногеномный анализ SNP актуален, как для фундаментальной биологии, так и для практической медицины.

Количественный полиморфизм генов — можно проиллюстрировать вариациями числа tandemных повторов (STR — Short Tandem Repeats), которые в зависимости от числа нуклеотидов в повторяющихся группах подразделяют на микросателлитную ДНК либо минисателлитную ДНК (пояснения ниже). Наконец, повторяющиеся группы нуклеотидов могут иметь достаточно протяженную последовательность — так называемые VNTR (Variable Number Tandem Repeats). В кодирующих участках гена (экзонах) присутствуют преимущественно тринуклеотидные повторы, поскольку они не приводят к сдвигу рамки считывания. Наиболее часто такие тринуклеотидные повторы представлены последовательностью нуклеотидов CAG (cytosine–adenine–guanine) — триплет, кодирующий глютаминовую кислоту. В большинстве случаев кратный прирост (или убавление) дополнительных повторов носит нейтральный характер и также, как и нейтральные SNP способствует индивидуальному разнообразию.

Между тем, генные полиморфизмы такого типа характерны для генов «болезней экспансии». В случае динамических мутаций (изменение числа, преимущественно, тринуклеотидных повторов) достижение определенного критического уровня кратности копий тринуклеотидного повтора генетический полиморфизм перестает быть функционально нейтральным и демонстрирует себя как специфический тип уже патологических мутаций -

«динамических мутаций». Патологическое увеличение числа повторов может быть локализовано в кодирующих или регуляторных частях гена, сопровождающееся нарушением его функции. Такого рода мутации могут приводить, либо к появлению патологического белка, либо к нарушению процессов экспрессии гена. Динамические мутации рассматриваются в качестве причины ряда нейродегенеративных заболеваний (хорея Гентингтона, болезнь Кеннеди, спиноцеребеллярная атаксия и др.). Характерными клиническими особенностями таких заболеваний являются: поздний возраст клинических проявлений, повышение тяжести течения заболеваний в последующих поколениях, отсутствие эффективных методов лечения.

SNP, SINGLE NUCLEOTIDE POLYMORPHISM – ОДНОНУКЛЕОТИДНЫЙ ПОЛИМОРФИЗМ

Однонуклеотидный полиморфизм это замена одного азотистого основания в последовательности ДНК, по сравнению с исходной формой гена. Полиморфизм одного нуклеотида присутствует в конкретном нуклеотидном сайте. Например, некоторые молекулы ДНК в одной и той же популяции могут иметь пару оснований Т–А в определенном нуклеотидном сайте, тогда как другие молекулы ДНК в той же популяции могут иметь пару оснований С–G в том же месте. Эта разница составляет SNP (single nucleotide polymorphism). SNP определяется двумя аллелями, для которых может быть три генотипа среди людей в популяции – гомозиготные хромосомы (или гетерозиготные хромосомы с Т–А в одной хромосоме и С–G в гомологичной хромосоме).

Таким образом, когда сравниваются аллели одного и того же локуса, то выявленная разница в одном нуклеотиде называется однонуклеотидным полиморфизмом (SNP). В среднем один SNP встречается примерно на каждые 1330 оснований в геноме человека. Благодаря специфичному ландшафту SNP в последовательности ДНК, каждый человек уникален. Присутствие SNP может быть обнаружено путем прямого сравнения данных последовательностей от разных индивидуумов. Одной из целей генетического картирования является получение каталога распространенных вариантов генов. Наблюдаемая частота SNP на один геном предсказывает, что в человеческой популяции в целом (учитывая геномы всех людей данной популяции) должно быть более 10 миллионов SNP, которые встречаются с частотой более 1% (т.е. являются полиморфными).

SNP составляют около 90% всех генетических вариаций человека. В определенном локусе может существовать множество различных вариантов последовательности. Некоторые из этих последовательностей оказывают влияние на фенотип, но другие "скрыты", потому что не оказывают видимого эффекта. В дополнение к информации о частоте встречаемости аллелей, также важно для выбора SNP их геномное местоположение и их предполагаемая функциональная значимость (кодирующий, интронный, промоторный и т.д.) (Xu H. et al., 2005). **В локусе может быть множество изменений, включая те, которые изменяют последовательность ДНК, но не изменяют последовательность полипептидного продукта, те, которые изменяют последовательность полипептида без изменения его функции, те, которые приводят к полипептидам с различными функциями, и те, которые приводят к измененным полипептидам, которые не являются функциональными.** Обычно считается, что полная человеческая последовательность ДНК содержит, по меньшей мере, миллион SNP, как кодирующих областей, так и в интронах и промоторах.

Не вызывает сомнений, что генетическая изменчивость является повсеместным явлением и формирует генетическую основу видового разнообразия. С завершением проекта "Геном человека" изучение генетических вариаций стало одним из краеугольных камней биомедицинских исследований. Это произошло не только потому, что варианты генов влияют на разнообразие антропометрических характеристик человека, но и потому, что они оказывают влияние на риск развития заболеваний и на реакцию организма в ответ на экологические проблемы. Информация, полученная в результате изучения наследственной изменчивости, не только углубляет наше понимание человеческих генов и эволюции, но также приносит пользу в выявлении и лечении наследственных заболеваний человека.

Поскольку наиболее распространенными вариантами являются SNP, этот термин часто используют, как синонимом вариации генетической последовательности. Однако здесь мы ограничиваем его использование формальным определением SNP: одно изменение положения азотистого основания в одной позиции. В результате высокопроизводительных методов генотипирования в последние годы были зарегистрированы миллионы SNP человека. Например, было проведено несколько исследований, посвященных оценке того, как SNP влияют на фенотип. В частности, наблюдаются биологические эффекты аминокислотных замен, выполняемых несинонимичными кодирующими SNP. Изучается влияние несинонимичных кодирующих SNP на функцию белка, вызванных изменением последовательности аминокислот в первичной структуре белка.

Другой областью генома, важной для функционирования генов, являются **регуляторные области генов**. Промоторы генов, посредством связывания специфических факторов транскрипции непосредственно участвуют в инициации и регуляции транскрипции генов. Таким образом, изменение последовательности в промоторах генов может изменять идентификацию и связывание факторов транскрипции, что, в свою очередь, может влиять на экспрессию генов и оказывать биологическое воздействие. Например, было обнаружено, что один аллель гена HLAG (главного комплекса гистосовместности), продукты которого ингибируют материнский антифетальный иммунный ответ, в значительной степени связан с повышением риска внутриутробной гибели плода. Одним из возможных объяснений является то, что SNP попадает в сайт связывания фактора ответа интерферона-1 (IRF-1), влияя на связывание IRF-1 и понижающую регуляцию транскрипции гена HLA-G. Поэтому, самостоятельным направлением исследований является изучение распределения SNP в регуляторных областях генов человека. Предполагаемые сайты связывания факторов транскрипции в промоторах генов были вычислены и сопоставлены с ранее идентифицированными SNP. Результаты показывают, что SNP обладают специфическими показателями распределения, как в регуляторных областях генов, так и в сайтах связывания факторов транскрипции по сравнению с частотой встречаемости SNP в других участках ДНК.

В регуляторных участках генов было обнаружено более 35 000 SNP. Более 99,8% из них представлены двумя аллелями. Частоты встречаемости замещения нуклеотидов промоторных SNP с двумя аллелями показаны в таблице 1. В литературе высказывается мнение о том, что объяснение высокой частоты встречаемости SNP в области промотора может состоять в том, что SNP участвуют в изменении экспрессии генов в процессе эволюции, оказывая влияние на последовательности сайтов

связывания промотора. Эффекты SNP в месте связывания факторов транскрипции, не поддаются однозначной интерпретации. В то время, как некоторые изменения последовательности связывания, вносимые SNP, могут полностью прерывать экспрессию гена, другие могут влиять только на уровень экспрессии (Y. Guo и D. C. Jamison, 2005).

Таблица №1 Частота встречаемости замещения нуклеотидов промоторных SNP с двумя аллелями в регионе промотора. Цитировано по Y. Guo и D. C. Jamison, 2005.

Purine A ↔ G 33.10%

Pyrimidine C ↔ T 33.10%

Purine ↔ Pyrimidine C ↔ G 8.93%

Purine ↔ Pyrimidine A ↔ C 8.77%

Purine ↔ Pyrimidine G ↔ T 8.82%

Purine ↔ Pyrimidine A ↔ T 7.42%

В процессе эволюции необходимость адаптации определенного уровня экспрессии генов к определенным факторам окружающей среды формирует естественный отбор для регуляции генов. В этом смысле SNP, встречающиеся в сайте связывания факторов транскрипции, оказывают прямое влияние на уровень экспрессии генов. Следовательно, тот факт, что большее количество SNP находится в сайтах связывания факторов транскрипции, может демонстрировать различные требования к экспрессии разных генов в разных условиях. Таким образом, промоторные SNP могут быть активным фактором естественного отбора при выборе режима транскрипции генов.

При этом необходимо особо подчеркнуть, что SNP, связанные с заболеванием, по данным GWAS (Genome wide association studies), чаще всего встречаются в некодирующих специфических регуляторных участках ДНК - transcription factor binding sites (TFBS) – сайтах связывания факторов транскрипции (Chen C. et al., 2014). Анализ дифференциального сродства к связыванию факторов транскрипции между обычными аллелями и аллелями риска выявил функциональные SNP с потенциальным нарушением или усилением связывания факторов транскрипции. Авторами установлены регуляторные SNP-кандидаты, один из которых обладает пограничной восприимчивостью к раку легких и со значительной восприимчивостью к раку молочной железы.

Многие функциональные исследования были сосредоточены на SNP, расположенных в кодирующих областях, которые могут влиять на фенотип, изменяя первичную структуру кодируемых белков. Вместе с тем, SNP также могут вызывать преждевременную терминацию трансляции и распад мРНК. Другая функция SNP заключается в том, что они влияют на сайты сплайсинга, что приводит к возникновению альтернативного сплайсинга.

Секвенирование полных индивидуальных геномов в настоящее время широко используется и позволяет оценивать индивидуальные вариации на уровне ДНК, как нейтральные SNP, так и те, которые связаны с заболеваниями или восприимчивостью к болезням. В настоящее время секвенированы сотни индивидуальных геномов человека всех основных расовых групп, включая геномы денисовцев (вид Номо эпохи палеолита, живший более 30 000 лет назад) и неандертальцев (возраст более 25 000 лет).

Проект "1000 геномов" проводился с 2008 по 2015 год с целью выявления распространенных генетических вариантов человека путем глубокого секвенирования по меньшей мере 1000 геномов человека; окончательное число фактически составило 2 504 анонимных последовательности генома человека, представляющих 26 человеческих популяций. Большая доля полиморфных сайтов означает, что каждый индивид обладает уникальным набором SNP. **Конкретная комбинация участков, обнаруженных в определенном регионе, называется гаплотипом и представляет собой небольшую часть полного генотипа.** Термин "гаплотип" был первоначально введен для описания генетического содержимого основного локуса гистосовместимости человека, области, определяющей белки, выполняющих важную роль в иммунных механизмах. В настоящее время этот термин расширен для описания конкретной комбинации аллелей или любых других генетических маркеров, присутствующих в некоторой определенной области генома. Используя SNP, была составлена подробная карта гаплотипов генома человека; это позволяет исследователям легче отображать гены, вызывающие болезни.

SNP И НАСЛЕДСТВЕННЫЕ БОЛЕЗНИ

Наследственные болезни человека вызываются различными типами мутаций генов, вставками/делециями (insertions/deletions), геномными структурными вариациями, а также патогенными однонуклеотидными полиморфизмами (SNP), которые тесно связаны с повышенным риском заболеваний. Среди них моногенные заболевания представляют собой серьезную проблему для проведения генной терапии. Эти пациенты унаследовали один патогенный аллель, вызывающий фенотип заболевания, особенно доминантно-негативным образом, а также один нормальный аллель. Таким образом, стратегия лечения обычно включает специфическое подавление патогенных аллелей, не затрагивая аллели дикого типа. Генетическое заболевание полностью или частично вызвано аномалией в генетическом составе генома индивида. В геноме человека наиболее частыми вариациями являются SNP. Средний геном человека содержит примерно 4-5 миллионов SNP, которые встречаются в среднем каждые 1000-2000 нуклеотидов, что обеспечивает достаточную плотность для всестороннего анализа гаплотипов. Поскольку однонуклеотидные полиморфизмы (SNP) являются наиболее распространенной формой полиморфизма в геноме человека, они широко используются для исследований в области антропогенеза, исследований наследственных заболеваний, в исследованиях в области фармакогенетики. В последние годы было обнаружено большое количество генов, способствующих моногенным и мультифакторным (полигенным) заболеваниям. В связи с этим специализированные базы данных биоинформатики и программное обеспечение, относящиеся к сбору и анализу генетических данных, значительно увеличились в количестве, размере и объеме. Однонуклеотидные полиморфизмы (SNP), составляющие наиболее распространенный тип генетической изменчивости, в настоящее время являются основным исходным материалом, лежащим в основе большинства генетических исследований и баз данных.

С одной стороны, другие типы вариаций, включая индели, варианты числа копий и эпигенетические маркеры, по-прежнему важны для рассмотрения и могут влиять на заболевание. С другой стороны, SNP в значительной степени легче всего определить и являются наиболее полезными и широко применяемыми маркерами в генетических исследованиях в современную эпоху (Johnson A.D., 2009). Как правило, многие SNP не влияют на функцию клеток, но сообщается, что некоторые SNP тесно связаны с заболеваниями или влияют на реакцию клеток на лекарственное средство. Хотя более 99% последовательностей ДНК человека одинаковы во всех популяциях, некоторые SNP могут оказывать существенное влияние на то, как люди реагируют на неблагоприятные факторы среды (патогенные бактерии, вирусы, токсины и химические вещества), а также лекарства и другие методы лечения. Это делает SNP очень ценными для биомедицинских исследований, разработки фармацевтических продуктов и медицинской диагностики. Генетические маркеры обеспечивают основу для чрезвычайно продуктивного метода идентификации генетических вариантов на молекулярном уровне. Типичная проблема касается мутации с известными эффектами на фенотип, когда соответствующий генетический локус может быть помещен на генетическую карту, но в отношении которого у нас нет никаких знаний о соответствующем гене или его продукте. Многие опасные и даже смертельные заболевания человека попадают именно в эту категорию. Например, муковисцидоз демонстрирует рецессивное менделевское наследование, но молекулярная природа мутантного белка была неизвестна до тех пор, пока ее не удалось идентифицировать на основе характеристики гена. В геноме SNP встречаются или рядом с каким-либо конкретным геном-мишенью, или внутри этого гена. Исследователи могут идентифицировать такие маркеры благодаря их тесной связи с геном, ответственным за мутантный фенотип. Если мы сравним ДНК пациентов, страдающих расстройством, с ДНК здоровых людей, мы можем обнаружить, что определенные маркеры всегда присутствуют (или всегда отсутствуют) у пациентов.

Патологический фенотип не всегда может определяться одним геном; чаще всего наследственная патология может быть полигенной или многофакторной. Кроме того, некоторые SNP, ассоциированные с заболеванием, могут не иметь функционального отношения к расстройству. Однако их полезность при анализе сложных генетических признаков может быть проиллюстрирована двумя SNP, которые были связаны с компонентами фенотипа заболевания. Эти SNP включают мутацию, которая приводит к усилению функции в гене рецептора интерлейкина-4, который связан с тяжелой формой астмы, и мутацию в гене β 2-адренергического рецептора (ADRB2), который связан с большей устойчивостью к адренергическим препаратам. Развитие современных технологий в молекулярной биологии привели к пониманию актуальности поиска SNP, ассоциированных с такими широко распространенными полигенными заболеваниями, как остеопороз, диабет, сердечно-сосудистые и воспалительные заболевания, психические расстройства и большинство видов рака (Bin Alwi Z., 2005).

Широко распространенный семейный характер сложных полигенных заболеваний ясно указывает на значительный генетический компонент. Однако, в отличие от моногенного типа наследования заболеваний, этот генетический элемент состоит из множества вариантов генов, каждый из которых вносит небольшой вклад. При наличии определенного набора генов предрасположенности в генотипе человека, вклад данных генов мультиплицируется, провоцируя патологию при определенном сочетании с факторами внешней среды. Следует принять во внимание, что SNP многочисленны и более стабильны, чем тандемные повторы из-за более низкой частоты мутаций. Многие SNP могут находить отражение в фенотипе, если они встречаются в кодирующих или регуляторных областях гена. Поэтому, используя SNP-маркеры, часто можно проверить наличие прямой связи между фенотипом и функциональным вариантом. По этим причинам SNP предпочтительнее для составления карт генетических маркеров высокой плотности, необходимых для одного из основных направлений исследований генетики человека: выявления сложных генетических признаков.

SNP И САХАРНЫЙ ДИАБЕТ

Диабет 2 типа представляет собой наследственное полигенное (мультифакторное) заболевание. Современные результаты исследований показывают, что наследственность является важным фактором, способствующим этому заболеванию. Во-первых, некоторые этнические меньшинства и группы коренного населения с низкой примесью населения (например, индейцы пима, жители Микронезии и других тихоокеанских островов, австралийские аборигены и американцы мексиканского происхождения) демонстрируют исключительно высокую распространенность диабета 2 типа (до 21% у индейцев Пима). Во-вторых, наблюдаются группы семей, предрасположенных к диабету 2 типа. Члены данных семей, состоящие в кровном родстве в целом, достоверно более предрасположены развитию заболевания, по сравнению с населением популяций, в которых они проживают. Наконец, использование близнецового метода медицинской генетики продемонстрировало заметно более высокую корреляцию степени риска диабета 2 типа у монозиготных по сравнению с дизиготными близнецами (70% против 10%) (Staiger H. et al., 2009). Диабет 2 типа, как и другие полигенные заболевания, не следует простому менделевскому наследованию и считается полигенным заболеванием. Сложные заболевания, такие как диабет 2 типа, вызваны одновременным возникновением общих вариаций последовательности ДНК во многих генах в сочетании с определенными факторами среды. Предполагается, что каждое из этих изменений ДНК оказывает лишь умеренное воздействие на функцию затронутых генов. Тем не менее, в сумме эти вариации придают повышенную восприимчивость к неблагоприятным факторам окружающей среды. Однонуклеотидные полиморфизмы (SNP), замены в одиночных парах оснований, охватывают примерно 90% вариаций последовательностей в геноме человека. По этой причине они рассматриваются, в качестве основных детерминант индивидуальной предрасположенности к таким сложным заболеваниям.

В течение последнего десятилетия было идентифицировано по меньшей мере 600 (подтвержденных и потенциальных) генов предрасположенности к диабету. Наибольший успех в генетике диабета 2 типа был достигнут благодаря исследованиям по принципу «случай/контроль» в больших группах пациентов и контрольных испытуемых. Большинство генов можно считать подтвержденными у многих этнических групп, например: *TCF7L2*, *SLC30A8*, *HHEX*, *CDKAL1*, *CDKN2A/B*, *IGF2BP2* и *FTO*. Лептин и грелин - два важных пептида, регулирующих аппетит и энергетический баланс. Показано, что общие полиморфизмы в генах, кодирующих эти пептиды и связанные с ними рецепторы, связаны с массой тела, различными маркерами ожирения и метаболическими нарушениями, которые рассматривают, как предпосылки сахарного диабета 2-го типа (Ghalandari H. et al., 2015). Адипонектин - белковый гормон, выделенный из адипоцитов, который повышает чувствительность к инсулину и улучшает состояние при диабете за счет усиления окисления жирных кислот и поглощения глюкозы в скелетных мышцах и снижения выработки глюкозы в печени. Низкие концентрации адипонектина в сыворотке крови связаны с диабетом, ожирением,

резистентностью к инсулину и метаболическим синдромом. Ген адипонектина расположен на хромосоме 3q27. Предполагается, что однонуклеотидные полиморфизмы (SNP) в гене адипонектина (*ADIPOQ*) были связаны с диабетом. SNP в *ADIPOQ* помогают оценить связь распространенных вариантов с уровнями адипонектина и риском развития диабета. Результаты исследований свидетельствуют о том, что rs1501299, rs2241766, rs266729, rs17366743, rs17300539, rs182052, rs822396, rs17846866, rs3774261 и rs822393 в значительной степени связаны с сахарным диабетом 2 типа, который является преобладающей формой заболевания. Кроме того, rs2241766 и rs266729 широко ассоциированы с гестационным диабетом, состоянием, которое развивается у женщин во время беременности. Следовательно, не конкретная единичная мутация, а ряд SNP в гене адипонектина могут быть фактором риска развития диабета у людей во всем мире. Это исследование свидетельствует о том, что адипонектин играет важную роль в патогенезе сахарного диабета 2 типа и гестационного диабета (Howlader M. et al, 2021).

SNP И ОНКОЛОГИЧЕСКИЕ ЗАБОЛЕВАНИЯ

По мнению некоторых авторов, использование маркеров полиморфизма генов, в частности однонуклеотидных полиморфизмов (SNP), обещает предоставить комплексный инструмент для анализа генома человека и выявления тех генов и областей генома, которые способствуют канцерогенезу (Engle L.J. et al., 2006). Авторами анализируются различные методы определения SNP с учетом того, как каждая из этих методологий может быть использована в целях определения местоположения генов и областей генома, представляющих интерес для различных типов рака. Высказывается мнение о том, что, как и при других мультифакторных наследственных заболеваниях, более 90% SNP, способных индуцировать онкологические заболевания, локализованы в областях ДНК, не принимающих участия в кодировании белка. Этот факт затрудняет расшифровку функциональных воздействий SNP на процесс развития патологии. Было предложено множество механизмов для объяснения того, как некодирующие варианты геном могут оказывать влияние на заболевание человека, такие как нарушение сплайсинга, трансляции или стабильности гена, кодирующего белок. В дополнение к генам, кодирующим белок, SNP, связанные с риском, идентифицированные GWAS, также влияют на ключевые некодирующие гены для miRNAs и lncRNAs (Liu S. et al., 2017). Недавние исследования показали, что из числа 996 SNP связанных с онкологическими заболеваниями, выявленными в исследованиях GWAS, идентифицировано 575 SNP во фрагментах, которые положительно регулируют экспрессию генов, и 758 SNP во фрагментах с отрицательной регуляторной активностью. Среди них 70 вариантов являются регуляторными вариантами, для которых два аллеля обеспечивают различную регуляторную активность. Анализ двух регуляторных вариантов —rs11055880, связанного с риском рака молочной железы, и rs12142375, связанного с риском лейкемии, позволил установить их регуляторную активность в отношении экспрессии генов *ATF7IP* и *PDE4B*. Возможно, ценность данных исследований заключается в том, что определяя регуляторные варианты, связанные с восприимчивостью к раку, и изучая их молекулярные функции, можно повысить достоверность оценки риска развития рака.

Высказывается мнение о том, что в процессах канцерогенеза, помимо накопления унаследованного полиморфизма (в т.ч. SNP) имеет значение утрата одного из аллелей (потере гетерозиготности - loss of heterozygosity LOH) через ряд возможных механизмов. В свою очередь, потеря гетерозиготности является наиболее распространенным молекулярно-генетическим изменением, наблюдаемым при раке человека с преобладанием инактивации генов-супрессоров опухолей (Zheng H.-T. et al., 2005).

VNTR (SSR) - ВАРИАЦИИ ЧИСЛА ТАНДЕМНЫХ ПОВТОРОВ НУКЛЕОТИДОВ

Геном человека содержит множество повторяющихся последовательностей ДНК, которые различаются по сложности повторяющейся единицы от одного нуклеотида до целого гена. Повторяющиеся последовательности могут быть широко распределены или в виде простых tandemных массивов. Массивы до 5 или 6 нуклеотидов известны как простые tandemные повторы, и они широко рассеяны и сильно полиморфны. Представители одной группы простых tandemных повторов, тринуклеотидных повторов, могут подвергаться увеличению числа копий в процессе мутации. **Микросателлиты, или tandemные простые повторы нуклеотидных последовательностей (SSR)**, широко распространены в геномах и демонстрируют высокий уровень полиморфизма. Геном человека содержит множество нуклеотидных последовательностей, которые встречаются неоднократно. Эти повторяющиеся последовательности различаются по сложности от полных генов (таких как гены рибосомной РНК) до простых последовательностей из одной или нескольких пар оснований. Организация повторяющихся последовательностей может варьировать от широко рассеянных копий относительно длинной сложной последовательности до tandemных массивов с простым составом

последовательностей. К числу простейших и наиболее распространенных повторов относятся динуклеотидные повторы, в основном имеющие основания АС на одной нити ДНК и GT на другой. Существуют и другие простые tandemные повторы (STR), которые включают моно-, три-, тетра- и пентануклеотидные повторяющиеся единицы. Геномные микросателлиты (простые повторы последовательности; SSR), итерации (повторение) нуклеотидных мотивов 1-6 п.н., были обнаружены в геномах каждого проанализированного в настоящее время организма. Вместе с тем, SSR составляют довольно большую долю некодирующей ДНК и относительно редки в областях, кодирующих белок. Например, все наблюдаемые моно-, ди- и тетрануклеотидные SSR были расположены в некодирующих областях у 54 видов растений. Все типы SSR (от моно- до гексануклеотидных повторов) были обнаружены в больших количествах в некодирующих областях генома беспозвоночных, растений и приматов (Li Y.-C. et al., 2002). По мнению авторов, тот факт, что все типы SSR (за исключением триплетов и гексануклеотидов) значительно реже встречаются в последовательностях, кодирующих белок, по сравнению с некодирующими участками ДНК, можно объяснить отрицательным отбором против мутаций в кодирующих областях. Тем не менее, триплетные повторы, связанные с заболеванием, в основном, встречаются в кодирующих областях генома человека. Некоторые три-повторы не сохраняются в течение длительных периодов времени, даже когда они образуют части последовательностей, кодирующих белки, поскольку длинные три-повторы (например, массивы CAG) могут быть дестабилизированы во время мейоза (кроссинговера). Большинство SSR (48-67%), обнаруженные у многих видов, представляют собой динуклеотиды, но у приматов моонуклеотидные (в основном, поли (A/T) треки) являются наиболее распространенными классами SSR. Динуклеотидные SSR могут быть обнаружены в интронах. Например, у мыши интрон А гена *Adhl* содержит (TA)₁₄ (TG)₈ и (TA)₁₉, а интрон гена *IL-5* содержит (AT)₁₇. Увеличение размеров ди- или тетрануклеотидных SSR в интронах может привести к разрушению исходного белка или образованию новых генов путем сдвига рамки считывания. Когда рассматривается встречаемость SSR в различных функциональных областях генома, оказывается, что большинство из них демонстрируют гораздо более высокую плотность в некодирующих областях. Исключением из этого правила являются **тримеры и гексамеры**, которые почти в **два раза чаще встречаются в экзонах по сравнению с интронами** и межгенными областями. Их высокая частота в кодирующих областях может быть объяснена тем фактом, что **они не изменяют рамок считывания и свойства генетического кода**. Таким образом, данный тип мутаций не оказывает существенного негативного влияния, в сравнении с другими типами SSR. Их положительный отбор в экзонах также предполагает некоторую функцию повторов (Jasinska A., Krzyzosiak W.J., 2004).

Локусы с более длинными повторяющимися единицами, по-видимому, подвергаются более сильному отбору из-за разницы в размерах, особенно в областях генома с высокой скоростью рекомбинации. Эти результаты также свидетельствуют о том, что различия между частотой встречаемости SSR в кодирующих и некодирующих участках генома возникают в результате специфического отбора против мутаций со сдвигом рамки считывания в кодирующих областях, возникающих в результате изменений длины в нетриплетных повторах. Тем не менее, **14% всех белков содержат повторяющиеся последовательности**, причем количество повторов у эукариот в три раза выше, чем у прокариот. Семейства прокариотических и эукариотических повторов сгруппированы в негомологичные белки. Это может указывать на то, что повторяющиеся последовательности возникли после того, как эти два царства разделились. Эукариоты, включающие больше повторов, могут иметь эволюционное преимущество в более быстрой адаптации к новым условиям. Высказывается мнение о том, что различия между кодирующими и некодирующими частотами SSR возникают в результате специфического отбора против мутаций со сдвигом рамки считывания в кодирующих областях.

Мутационные механизмы вариации SSR

Геномное изобилие SSR и различные функции и эффекты (предполагаемые или достоверно установленные) связаны с частотой мутаций, поскольку частота мутаций SSR (10^{-2} — 10^{-6} событий на локус за поколение) очень высока по сравнению с частотой точечных мутаций в локусах кодирующих генов. Хотя процесс мутации, по-видимому, демонстрирует явные различия между видами, типами повторов, локусами и аллелями, возрастом и полом, нестабильность проявляется преимущественно в виде изменений числа SSR-повторов. Для объяснения таких высоких показателей мутации можно использовать два мутационных механизма. Первый связан с проскальзыванием ДНК во время репликации ДНК. Второй включает рекомбинацию между нитями ДНК. Эффективность этих двух механизмов может предположительно зависеть от условий окружающей среды. Было обнаружено, что различные факторы влияют на частоту мутаций в локусах SSR, включая повторяющийся мотив, размер аллеля, положение хромосом, содержание GC в фланкирующей ДНК, деление клеток (митотическое против мейотического), пол и генотип (например, мутации в генах *MMR*). Для объяснения таких высоких показателей мутации можно использовать два мутационных механизма. Первый связан с состоянием процессов репарации ДНК во время репликации. Второй включает рекомбинацию между нитями ДНК во

время кроссинговера. Баланс эффективности этих двух механизмов может зависеть от условий окружающей среды.

SSR И ТРАНСКРИПЦИЯ

Многочисленные доказательства показывают, что SSR, расположенные в промоторных областях, могут влиять на активность генов. Было обнаружено, что (TC)_n-повторяющиеся последовательности нуклеотидов в промоторных областях служат транскрипционным элементом для гена белка теплового шока *hsp26* у различных видов. Делеция различных ди-, три- и тетра-SSR-повторов заметно изменяла транскрипционную активность. Например, транскрипционная активность промоторов была резко снижена путем удаления (TCCC)_n-повторов из промоторных областей гена *TGF-β3* в системе экспрессии *CAT*. С другой стороны, (GT)_n-повтор может усиливать активность гена, особенно, если GT-повтор находится ближе к промоторным последовательностям.

Лишь небольшая часть общей РНК, транскрибируемой в клетках человека, становится зрелой мРНК. Совокупность мРНК составляет транскриптом человека, который может меняться в зависимости от этапа онтогенеза, физиологического состояния и патологических процессов. Небольшая часть различных повторяющихся последовательностей, которые составляют более половины генома человека, сохраняется в зрелых транскриптах и формирует их функцию. В дополнение к мРНК, транскриптом также включает различные некодирующие РНК, играющие либо структурные, либо регуляторные функции в клетках. Следовательно, существуют тысячи различных транскриптомов в сотнях различных типов клеток и органов в их различных физиологических и патологических состояниях. Разнообразие транскриптомов по сравнению с генами увеличивается примерно на 50% благодаря альтернативному сплайсингу, синтезируются многочисленные бессмысловые транскрипты, и все некодирующие РНК, включая недавно признанное семейство микроРНК, также вносят значительный вклад в пул человеческих транскриптов. Анализ присутствия, в частности, триплетных повторов в мРНК показал, что по меньшей мере 718 повторов были обнаружены в 619 мРНК (Jasinska A., Krzyzosiak W.J., 2004). Короткие тандемные повторы (STR) часто встречаются во многих транскриптах, однако в некоторых случаях патологические события происходят при достижении критической длины повтора. Например, это явление наблюдается при различных неврологических расстройствах. Патологические эффекты этих повторов запускаются транскриптами мутантной РНК и/или кодируемыми этими РНК мутантными белками, которые определяются локализациями расширенных повторов в некодирующих или кодирующих областях. Показано, что структуры РНК, образованные этими мутантными повторяющимися трактами РНК, оказывают токсическое воздействие на клетки (Ciesiolka A. et al., 2017).

SSR И ТРАНСЛЯЦИЯ

Многие исследования показывают, что SSR могут влиять на трансляцию генов. Например, сокращение числа повторов (CA)_n увеличивает трансляцию мРНК в клетках *Escherichia coli*. Умеренное расширение повторов CGG может заметно уменьшить трансляцию гена *CAT* (pSVsCAT) при вставке человеческих повторов CGG. Распределение тетрануклеотидов AGCT в геномах *E. coli* и *Bacillus subtilis* предсказывает трансляционный сдвиг рамки в нескольких генах (*infB*, *ACEF/PDHC*, *eno*, *RPLI*, *OmpA*, *OMF* и *TOLA*). Связывание повтора CUG (*CUGBP1*) взаимодействует с 5'-областью мРНК *C/ebpβ* и регулирует трансляцию изоформ *C/EBPBETA*. Триплетные повторы UGG оказывают сильное ингибирующее действие на трансляцию мРНК в *E. coli*, что было доказано в репортерном гене *CAT*.

МЕДИЦИНСКОЕ ЗНАЧЕНИЕ SSR

Около 15% спорадических случаев колоректального рака, а также иных онкологических заболеваний, демонстрируют нестабильность SSR. Прогрессирующее накопление нестабильности SSR также может способствовать развитию рака желудка. Примеры изменений SSR, сопровождающих различные виды рака у человека, можно разделить на два подтипа: тип А, демонстрирующий относительно небольшие изменения в пределах шести пар оснований; и тип В, демонстрирующий резкие изменения в составе восьми пар оснований. При некоторых видах рака была предложена связь между нестабильностью SSR типа В и семейной предрасположенностью. Было показано, что 14 неврологических расстройств являются результатом расширения нестабильных тринуклеотидных повторов. Такие триплетные заболевания SSR включают случаи с изменениями либо в некодирующих (нетранслируемых), либо в кодирующих (экзоны) последовательностях. Расширенные повторы могут придавать генетическую нестабильность и, скорее всего, способствуют подавлению транскрипции. Человеческий массив повторов *FMRI* (CGG)_n может проявлять генетическую нестабильность,

характеризующуюся прогрессирующим расширением в течение поколений, приводящим к репрессии генов и развитию синдрома хрупкости X-хромосомы. Миотоническая дистрофия, аутосомно-доминантное неврологическое расстройство, вызвана расширением CTG-повторов в локусе *DMPK*, при этом у пострадавших лиц $n \geq 50$ повторов (CTG). Аномальное расширение (≥ 39) повторов (CAG), которое приводит к полиглутаминовым растяжениям, вызывает болезнь Мачадо Джозефа. Установлено также, что длина, а не специфический аллель (CA) n повтора гена альдозоредуктазы ассоциируется с диабетической ретинопатией, а полиморфизм аллеля (CA) n в первом интроне человеческого гена интерферона- γ связан с фиброзом аллотрансплантата легкого. Например, аллель (CCTTT) $_{14}$ в промоторе гена *NOS2A* значительно ограничивал диабетическую ретинопатию.

Вместе с тем, приводятся данные о том, что простая экспансия повторов может вызывать некоторые виды рака. Это может произойти путем изменения экспрессии гена-супрессора опухоли и изменением свойств кодируемых белков. Продемонстрированная мейотическая нестабильность может связать этот механизм мутации с семейными раковыми синдромами. Так установлена нестабильность STR в нескольких участках при наследственном неполипозном раке толстой кишки предполагает, что нестабильность последовательности может быть фактором прогрессирования рака. Дальнейшая идентификация генов-кандидатов, содержащих триплетные повторы, должна позволить провести готовую проверку гипотезы о том, что нестабильные последовательности простых повторов могут вызывать рак

РЕКОМЕНДОВАННАЯ ЛИТЕРАТУРА К РАЗДЕЛУ 2. ПОЛИМОРФИЗМ ГЕНОВ

Guo Y., Jamison D. C. The distribution of SNPs in human gene regulatory regions//BMC Genomics.-2005.-Т.6.-С.140 doi:10.1186/1471-2164-6-140

Kim B.-C., Kim W.-Y., Park D. SNP@Promoter: a database of human SNPs (Single Nucleotide Polymorphisms) within the putative promoter regions//BMC Bioinformatics.-2008.-Т.9(Suppl 1).-С. S2 doi:10.1186/1471-2105-9-S1-S2

Vignal A., Milan D., Magali SanCristobal, Eggen A. A review on SNP and other types of molecular markers and their use in animal genetics//Genet. Sel. Evol.-2002.-Т.34.-С.275–305 DOI: 10.1051/gse:2002009

Johnson A.D. SNP bioinformatics: a comprehensive review of resources//Circ Cardiovasc Genet.-2009.-Т.2,№5.-С.530–536 doi:10.1161/CIRCGENETICS.109.872010

Bin Alwi Z. The use of snps in pharmacogenomics studies//Malaysian Journal of Medical Sciences.-2005.-Т.12,№2.-С.4-12

Chen C., Chang I.-S., Hsiung C.A., Wasserman W.W. On the identification of potential regulatory variants within genome wide association candidate SNP sets//BMC Medical Genomics.-2014.-Т.7.-С.34 <http://www.biomedcentral.com/1755-8794/7/34>

Xu H., Gregory S.G., Hauser E.R. et al. SNPselector: a web tool for selecting SNPs for genetic association studies//Bioinformatics.-2005.-V.21,N22.-С.4181–4186 doi:10.1093/bioinformatics/bti682

Ghalandari H., Hosseini-Esfahani F., Mirmiran P. The Association of Polymorphisms in Leptin/Leptin Receptor Genes and

Ghrelin/Ghrelin Receptor Genes With Overweight/Obesity and the Related

Metabolic Disturbances: A Review// Int. J. Endocrinol. Metab.-2015.-Т.13,№3.- С.e19073 DOI: 10.5812/ijem.19073v2

Howlader M., Sultana I., Akter F., Hossain M. Adiponectin gene polymorphisms associated with diabetes mellitus: A descriptive review//Heliyon.- 2021.-Т.7,С8.-С. e07851 doi: 10.1016/j.heliyon.2021.e07851

Staiger H., Machicao F., Fritsche A., Haring H.-U. Pathomechanisms of Type 2 Diabetes Genes//Endocrine Reviews.-2009.-Т.30,№6.-С.557–585 doi: 10.1210/er.2009-0017

Liu S., Liu Y., Zhang Q. et al., Systematic identification of regulatory variants associated with cancer risk//Genome Biology.-2017.-Т.18.-С.194 DOI 10.1186/s13059-017-1322-z

Zheng H.-T., Peng Z.-H., Li S., He L. Loss of heterozygosity analyzed by single nucleotide polymorphism array in cancer//World Journal of Gastroenterology.-2005.-Т.11,№43.-С.6740-6744 <http://www.wjgnet.com/1007-9327/11/6740.asp>

Engle L.J., Simpson C.L., Landers J.E. Using high-throughput SNP technologies to study cancer//*Oncogene*.-2006.-T.25.-C.1594–1601
doi:10.1038/sj.onc.1209368

Li Y.-C., Korol A.B., Fahima T. et al. Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review//*Molecular Ecology*.-2002.-T.11.-C.2453—2465 DOI: 10.1046/j.1365-294x.2002.01643.x

Jasinska A., Krzyzosiak W.J. Repetitive sequences that shape the human transcriptome//*FEBS Letters*.-2004.-T.567.-C.136–141 doi:10.1016/j.febslet.2004.03.109

Ciesiolka A., Jazurek M., Draskowska K., Krzyzosiak W.J. Structural Characteristics of Simple RNA Repeats Associated with Disease and their Deleterious Protein Interactions//*Front. Cell. Neurosci*.-2017.-T.11.-C.97
doi: 10.3389/fncel.2017.00097

Panzer S., Kuhl D.P.A., Caskey C.T. Unstable Triplet Repeat Sequences: A Source of Cancer Mutations?//*STEM CELLS*.-1995.-T.13.-C.146-157

РАЗДЕЛ 3. ПОЛИГЕННЫЕ (МУЛЬТИФАКТОРНЫЕ) ЗАБОЛЕВАНИЯ. САХАРНЫЙ ДИАБЕТ 2-го ТИПА, КАК ПРИМЕР ПОЛИГЕННОГО ЗАБОЛЕВАНИЯ

ВВЕДЕНИЕ

Следствием изменения структуры генов (полиморфизма) является наследственная предрасположенность человека к тем или иным распространенным хроническим заболеваниям. Идентификация многих тысяч генов в геноме человека, выяснение молекулярных механизмов наследственных и полигенных (мультифакторных) заболеваний, доказательство генетической уникальности каждого индивида — основные достижения, которые заложили научный фундамент молекулярной медицины. Две основные особенности молекулярной медицины отличают ее от медицины традиционной (классической). Во-первых, индивидуальный характер (геном каждого человека уникален) и, во-вторых, ее профилактическая направленность (геном можно проанализировать до начала болезни). Фундамент одного из направлений молекулярной медицины — предиктивной (предсказательной) медицины — составляет анализ аллельных вариантов генов предрасположенности к полигенному заболеванию.

По имеющимся данным, величина генетического груза в популяциях Европы и России находится, примерно, на одном уровне, в диапазоне 5,5 %. При этом, около 1% приходится на долю **моногенных болезней** (патологические мутации одного гена). Ещё около 0,5% обусловлены **хромосомными абберациями**, а также 3–4% соответствуют **полигенными (мультифакторным) заболеваниям** с выраженным наследственным фактором.

Следовательно, из числа наследственных болезней, наиболее широким распространением характеризуются полигенные заболевания. Как правило, в патогенез полигенных заболеваний могут быть вовлечены не один, а несколько генов. Например, известно, что два гена (*BRCA1*, *BRCA2*) могут способствовать развитию наследственных форм рака молочной железы. Кроме того, по меньшей мере, три гена (*PS1*, *PS2*, *AAP*) принимают участие в развитии семейных форм болезни Альцгеймера, а также около 600 генов рассматриваются в качестве генов-кандидатов в диагностике диабета. Важной особенностью данной группы наследственных болезней является то, что патологическое проявление в фенотипе мутаций генов предрасположенности всегда зависит от воздействия на организм провоцирующих внешних факторов. По этой причине данную наиболее важную в медицинском отношении группу наследственных болезней называют **мультифакторными** наследственными болезнями. **Мультифакторные заболевания** имеют сложную генетическую структуру, представленную широким спектром вариантов аллелей. Данные гены различаются по происхождению (герминальные или соматические, возникшие *de novo*) и встречаются в популяции с различной частотой (от широко распространенных до редких, с частотой на уровне мутационных событий). Эти гены могут вносить разные по своей силе эффекты в риск развития заболевания у человека.

Ожирение, остеопороз, эндометриоз, инфаркт миокарда, большинство опухолей, психических и сердечно-сосудистых заболеваний входят в состав группы полигенных наследственных болезней на том основании, что они являются следствием взаимодействия многих генов с неблагоприятными факторами внешней среды. Это означает, что патогенез мультифакторных заболеваний определяется генными мутациями (полиморфизмом генов) не в одном, а сразу во многих генах. Значит, эти болезни будут фенотипически проявляться только при сочетании наследственной предрасположенности (наследственного риска развития заболевания) и соответствующих неблагоприятных факторов внешней среды.

Опираясь на выше изложенные факты можно сделать вывод о том, что лабораторный анализ (например, методом ПЦР) генетических маркеров и возможная фенотипическая корректировка их функций могут существенно снизить риск развития мультифакторного заболевания (атеросклероз, диабет, инфаркт миокарда, гипертония и прочие). Досимптоматическое выявление лиц, входящих в группу высокого риска по мультифакторной патологии и своевременная профилактика являются главными направлениями **предиктивной (предсказательной) медицины**.

Таким образом, мультифакторные заболевания являются результатом неблагоприятного сочетания генетических факторов и факторов окружающей среды. Поэтому, **мутантные гены (аллели), которые совместимы с рождением и жизнью, но при определенных неблагоприятных условиях способствуют развитию того или иного мультифакторного заболевания получили название «генов предрасположенности»**. Кроме генов, в патогенезе данной группы заболеваний принимают участие различные факторы риска, которые накапливаются с возрастом. Сердечно-сосудистые заболевания, хроническая почечная недостаточность, хронические заболевания легких, патологии печени, и раковые

заболевания являются наиболее распространенными многофакторными заболеваниями и формируют значительную нагрузку на систему здравоохранения.

Мультифакторные заболевания, также называемые полигенными или сложными наследственными заболеваниями, обусловлены многочисленными генетическими факторами и факторами окружающей среды. Важно отметить, что вовлеченные в патологический процесс аллели в подавляющем большинстве случаев не являются патологическими, а лишь повышают восприимчивость организма человека к болезням.

Подтверждение роли генов предрасположенности в патогенезе полигенных заболеваний, установление аллельного состава генов предрасположенности в каждой отдельной популяции – очень сложные задачи. Как правило, алгоритм таких исследований включает в себя три основных этапа:

- доказательство того, что заболевание является семейным,
- доказательство того, что эта семейная тенденция обусловлена генетическими факторами,
- и, наконец, выявление генов, вовлеченных в патогенез.

Картирование и отождествление генов предрасположенности к мультифакторному заболеванию проводятся, как правило, или по результатам исследования членов семей с повышенным риском развития данного заболевания. Или исследованием групп неродственных пациентов с данной патологией в сравнении с практически здоровыми испытуемыми (сопоставление групп случай/контроль). При этом, основной задачей поиска, чаще всего, является картирование однонуклеотидных полиморфизмов (SNP), расположенных, как внутри отдельного гена, так и за пределами структурного гена и наиболее тесно коррелирующих с данным заболеванием.

Данное направление исследований требует довольно существенных затрат времени и средств. Например, в масштабных исследованиях в области поиска генов предрасположенности к раку простаты были использованы образцы ДНК 1150 пациентов и 1150 контрольных образцов. Начинались исследования с анализа 500 000 SNP, количество которых в дальнейшем было сокращено до 200 SNP. Благодаря проделанной работе были идентифицированы 15 генов-кандидатов, изучение которых продолжается.

За прошедшее десятилетие, были идентифицированы генетические локусы (в основном однонуклеотидные полиморфизмы - SNP), связанные с распространенными сложными заболеваниями.

РОЛЬ ГЕННЫХ СЕТЕЙ В ПАТОГЕНЕЗЕ МУЛЬТИФАКТОРНЫХ БОЛЕЗНЕЙ

По мнению профессора Н. А. Колчанова, **генная сеть** — это группа координированно функционирующих генов, обеспечивающая формирование фенотипических признаков организма (молекулярных, биохимических, физиологических). С одной стороны, каждая из таких локальных генных сетей является дискретными и автономными сообществами генов и продуктов их экспрессии. С другой стороны **локальные генные сети** интегрированы в одну **глобальную сеть организма**. Благодаря успехам биоинформатики в настоящее время выявлен целый ряд генных сетей, обеспечивающих, например, регуляцию уровня глюкозы в организме, синтез стероидных гормонов, регуляцию функции адипоцитов (клеток жировой ткани), поддержание механизмов эритропоэза, а также генные сети окислительного стресса, регулирующие ответ органов и тканей на воздействия активных форм кислорода или азота.

Как уже отмечалось, все генные сети тесно взаимодействуют между собой, интегрированы в единую генную сеть организма. Такая интеграция автономных локальных генных сетей в организме может быть **горизонтальной или вертикальной (иерархической)**. Если речь идёт об интеграции генных сетей одного уровня, то это горизонтальная интеграция.

Примером горизонтальной интеграции может служить паракринный вариант интеграции инсулиновой и глюкагоновых генных сетей, совместно регулирующих уровень глюкозы в крови. В качестве иллюстрации вертикальной интеграции следует привести генную сеть контроля синтеза стероидных гормонов. Иерархическая природа данной вертикальной сети обусловлена тремя уровнями иерархии: гипоталамус, гипофиз и периферические эндокринные железы.

Необходимо отметить, что генные сети также находятся под внешним контролем. Действительно, генные сети являются объектом регуляции специфических механизмов-интеграторов генных сетей. В качестве таких интеграторов могут выступать специфические отдельные генные сети. Примером такого интегратора является генная сеть регуляции циркадного (суточного) ритма, которая воспринимает сигналы внешней среды и на основе полученных сигналов управляет ритмом работы огромного числа генных сетей: от регуляции обменных процессов до митотических циклов. Еще одним механизмом интеграции генных сетей могут являться различные **метаболиты** (глюкоза, формы активного кислорода и пр.) и **нейрогуморальные факторы** (гормоны, факторы транскрипции). На основе сказанного можно

сделать вывод о том, что интеграция генных сетей является важнейшим механизмом, осуществляющим адаптивный ответ всего организма и систем его жизнеобеспечения на воздействие окружающей среды.

Анализируя приведенные выше факты можно сделать вывод о том, что эффективность работы локальных генных сетей, а также возможности адаптации организма к постоянно меняющимся условиям внешней среды, будут зависеть от уникальных особенностей генов, входящих в состав данной сети. Каждый поступающий в организм сигнал вызывает ответную реакцию той или иной генной сети, которая с помощью соответствующих интеграторов передает сигнал другим генным сетям. Поломки в генных сетях, вызванные мутациями генов, или их функциональная неполноценность, обусловленная аллельными вариантами, может, если и не полностью разрушить, то существенно исказить работу всей интегральной генной сети, то есть всего организма. В патогенез каждого заболевания вовлекаются много разных функционально взаимосвязанных генов той или иной локальной генной сети. Наряду с главными генами, инициирующими болезнь, всегда имеются многочисленные второстепенные гены-модификаторы, фенотипические эффекты которых во многом определяются средовыми факторами. Выявление таких генов, установление характера функциональных нарушений на уровне локальных и интегральных генных сетей, особенностей ген-генных взаимодействий при мультифакторных заболеваниях — важная практическая задача биоинформатики.

Очевидно, что благодаря полиморфизму генов, индивидуальной уникальности аллельного состава, генные сети каждого отдельного человека функционируют по-разному. Поскольку гены, образующие основу каждой генной сети, полиморфны, в силу чего их аллельные варианты функционально различны у разных людей. В результате полиморфизма генов каждый человек имеет свой индивидуальный биохимический профиль, индивидуальную норму биохимических реакций, индивидуальные адаптивные возможности на фоне воздействия неблагоприятных факторов внешней среды. Следовательно, благодаря полиморфизму генов (индивидуальному генетическому составу) формируются индивидуальные реакции человека на внешние воздействия, инфекции, травмы, геофизические условия, пищевые продукты, на действие токсинов и лекарственных препаратов.

Всё сказанное имеет самое непосредственное отношение к патогенезу мультифакторных заболеваний. В современной медицинской литературе многократно описаны случаи семейного (наследственного) характера

предрасположенности к атеросклерозу, диабету, заболеваниям сердца, легких, почек, печени, психическим отклонениям, с высокой склонностью к онкологическим и аллергическим заболеваниям. Основа указанных наследственных болезней — гены. Поскольку фенотип (в том числе патологический) отдельного человека, есть проявление индивидуальных особенностей его генотипа, реализованного при участии локальных и интегральных генных сетей в данных конкретных условиях среды. Набор аллелей в генотипе человека индивидуален, но полиморфизм генов не

всегда является нейтральным. Иногда присутствие некоторых вариантов генов может привести к образованию белков, содержащих изменения физико-химических свойств, сопровождающихся нарушениями функциональной активности белка. Как уже отмечалось однонуклеотидные замены (SNP) в кодирующих участках гена могут оказывать влияние не только на изменение первичной структуры белка, но и третичной структуры белка, а значит на состояние его функциональной активности. В определенных условиях среды (геофизический фактор, рацион питания и т.д.) эффекты, определяемые полиморфизмом генов, могут способствовать или, противодействовать развитию различных заболеваний. Гены, аллельные варианты которых при наличии определённых факторов среды

предрасполагают к определенным заболеваниям, и получили название **генов предрасположенности**. Таким образом, гены предрасположенности — это мутантные гены (аллели), которые совместимы

с рождением и жизнью в постнатальном периоде, но при определенных неблагоприятных условиях способствуют развитию того или иного заболевания. Возможно, граница между моногенными и полигенными заболеваниями не столь значима и между этими группами наследственных болезней существует тесная взаимосвязь. Высказанное мнение базируется на том факте, что “общие” гены для менделевских и многофакторных заболеваний тесно связаны друг с другом на уровне белок-белковых взаимодействий и имеют более высокий уровень экспрессии в клетках и тканях, чем гены “только многофакторных заболеваний”. Кроме того, “общие” гены менделевских и многофакторных заболеваний имеют большое количество транскриптов и кодируют более длинные белки. При проведении ассоциативных исследований варианты в “общих” генах для менделевских и многофакторных заболеваний имеют более высокое значение рисков, чем гены “только многофакторных заболеваний”. Высказывается мнение о том, что исследования многофакторных заболеваний должны быть, прежде всего, направлены на изучение вариантов генов, вызванных мутациями в области промоторов генов менделевских болезней или в некодирующих регионах генома рядом с данными промоторами.

РОЛЬ МУЛЬТИФАКТОРНЫХ ЗАБОЛЕВАНИЙ СЕРДЦА, ПОЧЕК, ЛЁГКИХ И ПЕЧЕНИ В КАНЦЕРОГЕНЕЗЕ

Полигенное заболевание – достаточно сложное явление. Вместе с тем, наличие такого заболевания может стать причиной других патологий, например, онкопатологии. Дополнительные проблемы создаёт тот факт, что одновременное проявление двух или более заболеваний данной группы у одного пациента представляет собой различные клинические проблемы. При этом, терапевтическая коррекция одного заболевания может быть несовместимой с лечением другого заболевания. Помимо этого, следует учитывать повышение риска развития онкологических заболеваний среди пациентов с уже выявленными многофакторными заболеваниями. Говоря иначе, широкая распространенность мультифакторных заболеваний, может объяснить повышенный риск распространения онкологических заболеваний. По данным Отдела народонаселения Департамента ООН по экономическим и социальным вопросам на 2019 год, ожидается, что в 2050 году население мира старше 65 лет превысит 1,5 миллиарда человек. Следовательно, принимая к сведению тот факт, что возраст – один из основных факторов риска проявления полигенных болезней, можно предположить, что усилится манифестация мультифакторных заболеваний, в т.ч., заболеваний, способствующих развитию онкопатологий.

СВЯЗЬ МЕЖДУ ОКОЛОГИЕЙ И СЕРДЕЧНОЙ НЕДОСТАТОЧНОСТЬЮ

С одной стороны, кардиотоксичность, вызванная химиотерапией онкопатологии, все более очевидна, поскольку лечение на основе цитостатиков может привести к необратимому повреждению миокарда. С другой стороны, высказывается мнение о том, что сердечная недостаточность может быть онкогенным состоянием. Действительно, популяционные исследования показали, что распространенность онкологических заболеваний выше у пациентов с сердечной недостаточностью, в сравнении с общей популяцией. Возможно, риск развития рака у людей с сердечной недостаточностью повышен на 24–68%, по сравнению со здоровыми людьми. В частности, установлена взаимосвязь между сердечной недостаточностью и прогрессированием рака толстой кишки. Эта взаимосвязь, скорее всего, не зависит от изменений параметров гемодинамики, а определяется физиологически активными веществами, которые выделяет поврежденный миокард. Хроническое системное воспаление низкой степени интенсивности рассматривается в качестве основного патофизиологического процесса, связывающего сердечную недостаточность с многоступенчатым процессом канцерогенеза. Помимо рака толстой кишки сердечная недостаточность может способствовать развитию рака молочной железы и легких. Следует отметить, что в основе патогенеза, как сердечной недостаточности, так и рака лежат общие пути, такие как воспаление, окислительный стресс и соматические мутации. Следовательно, наличие общих, базовых механизмов может объяснить сосуществование этих двух нозологий.

СВЯЗЬ МЕЖДУ ОКОЛОГИЕЙ И ХРОНИЧЕСКОЙ ПОЧЕЧНОЙ НЕДОСТАТОЧНОСТЬЮ

В связи с риском онкологических заболеваний на фоне хронической почечной недостаточности, следует отметить, что, во-первых, риск рака наблюдается даже у пациентов с умеренными нарушениями функции почек. Во-вторых, эта тенденция наблюдалась только у мужчин. Привлекает внимание тот факт, что независимо от курения и возраста пациентов снижение скорости клубочковой фильтрации (СКФ) на 10 мл/мин/1,73 м² было связано с увеличением риска возникновения онкологического заболевания на 29%. Эта связь была ещё более значимой при оценке риска злокачественных новообразованиях легких, толстой кишки, почек и мочевыводящих путей. Можно предположить, что на фоне снижения величины СКФ в плазме крови пациентов с хронической почечной недостаточностью наблюдается более высокий уровень канцерогенных соединений и азотсодержащих веществ. Между тем, известно, что канцерогенные соединения, такие как 2-амино-6-метилдипиридо [1,2-а: 3',2'-d] имидазол (Glu-P-1) и 2-аминодипиридо [1,2-а:3',2'-d] имидазол (Glu-P-2) связываются с ДНК и вызывают мутации. При этом степень риска и локализация возникновения онкологических заболеваний на фоне прогрессирования почечной недостаточности может зависеть от тяжести основного заболевания. Действительно, среди пациентов, находящихся на диализе, риск рака мочевыводящих путей, эндокринной системы и пищеварительного тракта был значительно повышен.

В связи с изложенными фактами, некоторые авторы обращают внимание на ассоциированный с нейтрофильной желатиназой липокалин (NGAL). Это белок с массой 25 кДа, который секретируется в основном иммунными клетками, такими как нейтрофилы, макрофаги и дендритные клетки. Его выработка стимулируется в ответ на воспаление. Белок NGAL известен в основном, как биомаркер

острой почечной недостаточности и выделяется в результате повреждения канальцев нефрона, а также во время процессов регенерации почек. Уровень NGAL также повышен у пациентов с хроническим заболеванием почек. Присутствие NGAL усиливает воспалительные реакции, а его уровень повышается при артериальной гипертензии, ожирении, диабете и метаболических осложнениях, таких как резистентность к инсулину, а также участвует в канцерогенезе.

Что касается механизмов индукции канцерогенеза, то одним из основных является процесс хронического воспаления. С одной стороны, очаг хронического воспаления является источником активных форм кислорода, которые способствуют возникновению различных мутаций – базовому механизму канцерогенеза. Кроме того, продукция провоспалительных цитокинов, таких, как интерлейкин-1 β (IL-1 β), интерлейкин-6 (IL-6) и фактор некроза опухоли-альфа позволяет злокачественным клеткам ускользать от иммунного надзора организма, стимулируя ангиогенез, рост опухоли и инвазивность.

РАЗВИТИЕ РАКА ЛЕГКИХ ПРИ ХРОНИЧЕСКОЙ ОБСТРУКТИВНОЙ БОЛЕЗНИ ЛЕГКИХ (ХОБЛ)

Хроническая обструктивная болезнь легких (ХОБЛ) и рак легких являются двумя ведущими проблемами общественного здравоохранения и причинами заболеваемости и смертности. Оба заболевания имеют общие факторы риска, такие, например, как воздействие курения и генетическая предрасположенность.

В настоящее время установлено, что некоторые семейства генов протеиназ, детоксицирующих ферментов и воспалительных цитокинов могут играть важную роль в патогенезе ХОБЛ. Также выявлены некоторые гены-кандидаты, такие как гликофорин А (*GYPA*) и белок, взаимодействующий с Hedgehog (*HHIP*) могут быть связаны с восприимчивостью, как к ХОБЛ, так и к раку легких. ХОБЛ характеризуется хроническим воспалением легких, поэтому предполагают, что воспаление является связующим звеном между ХОБЛ и раком легких, но точный лежащий в основе механизм неизвестен. Медиаторы воспаления способствуют развитию рака, вызванного воспалением, стимулируя клеточную пролиферацию, устойчивость к апоптозу, инвазию и метастазирование.

Еще одним важным механизмом патогенеза ХОБЛ является окислительный стресс, который является причиной клеточной пролиферации при раке легких. Дисфункция митохондрий в паренхиме легких может влиять на патогенез ХОБЛ, включая дисбаланс окислительного стресса, который может усиливать хроническое воспаление и способствовать канцерогенезу.

ПРОГРЕССИРОВАНИЕ ЖИРОВОЙ БОЛЕЗНИ ПЕЧЕНИ, СВЯЗАННОЙ С МЕТАБОЛИЧЕСКОЙ ДИСФУНКЦИЕЙ И ОНКОЛОГИЧЕСКИЕ ЗАБОЛЕВАНИЯ

Жировая болезнь печени, связанная с метаболической дисфункцией (Metabolic dysfunction-associated fatty liver disease (MAFLD) - неалкогольная жировая болезнь печени — частный случай стеатогепатоза, возникающий у людей, не злоупотребляющих алкоголем. Данная патология обусловлена инсулинорезистентностью и метаболическим синдромом.

Метаболические нарушения, наблюдаемые у пациентов с MAFLD, являются наиболее важными факторами повышения риска рака, в первую очередь онкопатологии печени - гепатоцеллюлярной карциномой (ГЦК). При этом, наличие MAFLD у пациентов с хроническим гепатитом В дополнительно увеличивается риск развития ГЦК.

В последние годы MAFLD стал самой быстрорастущей причиной ГЦК в США. Ожирение и сахарный диабет 2-го типа играют важную роль в развитии и прогрессировании ГЦК при MAFLD. Диабет может оказывать влияние на канцерогенез посредством нескольких механизмов, включая гиперинсулинемию, гипергликемию, окислительный стресс и хроническое воспаление. В свою очередь, хроническое воспаление стимулирует секрецию IL-6 и TNF- α , связанных с онкогенезом. Помимо этого, стеатоз может привести к митохондриальной дисфункции и стрессу эндоплазматического ретикулума (ЭР), что приводит к увеличению продукции активных форм кислорода. В свою очередь, окислительный стресс приводит к повреждению ДНК и активации онкогенных сигнальных путей клетки, таких как NF- κ B.

Необходимо подчеркнуть, что MAFLD не только ассоциируется с повышенным риском ГЦК, но и с несколькими внепеченочными видами рака, включая колоректальный рак, рак почки, рак щитовидной железы и рак молочной железы. Привлекает внимание тот факт, что риск рака почки и щитовидной железы оставался значительным даже после коррекции метаболических процессов. Связь между MAFLD и колоректальным раком привлекает все большее внимание в последние несколько лет. В ходе

популяционных исследований установлено, что MAFLD является важным фактором риска колоректального рака.

ПОЛИГЕННАЯ ПРИРОДА МУЛЬТИФАКТОРНЫХ ЗАБОЛЕВАНИЙ НА ПРИМЕРЕ САХАРНОГО ДИАБЕТА 2-го ТИПА

Сахарный диабет относится к группе метаболических заболеваний, которые характеризуются хроническим повышением уровня глюкозы в плазме крови. По разным оценкам, более 90% всех случаев диабета относятся к диабету 2 типа (СД2), который ранее был известен как «диабет взрослых» или «инсулиннезависимый диабет». Заболеваемость диабетом 2-го типа достигла масштабов глобальной эпидемии. В настоящее время диабетом страдают около 537 миллионов взрослых (20–79 лет), и общее число людей с данным диагнозом постоянно растет. По прогнозам, заболеваемость СД2 увеличится до 643 миллионов к 2030 году и 783 миллион к 2045 году. Патогенетические механизмы СД2 включают резистентность к инсулину, нарушение секреции инсулина и повышенную выработку глюкозы печенью. Осложнения могут характеризоваться полиорганный недостаточностью, включая нефропатию, нейропатию, ретинопатию, пародонтоз и ускоренное развитие сердечно-сосудистых заболеваний. В 2010 году в Соединенных Штатах диабет 2-го типа был впервые диагностирован у 1,9 миллиона человек в возрасте от 20 лет и старше. СД2 становится все более распространенным во многих регионах мира. Распространенность СД2 среди подростков и юношей в возрасте от 10 лет и старше составляет 8,5 человек на 100 000 населения, тогда как среди детей младше 10 лет распространенность составляет 0,4 на 100 000 человек. В США 26,9%, всех людей в возрасте 65 лет и старше поражены данным заболеванием. Как показывают прогнозы, такие факторы, как старение населения и рост доли населения с избыточной массой тела еще больше увеличат распространенность СД2. Тогда, как изменение образа жизни (снижение массы тела, оптимизация питания и увеличение физической активности) может помочь предотвратить или существенно отсрочить развитие диабета.

Вместе с тем, накапливаются сведения о том, что в патогенезе СД2 важную роль играет наследственный фактор. В частности, широко признана отчетливая связь риска заболевания и принадлежности к определенной этнической группе. Кроме того, зачастую заболевание носит семейный характер. Подчеркнем, что семейный анамнез диабета является основным фактором риска заболевания и часто включается в инструменты количественной оценки риска диабета. В возрасте 80 лет риск заболевания, примерно, в 3,5 раза выше, если близким родственникам диагностирован СД2, по сравнению с лицами, родственники которых не болели этим заболеванием.

СД2 является полигенным заболеванием с наследуемым риском от 30 до 70%. СД2 является мультифакторным (полигенным) заболеванием. На проявление симптомов СД2 оказывают влияние, как генетические, так и экологические факторы. Степень участия генетических факторов и факторов среды может варьировать у разных людей. Но, заболевание развивается, когда сумма эффектов обоих факторов превышает определенный порог. Роль генетического фактора не вызывает сомнений. По данным близнецового и генеалогического методов медицинской генетики, наследуемость осложнений СД2 составляет около 40%. Кроме того, частота встречаемости полиморфизмов генов восприимчивости к заболеванию различается между группами пациентов и контрольными группами. Установлены популяционные особенности наличия генов предрасположенности к СД2. В то время как у европейцев развитию диабета 2 типа предшествует избыточная масса тела, у жителей Восточной Азии диабет развивается при более низком весе, чем у европейцев, из-за разного уровня секреции инсулина.

САХАРНЫЙ ДИАБЕТ 2-го ТИПА, КАК ПРИМЕР ПОЛИГЕННОГО ЗАБОЛЕВАНИЯ. ГЕНЫ ПРЕДРАСПОЛОЖЕННОСТИ К САХАРНОМУ ДИАБЕТУ 2-го ТИПА

Основным инструментом в изучении патологических вариантов генов предрасположенности к СД2 является секвенирование нового поколения. Большинство генов предрасположенности к СД2 связаны с регуляцией функционального состояния β -клеток поджелудочной железы. Генетические полиморфизмы, которые оказывают влияние на важные белки, участвующие в метаболизме глюкозы и секреции инсулина, также могут влиять на предрасположенность к СД2. Геномные полигенные оценки для СД2 могут предсказать риск возникновения диабета. Некоторые авторы высказывают мнение о том, выявление генов предрасположенности развития СД2, с точки профилактического мероприятия, является слишком громоздким, дорогостоящим и избыточным. Дополнительной серьезной проблемой является подбор маркерных генов для данной конкретной популяции. По их мнению, наиболее целесообразным и вполне эффективным является скрининг, проводимый на базе показателей уровня глюкозы в плазме крови, гликозилированного гемоглобина, величины кровяного давления и индекса массы тела. Тем не

менее, объем исследований генетических аспектов предрасположенности к СД2 неуклонно увеличивается во всем мире. Причина данной тенденции обусловлена тем, что масштабы проблемы носят глобальный характер, количество пациентов с СД2 стремительно растет, также, как и размер наносимого ущерба и необходимо принятие неотложных мер, направленных на разработку эффективных методов ранней диагностики и лечения заболевания. Необходимо подчеркнуть, что основным направлением поиска генов предрасположенности к СД2 являются генные мутации, преимущественно SNP.

Одним из первых общепринятых генетических маркеров СД2 является ген *TCF7L2*, локус которого располагается на длинном плече 10-й хромосомы. В настоящее время среди европейских и азиатских популяций, выявлено более 600 локусов, восприимчивых к СД2. При этом, ген *TCF7L2* имел одинаковую диагностическую ценность в европейских популяциях и популяциях Восточной Азии. Также и ген *KCNQ1* — локус восприимчивости к болезням с относительно высоким отношением рисков СД2 - 1,4 раза, выявленный у японской популяции, но он высоко воспроизводим и в европейских популяциях. Напротив, гены *DGKD* и *ZNF257* обладают диагностической ценностью только в популяции Восточной Азии.

В современной литературе уделяется существенное внимание вопросу частоты встречаемости генов предрасположенности к СД2 в различных популяциях. В качестве иллюстрации можно привести сведения о динамике уровня ожирения и СД2 в странах Ближнего Востока. В связи с этим привлекает внимание несколько закономерностей. Во-первых, наблюдается стремительный рост количества людей с избыточной массой тела – начиная с пороговых величин (менее 5% населения) в середине 1970 гг, до более, чем 20% в настоящее время. В Кувейте доля людей с избыточной массой тела достигает почти 40%. Во-вторых, в большинстве арабских стран уровень браков между двоюродными братьями и сестрами может достигать 30%, что делает архитектуру арабского генома уникальной с точки зрения их восприимчивости к различным заболеваниям, включая как менделевское наследование, так и полигенные заболевания. В-третьих, из, примерно, 100 вариантов генов, ассоциированных с избыточной массой тела, два были описаны как уникальные для арабов, поскольку они ранее не были зарегистрированы в других популяциях планеты. В то же время, 19 аллелей были зарегистрированы как отчетливые маркеры избыточной массы тела у арабов, но не обладали патогенными эффектами в неарабских популяциях.

Привлекает внимание и тот факт, что существует и гендерная специфика генетических маркеров. Например, ген *ALDH2* был тесно связан с риском развития диабета 2 типа у мужчин, но не у женщин.

Ниже мы предлагаем к рассмотрению характеристики наиболее актуальных SNP, ассоциированных с развитием СД2 у человека.

ГЕН *TCF7L2* (rs7903146). Данный ген кодирует фактор транскрипции, который играет ключевую роль в сигнальном пути Wnt. Белок участвует в гомеостазе глюкозы в крови. Генетические варианты этого гена связаны с повышенным риском диабета 2 типа. Для этого гена было обнаружено несколько вариантов транскриптов, кодирующих несколько различных изоформ. Нарушение регуляции сигнального пути Wnt играет ключевую роль в развитии резистентности к инсулину. Генотип rs7903146 является хорошо изученным маркером риска СД2 в европейских и азиатских популяциях. Пациенты с СД2 с большей вероятностью являются носителями генетического варианта (rs7903146) этого гена в разных популяциях.

ГЕН *KLF14* (rs4731702). Данный неинтронный ген кодирует белок, относящийся к семейству транскрипционных факторов, подобных Круппелю. Кодируемый белок функционирует, как корепрессор транскрипции и индуцируется трансформирующим фактором роста-бета (TGF-бета) для подавления экспрессии гена рецептора II TGF-бета. Данный ген регулирует генные сети, которые участвуют в метаболизме липидов. Высокие уровни инсулина натощак и резистентность к инсулину ассоциируются с аллелем риска rs4731702. Кроме того, он может оказывать влияние на экспрессию генов, определяющих индекс массы тела.

ГЕН *KCNQ1* (rs2237895). Данный ген кодирует альфа-субъединицу потенциалзависимого калиевого канала Kv7.1, является членом суперсемейства каналов Kv и расположен на хромосоме 11p15.5. Белок, кодируемый геном *KCNQ1*, является порообразующей альфа-субъединицей калиевых каналов KCNQ1/KCNE1, KCNQ1/KCNE2 и KCNQ1/KCNE3. Ген *KCNQ1* экспрессируется не только в миокарде, но и в жировой ткани, поджелудочной железе и мозге. В островках поджелудочной железы и играет важную роль в гомеостазе глюкозы, поскольку он функционирует как регулятор секреции инсулина. Распологающийся в интроне rs2237895 *KCNQ1* связан с пониженным риском абдоминального ожирения у пациентов с СД2.

ГЕН *FTO*. Этот ген кодирует ядерный белок суперсемейства негемового железа и 2-оксоглутарат-зависимой оксигеназы, родственного AlkB. Как и другие негемовые железо-содержащие ферменты этого семейства регулирует процессы окислительного деметилирования ДНК. Высокие уровни экспрессии гена наблюдаются в нервной системе, надпочечниках и миокарде. Выявлена высокая степень связи с индексом массы тела, риском ожирения и диабетом 2 типа. Показало, что rs1558902 *FTO* тесно коррелировал с заболеваемостью СД2 у людей, причём, высокие значения корреляции оставались значительными даже после корректировки на текущий возраст и индекс массы тела.

Ген *HNF4A* является членом суперсемейства рецепторов стероидных гормонов, которое в основном экспрессируется в почках, печени, поджелудочной железе (включая β -клетки) и тонком кишечнике и влияет на метаболизм и транспорт липидов. Он также принимает участие в регуляции функции печени и дифференцировке гепатоцитов. *HNF4A* необходим для метаболизма глюкозы, а также экспрессии и секреции инсулина, в то время как в печени *HNF4A* необходим для печеночного глюконеогенеза. В настоящее время известно, что некодирующие варианты гена *HNF4A* rs6017317 и rs4812829, а также кодирующий миссенс-вариант rs1800961 являются факторами риска развития СД2. Патологические варианты гена *HNF4A* могут участвовать в нарушениях секреции инсулина, как это наблюдается у пациентов с СД2 и MODY1. Форма диабета MODY1 - это диабет зрелого возраста у молодых. MODY1 характеризуется ранним возрастом начала (обычно в подростковом или детском возрасте), доминантным наследованием и дефектами функции β -клеток. MODY, возникающий в результате мутаций в гене фактора транскрипции *HNF4A*.

ГЕН *IGF2BP2* кодирует белок, который связывает 5' UTR мРНК инсулиноподобного фактора роста 2 (*IGF2*) и регулирует ее трансляцию. Ген характеризуется высокими показателями экспрессии в инсулярном аппарате поджелудочной железы. В жировой ткани и поджелудочной железе *IGF2BP2* может снижать экспрессию *IGF2*, который играет решающую роль в контроле гистогенеза поджелудочной железы и адипогенеза. Ген играет важную роль в регуляции обменных процессов, а патологические варианты гена связаны с восприимчивостью к диабету. Полиморфизмы гена *IGF2BP2* отражаются на функциональном состоянии β -клеток поджелудочной железы. Некоторые формы гена тесно связаны с избыточным весом и ожирением. Некоторые формы гена *IGF2BP2* могут повышать риск развития СД2 через изменения в жировой ткани или нарушения функции β -клеток, которые приводят к снижению секреции инсулина. В частности, наблюдаются более высокие уровни глюкозы в плазме натощак, общего холестерина и сывороточного инсулина у пациентов с СД2, которые были носителями аллеля C rs1470579. Предполагает возможную роль *IGF2BP2* в патогенезе резистентности к инсулину.

ГЕНЫ *CDKN2A* и *CDKN2B* (rs10811661) Гены *CDKN2A* и *CDKN2B* являются генами-супрессорами опухолей, участвующими в апоптозе клеток, опухолеобразовании и регуляции пролиферации. Ген *CDKN2A* кодирует ингибитор p16 циклинзависимой киназы p16INK4A с максимально высокой экспрессией в жировой ткани и напочечниках. Ген *CDKN2B* кодирует p15INK4B. Ген *CDKN2B* расположен рядом с геном-супрессором опухолей *CDKN2A* в области, которая часто мутирует и подвержена делециям в самых разных опухолях. Этот ген кодирует ингибитор циклин-зависимой киназы p15INK4B, который образует комплекс с CDK4 или CDK6 и предотвращает активацию киназ CDK. Таким образом, кодируемый белок функционирует как регулятор роста клеток, который контролирует прогрессию клеточного цикла в фазе G1. Полиморфизмы в генах *CDKN2A/B* оказывают влияние на состояние белкового обмена, который определяет состояние β -клеток, секреции ими инсулина и пролиферацию данной популяции клеток. Результаты популяционных исследований в Азии и Европе подтвердили, что полиморфизм генов *CDKN2A/B* тесно связан с риском развития СД2. Кроме этого, при ожирении макрофаги проникают в жировую ткань и приводят к развитию хронического слабовыраженного воспаления. Эти макрофаги жировой ткани (МЖТ) стимулируют секрецию провоспалительных цитокинов и в дальнейшем способствуют резистентности к инсулину. Известно, что SNP rs10811661, который связан с экспрессией длинной некодирующей РНК, известной как антисмысловая некодирующая РНК в локусе INK4 (ANRIL или антисмысловой РНК *CDKN2B* 1 (*CDKN2B*-AS1)) тесно связан с риском развития диабета у человека. Необходимо также отметить, что rs10811661 *CDKN2A/B* демонстрирует значимую корреляцию с повышенным риском рака молочной железы, и может быть потенциально ценным для применения в качестве прогностического фактора для индивидуальной медицинской помощи.

ГЕН *KCNJ11* (rs5219; p.Glu23Lys) Белок, кодируемый этим геном, является интегральным мембранным белком Kir6.2 с функцией калиевого канала. Кодируемый белок контролируется G-белками и отвечает за проникновение калия в клетку. Повышенные уровни глюкозы стимулируют открытие канала Kir6.2 и

проникновение ионов K^+ в клетку. Повышение уровня ионов K^+ в клетке деполяризует клеточную мембрану и индуцирует каналы Ca^{2+} для увеличения уровня свободного внутриклеточного Ca^{2+} . Ионы Ca^{2+} запускают механизмы секреции инсулина, высвобождая гормон из секреторных гранул. Полиморфизм гена *KCNJ11* может приводить к снижению выработки инсулина из-за снижения или отсутствия экспрессии белка Kir6.2. Поэтому мутации в данном гене являются причиной различных форм диабета, как моногенной, так и полигенной природы: семейной персистирующей гиперинсулинемической гипогликемии младенцев (PHHI), ауtosомно-рецессивного заболевания, характеризующегося нерегулируемой секрецией инсулина. Дефекты в этом гене также могут способствовать ауtosомно-доминантному инсулиннезависимому сахарному диабету II типа (NIDDM), транзиторному неонатальному сахарному диабету 3 типа (TNDM3) и постоянному неонатальному сахарному диабету (PNDM). В частности, аллель гена *KCNJ11* rs5219 может снижать чувствительность канала к АТФ. Установлена тесная связь между полиморфизмами rs5219 и восприимчивостью к СД2, как в восточноазиатских, так и в европейских популяциях. При наличии в генотипе человека rs5219 первоначально чувствительность к инсулину повышается, тем самым поддерживая нормальную толерантность к глюкозе. Однако, со временем, по мере дальнейшего снижения секреции инсулина или развития инсулинорезистентности, уровень глюкозы повышается, что приводит к сахарному диабету 2 типа.

ГЕН *CDKAL1* (rs7756992) кодирует CDK5, регуляторный субъединичный ассоциированный белок-1. Этот регуляторный субъединичный ассоциированный белок CDK5 демонстрирует высокие уровни экспрессии в нервной ткани и бета-клетках поджелудочной железы. Данный белок участвует в глюкозозависимой регуляции секреции инсулина в инсулярном аппарате поджелудочной железы. Ранее было показано, что rs7756992 связан с риском нарушения секреции инсулина. Полиморфизм rs775480 расположен в интроне 5 гена *CDKAL1*. Данный SNP связан со сниженной чувствительностью к глюкозе и секрецией инсулина в β -клетках. Более того, полиморфизм rs10946398 гена *CDKAL1* был предложен в качестве маркера нарушенной секреции инсулина, поскольку генотипы CC/CA (присутствие аллели C) способствуют восприимчивости к СД2 у людей с ожирением. В настоящее время установлено, что аллели rs4712524, rs10946398 и rs775480 в *CDKAL1* являются подтвержденными факторами риска СД2.

ГЕН *PPAR γ 2* (rs1801282) PPAR γ 2 (рецептор, активируемый пролифератором пероксисом-гамма 2) — это активируемый лигандом фактор транскрипции суперсемейства ядерных гормональных рецепторов. Ген *PPAR γ 2* играет важную роль в гомеостазе глюкозы, липидном обмене, ожирении, чувствительности к инсулину, в контроле экспрессии различных генов, специфичных для адипоцитов и непосредственно вовлечен в патогенез сахарного диабета 2 типа. Одним из первых был установлен патологический вариант гена rs1801282, связанный с ожирением и развитием СД2. На основании ряда исследований было установлено, что аллель Pro12 PPAR γ был связан с 1,25-кратным повышением риска СД2. Кроме того, показало, что полиморфизм Pro12Ala был связан с риском СД2 в многоэтническом исследовании.

ГЕН *SLC30A8* (rs13266634) Белок ZNT8, кодируемый этим геном, является переносчиком ионов цинка, участвующим в накоплении цинка во внутриклеточных везикулах. Это протон-связанный антипортёр ионов цинка, опосредующий поступление цинка в просвет секреторных гранул бета-клеток поджелудочной железы, тем самым регулируя секрецию инсулина. Этот ген экспрессируется на высоком уровне только в поджелудочной железе, особенно в островках Лангерганса. Кодируемый белок локализуется с инсулином в гранулах секреторного пути клеток, секретирующих инсулин. Низкая экспрессия белка ZnT8 приводит к снижению выработки инсулина β -клетками. Полиморфизм *SLC30A8* rs13266634 был тесно связан с риском развития СД2. Кроме того, rs13266634 *SLC30A8* был связан с возрастом, как фактором риска развития СД2.

ГЕН *POMC* Этот ген кодирует препропротеин проопиомеланокортин, который подвергается обширному, тканеспецифичному посттрансляционному процессингу. В препропротеине есть несколько потенциальных участков расщепления, и в зависимости от типа ткани и доступных конвектаз процессинг может дать до десяти биологически активных пептидов, участвующих в различных клеточных функциях. Кодируемый белок синтезируется в основном в кортикотрофных клетках передней доли гипофиза, где используются четыре участка расщепления; аденокортикотропин, необходимый регуляции стероидогенеза в надпочечниках, и липотропин-бета. В других тканях, включая гипоталамус, плаценту и эпителий, могут использоваться все участки расщепления, что приводит к образованию эндорфинов, гормона стимуляции меланоцитов и иммуномодуляции. Они включают несколько отдельных меланотропинов, липотропинов и эндорфинов. У пациентов с СД2 и ожирением, связанным с мутациями

в генах *POMC*, наблюдалась миссенс-мутация в сигнальном пептиде. Эта мутация привела к гетерозиготной замене аргинина на глицин в A15G-*POMC* (кодон 15), что подавляло выработку и секрецию белка *POMC*. У людей дефицит *POMC* может вызывать резистентность к инсулину и гиперинсулинемию.

ГЕН *ADIPOQ* кодирует регуляторный пептид жировой ткани адипонектин, который является адипоцитокинем. Жировая ткань играет важную роль в развитии сахарного диабета и ожирения. Адипонектин, основной секреторный белок адипоцитов в плазме человека, функционирует как регулятор энергии и участвует в формировании толерантности к глюкозе. Адипонектин, вырабатываемый жировой тканью, существует в различных формах, таких как мономеры, тримеры и гексамеры. Он действует через свои рецепторы, *adipoR1* и *adipoR2*, влияя на множественные метаболические процессы. Активация *adipoR1* в основном способствует митохондриальному биогенезу и усвоению глюкозы через пути AMPK и p38 MAPK. Активация *adipoR2* участвует в окислении жирных кислот через путь PPAR α и обеспечивает цитопротекторное и противовоспалительное действие посредством модуляции NF- κ B. В совокупности эти пути способствуют регуляции энергетического метаболизма и воспалительных реакций в различных тканях, включая мышцы, печень, почки, мозг и кости. Ген *ADIPOQ* расположен на хромосоме человека 3q27 и, как сообщается, является локусом восприимчивости к СД2. Полиморфизмы гена *ADIPOQ* в rs1501299, rs182052 и rs7627128 демонстрируют значительную связь с СД2. Также показано, что полиморфизм ген *ADIPOQ*, в частности аллель rs2241766 способствует увеличению риска развития СД2 и принимает участие в патогенезе диабетической нефропатии.

ГЕНЫ ФЕРМЕНТОВ АНТИОКСИДАНТНОЙ ЗАЩИТЫ Нарушение баланса антиоксидантов и активных форм кислорода (ROS) приводит к усилению окислительного стресса, что может привести к диабету. Образование и накопление ROS в β -клетках может вызвать дисфункцию β -клеток, дефекты в выработке инсулина и нарушение функции, что приводит к диабету. Однако воздействие окислительного стресса может быть уменьшено или изменено ферментативными антиоксидантами, включая каталазу (CAT), глутатион-S-трансферазу (GST), глутатионпероксидазу (GPx), супероксиддисмутазу (SOD), синтазу оксида азота и никотинамидадениндинуклеотидфосфатоксидазу. Люди с полиморфизмом, влияющим на генетическую регуляцию этих шести ферментов, подвергаются повышенному риску развития СД2. При этом, риск развития СД2 увеличивается по мере увеличения вариации генов, регулирующих антиоксидантные ферменты.

Помимо перечисленных выше генов к генам предрасположенности к сахарному диабету 2-го типа можно отнести гены: ген *ANK1/NKX6-3* (rs33981001, rs62508166), ген *HHEX* (rs1111875), ген *WFS1* (rs10010131), ген *MTNR1B* (rs1387153) и ген *DGKB (TMEM195)* (rs10244051).

РЕКОМЕНДОВАННАЯ ЛИТЕРАТУРА К РАЗДЕЛУ 3. ПОЛИГЕННЫЕ ЗАБОЛЕВАНИЯ

Генетический паспорт — основа индивидуальной и предиктивной медицины / Под ред. В. С. Баранова. — СПб.: Изд-во Н-Л, 2009. — 528 с.

Shi C. и соавт. Multifactorial Diseases of the Heart, Kidneys, Lungs, and Liver and Incident Cancer: Epidemiology and Shared Mechanisms//Cancers (Basel). 2023 Feb; 15(3): 729
doi: 10.3390/cancers15030729

Zamorano J.L., Lancellotti P., Rodriguez Munoz D. et al. 2016 ESC Position Paper on cancer treatments and cardiovascular toxicity developed under the auspices of the ESC Committee for Practice Guidelines: The Task Force for cancer treatments and cardiovascular toxicity of the European Society of Cardiology (ESC) Eur. Heart J. 2016;37:2768–2801. doi: 10.1093/eurheartj/ehw211

Hasin T., Gerber Y., McNallan S.M. et al. Patients with heart failure have an increased risk of incident cancer. J. Am. Coll. Cardiol. 2013;62:881–886. doi: 10.1016/j.jacc.2013.04.088

Meijers W.C., Maglione M., Bakker S.J.L. et al. Heart Failure Stimulates Tumor Growth by Circulating Factors. Circulation. 2018;138:678–691. doi: 10.1161/CIRCULATIONAHA.117.030816

Wong G., Hayen A., Chapman J.R. et al. Association of CKD and cancer risk in older people. J. Am. Soc. Nephrol. 2009;20:1341–1350. doi: 10.1681/ASN.2008090998

Wong G., Staplin N., Emberson J. et al. Chronic kidney disease and the risk of cancer: An individual patient data meta-analysis of 32,057 participants from six prospective studies. *BMC Cancer*. 2016;16:488. doi: [10.1186/s12885-016-2532-6](https://doi.org/10.1186/s12885-016-2532-6)

Locatelli F., Canaud B., Eckardt K.U. et al. Oxidative stress in end-stage renal disease: An emerging threat to patient outcome. *Nephrol. Dial. Transpl.* 2003;18:1272–1280. doi: [10.1093/ndt/gfg074](https://doi.org/10.1093/ndt/gfg074)

Igo R. P. Jr., Kinzy T.G., Bailey J.N. Genetic Risk Scores. *Curr Protoc Hum Genet*. 2019; 104(1): e95. doi: [10.1002/cphg.95](https://doi.org/10.1002/cphg.95)

Palomaki G.E., Melillo S., Marrone M., Douglas M.P. Use of genomic panels to determine risk of developing type 2 diabetes in the general population: a targeted evidence-based review. *Genet Med*. 2013; 15(8): 600–611. doi: [10.1038/gim.2013.8](https://doi.org/10.1038/gim.2013.8)

Akhlaghipour I., Bina A.R., Mogharrabi M.R. Single-nucleotide polymorphisms as important risk factors of diabetes among Middle East population. *Hum Genomics*. 2022; 16: 11. doi: [10.1186/s40246-022-00383-2](https://doi.org/10.1186/s40246-022-00383-2)

Shojima N., Yamauchi T. Progress in genetics of type 2 diabetes and diabetic complications. *J Diabetes Investig*. 2023; 14(4): 503–515. doi: [10.1111/jdi.13970](https://doi.org/10.1111/jdi.13970)

Witka B.Z., Oktaviani D.J., Marcellino M. и соавт. Type 2 Diabetes-Associated Genetic Polymorphisms as Potential Disease Predictors. *Diabetes Metab Syndr Obes*. 2019; 12: 2689–2706. doi: [10.2147/DMSO.S230061](https://doi.org/10.2147/DMSO.S230061)

Younes S., Ibrahim A., Al-Jurf R., Zayed H. Genetic polymorphisms associated with obesity in the Arab world: a systematic review. *Int J Obes (Lond)*. 2021; 45(9): 1899–1913. doi: [10.1038/s41366-021-00867-6](https://doi.org/10.1038/s41366-021-00867-6)

Ausoni S., Azzarello G. Development of Cancer in Patients With Heart Failure: How Systemic Inflammation Can Lay the Groundwork. *Front Cardiovasc Med*. 2020; 7: 598384. doi: [10.3389/fcvm.2020.598384](https://doi.org/10.3389/fcvm.2020.598384)

Romejko K., Markowska M., Niemczyk S. The Review of Current Knowledge on Neutrophil Gelatinase-Associated Lipocalin (NGAL). *Int J Mol Sci*. 2023; 24(13): 10470. doi: [10.3390/ijms241310470](https://doi.org/10.3390/ijms241310470)

del Bosque-Plata L., Martínez-Martínez E., Espinoza-Camacho M.Á., Gragnoli C. The Role of *TCF7L2* in Type 2 Diabetes. *Diabetes*. 2021; 70(6): 1220–1228. doi: [10.2337/db20-0573](https://doi.org/10.2337/db20-0573)

Lorenz K., Voight B.F. Dissecting an adiposity locus with an arsenal of genomics. *Genome Biol*. 2018; 19: 74. Published online 2018 Jun 7. doi: [10.1186/s13059-018-1460-y](https://doi.org/10.1186/s13059-018-1460-y)

Akash M.S.H., Rasheed S., Rehman K. и соавт. Biochemical Activation and Regulatory Functions of Trans-Regulatory KLF14 and Its Association with Genetic Polymorphisms. *Metabolites*. 2023; 13(2): 199. doi: [10.3390/metabo13020199](https://doi.org/10.3390/metabo13020199)

Yu X., Liao M., Zeng Y. и соавт. Associations of *KCNQ1* Polymorphisms with the Risk of Type 2 Diabetes Mellitus: An Updated Meta-Analysis with Trial Sequential Analysis. *J Diabetes Res*. 2020; 2020: 7145139. doi: [10.1155/2020/7145139](https://doi.org/10.1155/2020/7145139)

Popović A.-M., Huđek Turković A., Žuna K. и соавт. *FTO* Gene Polymorphisms at the Crossroads of Metabolic Pathways of Obesity and Epigenetic Influences. *Food Technol Biotechnol*. 2023; 61(1): 14–26. doi: [10.17113/ftb.61.01.23.7594](https://doi.org/10.17113/ftb.61.01.23.7594)

Radi S.H. и соавт. HNF4α isoforms: the fraternal twin master regulators of liver function. *Front Endocrinol (Lausanne)*. 2023; 14: 1226173.

doi: [10.3389/fendo.2023.1226173](https://doi.org/10.3389/fendo.2023.1226173)

Zhang P., Wu W., Ma C. и соавт. RNA-Binding Proteins in the Regulation of Adipogenesis and Adipose Function. *Cells*. 2022; 11(15): 2357.
doi: [10.3390/cells11152357](https://doi.org/10.3390/cells11152357)

Davidson H.W., Wenzlau J.M., O'Brien R.M. Zinc transporter 8 (znt8) and beta cell function. *Trends Endocrinol Metab*. 2014; 25(8): 415–424.
doi: [10.1016/j.tem.2014.03.008](https://doi.org/10.1016/j.tem.2014.03.008)

Yi B., Huang G., Zhou Z. Different role of zinc transporter 8 between type 1 diabetes mellitus and type 2 diabetes mellitus. *J Diabetes Investig*. 2016; 7(4): 459–465.
doi: [10.1111/jdi.12441](https://doi.org/10.1111/jdi.12441)

Hamidi Y. и соавт. A Meta-analysis of ADIPOQ rs2241766 polymorphism association with type 2 diabetes. *J Diabetes Metab Disord*. 2022; 21(2): 1895–1901.
doi: [10.1007/s40200-022-01086-0](https://doi.org/10.1007/s40200-022-01086-0)

Baldelli S. и соавт. The Role of Adipose Tissue and Nutrition in the Regulation of Adiponectin. *Nutrients*. 2024; 16(15): 2436.
doi: [10.3390/nu16152436](https://doi.org/10.3390/nu16152436)

РАЗДЕЛ 4. ПРЕЭКЛАМПСИЯ, КАК ПОЛИГЕННОЕ (МУЛЬТИФАКТОРНОЕ) ЗАБОЛЕВАНИЕ

ВВЕДЕНИЕ

Преэклампсия (ПЭ) является наиболее актуальным заболеванием у женщин во время беременности во всём мире. ПЭ специфическое для беременности заболевание, характеризующееся развитием гипертензии *de novo* и протеинурии во второй половине беременности. Данное заболевание является серьезной акушерской проблемой, которая вносит значительный вклад в материнскую и перинатальную заболеваемость и смертность во всем мире. ПЭ описывается как плацентарная патология, поскольку после родов у большинства женщин, как правило, восстанавливается нормальное артериальное давление. Риск развития ПЭ в различных популяциях может варьировать от 3 до 14—21%. Глобальная тенденция указывает на повышение риска развития ПЭ во время беременности у женщин во всем мире. В результате преэклампсии, повышается риск преждевременных родов, перинатальной смертности и нарушений развития нервной системы, а также сердечно-сосудистых и метаболических заболеваний в более позднем возрасте. В России ПЭ встречается с частотой до 11—16% случаев и занимает третье место среди причин материнской смертности. В настоящее время не существует универсальной международной системы классификации или определения ПЭ. Однако, гипертензивные расстройства беременности принято подразделять на четыре категории:

-Гестационная гипертензия. Симптомы: артериальное давление $\geq 140/90$ мм рт. ст. после 20-й недели беременности у ранее нормотензивной женщины.

-Преэклампсия. Симптомы: артериальное давление $\geq 140/90$ мм рт. ст. после 20-й недели беременности у женщины, которая ранее имела нормотензию. Протеинурия, экскреция белка с мочой ≥ 300 мг/л или 500 мг/24 ч при отсутствии инфекции мочевыводящих путей.

-Эклампсия. Встречается у женщин с преэклампсией. Характеризуется приступами, не связанными с другими причинами.

-Наложенная преэклампсия. Хроническая гипертензия с развитием протеинурии во время беременности.

Первоначально ПЭ была разделена на 2 подгруппы на основе гистологических, клинических, биологических, гемодинамических и эпидемиологических данных. Эти 2 подгруппы были обозначены как ранняя (<34 недель) и поздняя (≥ 34 недель) ПЭ в зависимости от гестационного возраста на момент постановки диагноза или родов. Гистологические поражения плаценты, сопряженные с материнской сосудистой патологией, обусловленной нарушением ремоделирования спиральных артерий и задержкой роста плода. Такой сценарий чаще встречается при ранней ПЭ, а извлечение плода и плаценты до 34 недель беременности увеличивает риск неонатальной смертности. В условиях ранней плацентарной ПЭ ремоделирование спиральных артерий происходит не в полной мере, артерии сохраняют свою способность к сокращению, что на поздних сроках беременности может приводить к снижению кровоснабжения плода – плацентарной недостаточности и, как следствие, внутриутробной задержке роста плода. При этом плод из-за снижения поступления к нему кислорода подаёт определённые сигналы материнскому организму, в ответ на которые организм женщины реагирует повышением кровяного давления. При повышении артериального давления кровоснабжение плода в норме должно усиливаться, чего не происходит в том случае, если спиральные артерии матки не потеряли своей способности к сокращению. В отличие от поздней ПЭ, гемодинамическое состояние при ранней ПЭ характеризуется периферической вазоконстрикцией без увеличения сердечного выброса.

В плацентарной, или ранней, стадии преэклампсии этиология обусловлена аномальной плацентацией в гипоксических условиях. Плацентарная, или ранняя начальная преэклампсия несёт значительно более высокий риск осложнений у матери и плода. У них также наблюдается более высокая распространённость плацентарных поражений, особенно между 28 и 32-й неделями беременности. Следовательно, плацентарная, или ранняя, начальная преэклампсия связана с ограничением роста плода и неблагоприятными материнскими и неонатальными исходами.

В материнской преэклампсии, или поздней стадии преэклампсии, проблема возникает из-за взаимодействия между предположительно нормальной плацентой и материнскими факторами, которые спровоцированы эндотелиальной дисфункцией. Что создаёт предпосылки к микрососудистым повреждениям. Преэклампсия матери, или позднего начала, по-видимому, является декомпенсированным ответом на окислительный стресс в плаценте, вызванный дисфункциональным материнским эндотелием. Эндотелиальная дисфункция, которая является одним из аспектов системного воспалительного ответа матери, может привести к генерализованной вазоконстрикции и уменьшению тока крови к нескольким органам, включая сердце, почки и мозг. С другой стороны, поздняя ПЭ, также известная как материнский подтип, развивается после 34 недель беременности. Было высказано предположение, что этот подтип является следствием материнской генетической предрасположенности к

сердечно-сосудистым и метаболическим заболеваниям, связанным с хроническим системным воспалением в сочетании с измененным соотношением между перфузией материнской крови и повышенными метаболическими потребностями плаценты и плода. Эти нарушения могут привести к плацентарной дисфункции и связанной с ней плацентарной гипоперфузии без нарушений плацентации, как это наблюдается при ранней ПЭ.

Однако, поскольку причина патологии, по-видимому, не находится в плаценте, наблюдается более низкая частота вовлеченности плода и более благоприятные перинатальные исходы.

Несмотря на очевидные патофизиологические различия между этими подтипами преэклампсии, зачастую их трудно разделить, потому что два подтипа могут наблюдаться одновременно, например, у поздней первородящей с сосудистыми заболеваниями, у которой отмечена аномальная плацентация. Таким образом, хотя разделение типов может быть полезным для понимания и прогнозирования состояния, большинство пациенток с преэклампсией имеют элементы обеих патологий.

К достаточно частым осложнениям ПЭ можно отнести поражение внутренних органов, включая почки, печень и мозг. ПЭ может вызвать синдром HELLP (гемолиз, повышенный уровень ферментов печени в крови и низкий уровень тромбоцитов), отек мозга, нарушение функции печени, почек или сердца, отслоение плаценты, задержку внутриутробного развития, преждевременные роды и даже смерть матери и плода.

Патогенетические механизмы, лежащие в основе ПЭ, еще предстоит выяснить; однако считается, что аномалии развития плаценты и инвазии трофобласта, плацентарная ишемия, окислительный стресс, воспаление, иммунная дезадаптация и тромбоз представляют собой ключевые факторы в развитии заболевания. Возможно, основная причина развития ПЭ заключается в нарушении процессов формирования плацентарной ткани на ранних сроках гестации вследствие аномальной инвазии цитотрофобласта и неполного ремоделирования спиральных артерий матки. Общепризнано, что патогенез ПЭ начинается во время инвазии трофобласта и ремоделирования спиральных артерий в течение первого и начала второго триместра беременности. Ранняя ПЭ характеризуется аномальной трофобластической инвазией из-за неполной трансформации спиральных артерий, что приводит к плацентарной ишемии, последующей воспалительной реакции и усиленному окислительному стрессу. Фактор роста эндотелия сосудов (VEGF), эндоглин и плацентарный фактор роста (PlGF) являются ангиогенными факторами, имеющими большое значение в этиологии ПЭ. Все указанные факторы способствуют системному дисбалансу антиангиогенных и проангиогенных факторов, которые вызывают клинический синдром. Следовательно, повышенная активация эндотелия у матери и общее гипервоспалительное состояние предшествуют появлению клинических симптомов ПЭ.

Женщины с ПЭ имеют потенциально опасное для жизни состояние здоровья, поскольку это заболевание связано с повышенным риском развития долгосрочных осложнений, таких как сердечно-сосудистые заболевания.

Установлено, что все эти компоненты в значительной мере определяются генетическими факторами, которые могут быть вовлечены в происходящие патогенные изменения. Семейная природа ПЭ известна уже много лет, в этой области были проведены обширные генетические исследования. Применение близнецового метода позволило установить, что наследуемость ПЭ составляет около 55%, причем вклад в развитие заболевания вносят, как материнские, так и фетальные гены. Наиболее актуальной является теория, определяющая ПЭ, как мультифакторное заболевание, в основе которого лежат наследственность и факторы среды.

Несмотря на большое количество исследований, проводимых в области взаимосвязи генетических полиморфизмов и ПЭ, утверждать, что ведущую роль играет только определенный генетический полиморфизм, нет оснований. Большинство исследователей высказывают мнение о том, что вероятнее всего только множественные комбинации полиморфных

вариантов генов в сочетании с факторами внешней среды приводят к развитию такого опасного осложнения беременности, как ПЭ. Воздействие факторов окружающей среды, включая возраст, стрессовые ситуации, избыточное потребление соли, хронические заболевания почек и сердечно-сосудистой системы, гипоксия, гиподинамия, ожирение, диабет, алкоголь, черепно-мозговая травма, в конечном счете, определяют, приведут ли полиморфизмы генов предрасположенности с низкой пенетрантностью к фенотипическим проявлениям заболевания. Вполне вероятно, что ни один из факторов среды или генетических вариантов не является универсальным и не может объяснить все случаи ПЭ. В то же время, возможно, что разнообразие их комбинаций, в свою очередь, оказывает влияние на разнообразие фенотипического проявления конкретных форм течения заболеваний (например, ПЭ в сочетании с задержкой внутриутробного роста).

НАСЛЕДСТВЕННАЯ ПРИРОДА ЗАБОЛЕВАНИЯ

В некоторых редких случаях семейной ПЭ патология, по-видимому, наследуется в соответствии с менделевскими моделями наследования заболеваний. В этом случае ПЭ определяется редким моногенным патологическим вариантом гена с высокой пенетрантностью. Однако в большинстве случаев ПЭ, по-видимому, представляет собой сложное генетическое расстройство и возникает в результате многочисленных общих вариантов генов в разных локусах, которые по отдельности имеют небольшие эффекты, но в совокупности способствуют восприимчивости беременной женщины к заболеванию.

С целью выявления генетической предрасположенности к ПЭ необходимо анализировать, как материнский, так и фетальный генотипы. Роль плаценты в первичном патогенезе ПЭ указывает на вклад генов плода в восприимчивость к заболеванию. Тесная связь тяжелой ранней ПЭ со случаями фетальных хромосомных аномалий также указывает на это. К тому же установлен повышенный риск ПЭ у женщин, отцы которых родились в условиях ПЭ. На основании сказанного была разработана гипотеза генетического конфликта, объясняющая генетическую предрасположенность к ПЭ. Согласно этой гипотезе фетальные (в т.ч. отцовские) гены оказывают стимулирующее влияние на плацентарный кровоток с целью увеличения передачи питательных веществ плоду от материнского организма. Напротив, материнские гены будут стремиться ограничить передачу питательных веществ с кровью сверх определенного оптимума, допустимого для материнского организма. Другими словами, фетальные гены повышают материнское кровяное давление, чтобы усилить маточно-плацентарный кровоток, тогда как материнские гены действуют противоположным образом. Таким образом, эндотелиальная дисфункция у матерей с ПЭ может быть интерпретирована как попытка плода компенсировать недостаточное маточно-плацентарное снабжение плода питательными веществами.

Первоначально, большинство генетических исследований ПЭ были проведены преимущественно с учетом материнского генотипа. Кроме того, снижение риска развития ПЭ при второй и последующих беременностях затрудняло анализ вклада генотипа плода. Тем не менее, в настоящее время существенное внимание уделяется генам, экспрессируемым в плаценте. По современным данным около 35% случаев ПЭ можно объяснить спецификой генотипа матери, 20% — спецификой генотипа плода, 13% — спецификой генотипа отца и 32% — факторами окружающей среды.

Анализ экспрессии генов является важным и эффективным инструментом для идентификации генов, дифференциально экспрессируемых между тканями нормотензивной и преэклампсической беременностей. В последнее время установлена связь более 100 полиморфных вариантов генов с патогенезом ПЭ. Генетическая карта заболевания достаточно разнообразна и включает группы генов системы свертывания крови, метаболизма, главного комплекса гистосовместимости, цитокинов и ростовых факторов, а также гены эндотелия и сосудистой системы.

ГЕНЫ ПРЕДРАСПОЛОЖЕННОСТИ К ПРЕЭКЛАМПСИИ

Выявление генов предрасположенности к ПЭ осуществляется преимущественно методом секвенирования нового поколения (иногда с использованием метода ПЦР) путем сравнительного анализа полиморфизма генов между пациентками с ПЭ и женщинами с физиологическим течением беременности. Большинство исследований анализируют один полиморфизм в гене-кандидате, в то время как в литературе имеются данные нескольких генах предрасположенности. Гены-кандидаты в данном случае выбираются на основе современного понимания патофизиологических механизмов ПЭ. При этом проведение популяционных исследований по принципу случай ПЭ/контроль позволяет выявить не только гены предрасположенности к ПЭ, но и гены, которые напротив, оказывают сдерживающий (защитный) эффект.

С другой стороны, популяционный анализ частоты встречаемости генетических факторов высокого или низкого риска может улучшить качество прогноза риска ПЭ. Кроме того, данная тактика способствует более эффективному раннему родовому скринингу и вмешательству в популяциях с более высокой восприимчивостью к ПЭ.

ГЕНЫ РЕГУЛЯЦИИ ФУНКЦИЙ ЭНДОТЕЛИЯ И КРОВЯНОГО ДАВЛЕНИЯ В ПАТОГЕНЕЗЕ ПЭ

Неблагоприятные прогнозы, связанные с гипертензивными расстройствами беременности являются результатом чрезмерного воспаления и повреждения эндотелия и включают эклампсию, инсульт, отслоение сетчатки, острую почечную недостаточность, отслойку плаценты, отек легких, гематому печени, диссеминированное внутрисосудистое свертывание и цереброваскулярное кровоизлияние. Осложнения для плода включают мертворождение, внутричерепное кровоизлияние,

маловодие и задержку роста плода. При этом, прогнозы состояния здоровья матери и плода, как правило, группируются вокруг диагноза ПРЕЭКЛАМПСИЯ. Таким образом, одним из важных симптомов ПЭ является повышенное артериальное давление во время беременности. Поэтому, гены, вовлеченные в регуляцию кровяного давления и гены-маркеры эндотелиальных дисфункций входят в состав всех генетических панелей анализа предрасположенности к ПЭ. Возможно, что плацентарное расстройство, в основном характеризующееся маточно-плацентарной ишемией, вторичной по отношению к нарушению сосудистого ремоделирования, является первой аномалией в патогенезе ПЭ. Накапливаются данные, которые свидетельствуют о том, что сердечно-сосудистое расстройство, вероятно, является первой аномалией, вызывающей плацентарную ишемию, которая в свою очередь и дает начало ПЭ.

ГЕН *VEGFA* VEGFA, фактор роста эндотелия сосудов А является продуктом гена *VEGFA*, который расположен на хромосоме 6 в цитогенетической полосе 6p21.3. Ген является членом семейства факторов роста PDGF/VEGF. Он кодирует гепарин-связывающий белок, который существует в виде гомодимера с дисульфидной связью. Этот фактор роста вызывает пролиферацию и миграцию эндотелиальных клеток сосудов и необходим, как для физиологического, так и для патологического ангиогенеза. Полиморфизм данного гена ассоциируется с опасными заболеваниями. В частности, данный ген активируется во многих известных опухолях, и его экспрессия коррелирует со стадией и прогрессированием опухоли. Повышенные уровни этого белка обнаружены у пациентов с синдромом ROEMS, также известным как синдром Кроу-Фуказе. Аллельные варианты этого гена связаны с микрососудистыми осложнениями диабета 1 типа (MVCD1) и атеросклерозом. Фактор роста эндотелия сосудов А необходим для развития плаценты и ее васкуляризации. VEGF жизненно важен для формирования трофобластов, эмбриональной сосудистой сети и клеток крови матери и плода. Возможно, аномальный рост сосудов и эндотелиальная дисфункция являются частью патогенеза ПЭ. Известно, что *VEGFA* экспрессируется на значительно более высоких уровнях в плацентах женщин с ПЭ по сравнению с плацентами здоровых женщин (соответственно 49,6% против 14,9%). Получены данные о том, что два полиморфизма гена VEGF (rs3025039, rs2010963) связаны с повышенным риском ПЭ в различных этнических группах, что означает, что эти мишени могут быть полезными генетическими маркерами для раннего прогнозирования ПЭ. В частности, имеется ряд сообщений о том, что rs3025039, расположенный в 3'-нетранслируемой области *VEGFA* и способный модулировать экспрессию гена, рассматривается, как фактор предрасположенности к ПЭ в популяциях Китая, Южной Европы и Северо-Восточной Африки, повышая риск заболевания. По вопросу ассоциации rs3025039 (NC_000006.12:g.43784799C>T) с ПЭ и ее возможной роли в патогенезе ПЭ в литературе были высказаны противоречивые мнения. Тем не менее, есть основания предположить, что в отдельных популяциях может присутствовать значимая связь между полиморфизмом гена *VEGFA* и ПЭ. Возможно, полиморфизм гена VEGF rs2010963, с.-94C>G также связан с повышенным риском ПЭ в различных этнических группах, что означает, что данный SNP может быть полезным генетическим маркером для раннего прогнозирования ПЭ.

ГЕН *ACE* Данный ген кодирует ангиотензинпревращающий фермент (АПФ), участвующий в регуляции артериального давления и водно-солевого баланса. Он катализирует превращение ангиотензина I в физиологически активный октапептид ангиотензин II. Ангиотензин II является мощным вазопрессором и пептидом, стимулирующим альдостерон, который контролирует артериальное давление и водно-электролитный баланс. Ангиотензинпревращающий фермент (АПФ) также инактивирует вазодилататорный белок брадикинин. Соответственно, кодируемый геном фермент повышает артериальное давление и является мишенью препаратов - ингибиторов АПФ, которые часто назначают для снижения артериального давления. Полиморфизм гена (мутации в этом гене) вовлечен в широкий спектр заболеваний, включая сердечно-сосудистые заболевания, псориаз, почечную недостаточность, инсульт и болезнь Альцгеймера. Вместе с тем, полиморфизм гена (rs4344) является широко известным маркером гипертонической болезни и других сердечно-сосудистых заболеваний, а также связан с риском ПЭ.

ГЕН *AGT* Белок, кодируемый этим геном, преангиотензиноген или предшественник ангиотензиногена, экспрессируется в печени и расщепляется ферментом ренином в ответ на пониженное артериальное давление и изменение транспорта натрия в почечных канальцах. Полученный продукт, ангиотензин-I, затем расщепляется ангиотензинпревращающим ферментом (АПФ) для получения физиологически активного фермента ангиотензина-II. Преангиотензиноген принимает участие в поддержании артериального давления, гомеостаза жидкости и электролитов в организме, а также в патогенезе эссенциальной гипертензии. Мутации в этом гене связаны с восприимчивостью к эссенциальной гипертензии и могут вызывать дисгенезию почечных канальцев, тяжелое нарушение развития почечных канальцев. Дефекты в этом гене также связаны с несемейной мерцательной аритмией и воспалительным

заболеванием кишечника. Полиморфизм гена *AGT* также тесно вовлечен в патогенез ПЭ. У беременных с ПЭ наблюдается повышение уровня белка AGT в моче. По имеющимся данным присутствие аллелей 217A, 235T и 704C гена *AGT* значительно увеличивало риск тяжелой ПЭ в 2,23 раза. В частности, частота генотипа CC полиморфизма T704C способствовала повышению риска ранней ПЭ по сравнению с таковой у пациенток с поздней преэклампсией.

ГЕН *NOS3* Кодирует фермент eNOS (эндотелиальная изоформа фермента NO-синтазы). Монооксид азота (NO) — это реактивный свободный радикал, который действует как биологический медиатор в нескольких процессах, включая нейротрансмиссию, антимикробную и противоопухолевую активность. Оксид азота синтезируется из аминокислоты L-аргинина с помощью синтаз оксида азота. Монооксид азота (NO), который оказывает влияние на сосудистый тонус и агрегацию тромбоцитов, стабилизацию реологических свойств крови, нейротрансмиссию, участвует в устранении метаболического синдрома, а также в ряде иммунных реакций.

Изменения в этом гене связаны с восприимчивостью к целому ряду заболеваний сердечно-сосудистой системы, в том числе к ПЭ. Аллельные варианты гена *NOS3* приводят к снижению уровня NO, следствием чего является развитие гипертензивных состояний, которые в сочетании с неблагоприятными условиями внешней и внутренней среды организма ассоциированы с изменением сосудистого тонуса, что может привести к ПЭ, плацентарной недостаточности, задержке роста плода. В патогенезе преэклампсии большое значение имеет эндотелиальная дисфункция, которая проявляется увеличением «чувствительности» сосудистой стенки к прессорным влияниям медиаторов с одновременным уменьшением продукции вазодилаторов, таких как оксид азота (NO). Исследования, проведенные в популяциях Северной Европы показали, что женщины, гомозиготные по rs1799983 в *NOS3*, в 1,6 раза чаще страдают тяжелой формой ПЭ. В частности, установлено, что генотип для SNP *NOS3* rs1799983 не только связан с риском ПЭ, но и снижением эффективности антигипертензивных препаратов у пациенток с ПЭ. Кроме того, полиморфизм 894G>T в 7 экзоне гена *NOS3* (маркер rs1799983) приводит к замене глутамина на аспарагин в 298 позиции белка, следствием чего являются уменьшение концентрации окиси азота в кровяном русле, и снижение вазодилатации. Важно отметить, что именно этот вариант чаще встречается в российских популяциях и ассоциирован с ПЭ. Сообщается о диагностической ценности rs2070744 в *NOS3* в раннем прогнозе ПЭ. Анализ мультилокусного взаимодействия семи аллелей, включая rs2070744 в *NOS3*, связанных с ПЭ (rs4762 в *AGT*, rs1800896 в *IL-10*, rs1800629 и rs1799724 в *TNFA*, rs2070744 в *NOS3*, rs7412 в *APOE* и rs2549782 в *ERAP2*) продемонстрировал достаточно высокую точность прогноза риска ПЭ.

ГЕНЫ РЕГУЛЯЦИИ ЛИПИДНОГО ОБМЕНА В ПАТОГЕНЕЗЕ ПЭ

Нарушения липидного обмена, индуцированные перекисным окислением липидов, вызванным окислительным стрессом, рассматривают в качестве важного патогенетического механизма ПЭ. Известно, что фермент липопротеинлипаза (*LPL*) и белок аполипопротеин Е (*ApoE*), основной функцией которого является транспортировка липидов между различными клетками и тканями организма, рассматриваются в качестве важных регуляторов липидного обмена, которые обильно экспрессируются в плаценте. В частности, *LPL* участвует в гидролизе ТГ в хиломикронах и гранулах липопротеинов очень низкой плотности. В связи с этим гены, кодирующие данные белки были предложены в качестве возможных генов-кандидатов. Также известно, что на фоне ПЭ происходит нарушение ковалентной модификации циркулирующих изоформ белка ApoE. В частности, доля дегликозилированной основной изоформы ApoE была увеличена при ПЭ, а кислая сialiлированная изоформа ApoE была уменьшена. Функционально это может увеличить риск развития атеротических изменений плаценты. Наиболее многообещающим генетическим вариантом в этом контексте является мутация-миссен, Asn318Ser, в *LPL*, которая коррелирует с пониженной активностью фермента *LPL* и повышенной дислипидемией. Также сообщалось, что фетальный генотип этих двух генов вносит вклад в метаболизм материнских липопротеинов. Транспорт липидов в плаценте является не только основным источником снабжения плода энергией, но и ключевым пластическим материалом, который способствует развитию нервной системы плода. Нарушение метаболизма липидов может вызвать воспалительные реакции и накопление продуктов окислительного стресса, влияющих на инвазию трофобласта и развитие плаценты, а также вызывающих повреждение эндотелия сосудов, что приводит к ряду патологических изменений. По сравнению с беременными женщинами с физиологическим течением беременности, уровень липидов в сыворотке и параметры метаболизма липопротеинов подвергаются нарушениям у пациенток с ПЭ. Высокий уровень липидов в крови у пациенток с ПЭ приводит к атеросклеротическим изменениям во время беременности, которые могут сохраняться в течение нескольких лет после родов, увеличивая риск сердечно-сосудистых заболеваний.

Некоторые авторы в качестве фактора риска ПЭ выделяют полиморфизмы гена *ADIPOQ*. Этот ген кодирует физиологически активный белок адипонектин и экспрессируется исключительно в жировой ткани. Ожирение оказывает негативное влияние на течение беременности и изменяет качество родов (преждевременные роды или затяжные роды) и более высокие потребности в окситоцине у беременных женщин. Жировая ткань, которая является крупнейшим эндокринным и паракринным органом, особенно при ожирении, отвечает за выработку адипокинов. Низкие концентрации адипонектина в сыворотке связаны с диабетом, избыточной массой тела, резистентностью к инсулину и метаболическим синдромом. Нарушение регуляции адипонектина во время беременности способствует гестационному сахарному диабету и преэклампсии.

Необходимо отметить, что нарушения липидного обмена организма, вызванные избыточной массой тела оказывают крайне негативное влияние на течение беременности, изменяя чувствительность тканей к половым стероидным гормонам и создают дополнительные риски во время самих родов.

ГЕН *LPL* кодирует липопротеинлипазу, которая экспрессируется в сердце, мышцах и жировой ткани. Фермент LPL функционирует как гомодимер и имеет двойную функцию триглицеридгидролазы и лиганд/мостикового фактора для рецептор-опосредованного захвата липопротеинов. Тяжелые мутации, вызывающие дефицит LPL, приводят к гиперлипопротеинемии I типа. Менее опасные формы мутаций в *LPL* связаны со многими нарушениями метаболизма липопротеинов. Липопротеинлипаза (LPL) принимает участие в процессах гидролиза триглицеридов и играет важную роль в транспорте липидов в плаценте. Выявлено, что материнский аллель *LPL* N291S (rs268, с.953A>G (p.Asn318Ser)) был связан с повышенным риском преэклампсии и с повышенным риском синдрома HELPP. Помимо этого присутствие аллеля *LPL* N291S в генотипе плода был связан с более высокими уровнями триглицеридов в организме матери и с повышенным уровнем риска ПЭ.

ГЕН *APOE* Белок, кодируемый этим геном, является основным апопротеином хиломикрона. Он связывается со специфическим рецептором печени и периферических клеток и необходим для нормального катаболизма богатых триглицеридами липопротеиновых компонентов. Этот ген локализован на хромосоме 19. Мутации в этом гене приводят к семейной дисбеталипопротеинемии или гиперлипопротеинемии типа III (HLP III), при которой повышенный уровень холестерина и триглицеридов в плазме является следствием нарушенного клиренса остатков хиломикрона и ЛПОНП. Известно, что уровни триглицеридов, холестерина липопротеинов низкой плотности и Apo-E с 4–16 недель до 28–42 недель беременности в группе беременных с ПЭ увеличились больше, чем в группе беременных с хронической гипертензией. Большинство авторов высказывают мнение о том, что повышенные уровни триглицеридов и Apo-E увеличивают риск развития преэклампсии. По некоторым данным аллель *ApoE* rs440446 ассоциируется с повышенным риском ПЭ, в то же время, указывается, что данная форма гена может участвовать в патогенезе желчекаменной болезни и связана с риском онкологических заболеваний желчевыводящих путей у мужчин.

ГЕН *ADIPOQ* Ген *ADIPOQ* состоит из трех экзонов и двух интронов, расположенных на хромосоме 3q27, где был картирован локус восприимчивости к диабету. Этот ген кодирует физиологически активный белок адипонектин и экспрессируется исключительно в жировой ткани. Он кодирует белок, схожий с коллагенами X и VIII и фактором комплемента C1q. Кодируемый белок циркулирует в плазме и участвует в метаболических и гормональных процессах. Адипонектин — наиболее распространенный циркулирующий белок, синтезируемый только в жировой ткани, который действует как гормон с противовоспалительными, антидиабетическими и инсулиносенситизирующими свойствами, как известно, играет важную роль в различных метаболических процессах, включая контроль глюкозы и катаболизм жирных кислот. Снижение уровня адипонектина в плазме тесно связано с патогенезом сахарного диабета 2 типа (СД2), ряда сердечно-сосудистых заболеваний, ожирения и нарушениях физиологического течения беременности. Адипонектин встречается в плазме в олигомерных комплексах, состоящих из тримеров (низкомолекулярный вес), гексамеров (средний молекулярный вес) и крупных мультимеров из 12–18 субъединиц (высокомолекулярный вес (HMW)). Кроме того, изоформа HMW более эффективно связывается со своими рецепторами и запускает АМР-активируемую протеинкиназу, один из основных ферментов, контролирующих метаболические эффекты адипонектина. Форма HMW адипонектина, опосредующая чувствительность к инсулину в периферических тканях, является наиболее физиологически активной формой с точки зрения гомеостаза глюкозы. Мутации в этом гене зачастую связаны с дефицитом адипонектина. Известно, что вариант гена rs2241766 (+45T/G) тесно коррелирует с риском нарушения течения беременности и ПЭ, вызванной ожирением и гестационным сахарным диабетом на фоне более низких уровней адипонектина в плазме крови.

ГЕНЫ СИСТЕМЫ РЕГУЛЯЦИИ АГРЕГАТНОГО СОСТОЯНИЯ КРОВИ В ПАТОГЕНЕЗЕ ПЭ

Тромбофилия — патологическое состояние, характеризующееся нарушением системы свёртываемости крови, при котором увеличивается риск развития тромбоза. Однако, для физиологического течения беременности необходимо развитие адекватного плацентарного кровообращения. Между тем, тромбофилии могут повышать риск плацентарной недостаточности из-за микрососудистого тромбоза плаценты, макрососудистых тромбозов или обоих этих факторов, а также оказывать влияние на рост и дифференцировку трофобласта. Широко признана роль аномалий механизмов свертывания крови в патогенезе ПЭ. Повреждение эндотелия при ПЭ тесно связано с изменением фенотипа с антикоагулянтного на прокоагулянтный и со сниженной эндотелиально опосредованной вазорелаксацией. Возможно, что этот фенотип присутствует до ПЭ во время беременности или может развиваться вследствие повреждения, инициированного во время плацентации. На фоне нарушения механизмов свертывания крови при ПЭ у женщин развивается выраженная тромбоцитопения, часто связанная с гемолизом, повышением уровня ферментов печени в плазме крови и синдромом низкого количества тромбоцитов (HELLP). В настоящее время широко известна роль трех генов тромбофилических факторов, фактора V Лейдена (*FVL*), метилентетрагидрофолата (*MTHFR*) и протромбина (*F2*).

ГЕН *FV* Ген кодирует фактор V, важный кофактор каскада свертывания крови. Этот фактор циркулирует в плазме и преобразуется в активную форму путем высвобождения активационного пептида тромбином во время коагуляции. Это создает тяжелую цепь и легкую цепь, которые удерживаются вместе ионами кальция. Активированный белок является кофактором, который участвует вместе с активированным фактором свертывания X для активации протромбина в тромбин. Дефекты этого гена приводят либо к аутосомно-рецессивному геморрагическому диатезу, либо к аутосомно-доминантной форме тромбофилии, которая известна как резистентность к активированному протеину C. Аномалия гена коагуляционного фактора *FV* заключается в однонуклеотидной замене гуанина на аденин в положении 1601 (rs6025, c.1601G>A (p.Arg534Gln)), вследствие этого *FV* теряет способность разрушаться. Современные исследования показывают, что *FV* rs6025 связан с повышенным риском преэклампсии в европейских и некоторых азиатских популяциях. У беременных женщин с мутацией *FV* чаще обнаруживаются тромбы в плаценте, что с большей долей вероятности повышает риск ПЭ и другой акушерской патологии, антенатальной гибели плода. Помимо этого, полиморфизм гена *FV*, в частности rs6022 (c.319G>C (p.Asp107His)) в сочетании с курением создаёт риск нарушения течения беременности и преждевременных родов.

ГЕН *F2* кодирует белок протромбин (также известный как фактор свертывания крови II). Этот белок протеолитически расщепляется в несколько этапов, образуя активированную сериновую протеазу тромбин. Активированный фермент тромбин играет важную роль в тромбообразовании и гемостазе, превращая фибриноген в фибрин во время образования сгустка крови, стимулируя агрегацию тромбоцитов и активируя дополнительные факторы свертывания крови. Тромбин также играет роль в пролиферации клеток, восстановлении тканей и ангиогенезе, а также в поддержании целостности сосудов во время эмбрионального развития и постнатальной жизни. Пептиды, полученные из C-конца этого белка, обладают антимикробной активностью против *E. coli* и *P. aeruginosa*. Мутации в этом гене приводят к различным формам тромбоза.

Наследственная мутация протромбина (G20210A) во время беременности усугубляет уже существующее состояние гиперкоагуляции. Замена нуклеотидного основания гуанина на аденин в положении 20210 (G20210A) гена протромбина (*F2*) является причиной повышения уровня синтеза протромбина и риска развития тромбозов. Мутация гена протромбина *F2* в сочетании с мутацией гена *FV* встречается достаточно часто и достоверно увеличивает риск развития осложнений в период беременности. Имеются указания на тесную связь мутации гена *F2* с неблагоприятными исходами беременности, такими как ПЭ, самопроизвольный аборт, и задержка внутриутробного развития плода.

ГЕНЫ ФАКТОРОВ ИММУННОЙ СИСТЕМЫ В ПАТОГЕНЕЗЕ ПЭ

В настоящее время в литературе приведено большое количество аргументов, подтверждающих иммунологические причины преэклампсии. Иммуногенетическая память предполагает повышенный риск ПЭ в случае уже имевшей место ПЭ в предыдущей беременности. При этом риск ПЭ увеличивается с 4,1% во время первой беременности до 14,7% для второй беременности после первой беременности,

осложненной ПЭ, и до 31,9% для третьей беременности в случае двух последовательных предыдущих беременностей, осложненных ПЭ. В отличие от защитной иммунологии, репродуктивная иммунология основана не на иммунологических системах распознавания своего/чужого, а на иммунотолерантности во взаимодействиях клеток матери и плода. Адаптация иммунной системы матери к протеканию беременности имеет решающее значение для определения исхода и физиологического протекания беременности. Увеличение случаев ПЭ у первородящих женщин, указывает на тесную связь между иммунными факторами и риском ПЭ. Кроме того, успехи в области вспомогательных репродуктивных технологий (широкое распространение экстракорпорального оплодотворения - ЭКО) создают новые проблемы для иммунной системы матери. Использование донорской спермы или яйцеклеток увеличивает риск ПЭ в три раза. Показано, что плаценты при ЭКО с использованием донорских ооцитов имеют значительное увеличение поражений, связанных с хроническим воспалением, по сравнению с плацентами при ЭКО с использованием недонорских ооцитов.

Основными механизмами ускользания иммунитета плода из-под контроля материнской иммунной системы на границе мать-плод являются снижение экспрессии молекул главного комплекса гистосовместимости клетками трофобласта, присутствие регуляторов комплемента, повышенная продукция индоламин-2,3-диоксигеназы, активация регуляторных Т-клеток и увеличение иммунных контрольных точек. Эта иммунная защита больше похожа на иммунные реакции, наблюдаемые в биологии опухолей, чем в биологии аллотрансплантата.

В частности, с точки зрения риска ПЭ представляют интерес гены главного комплекса гистосовместимости (МНС - человеческие лейкоцитарные антигены). МНС располагается в хромосоме 6. HLA молекулы, кодируемые этими генами, разделяют на антигены (АГ) I и II классов. Молекулы МНС I класса (HLA-A, HLA-B, HLA-C) экспрессируются на клеточной поверхности всех соматических клеток, исключением являются клетки трофобласта и эритроциты. Аллельные варианты HLA-молекулы МНС II класса: HLA-DP, HLA-DQ и HLA-DR. Основная функция белков МНС представление пептидных АГ Т-лимфоцитам. Наиболее достоверна ассоциация HLA АГ с ПЭ для HLA-DR4 антигена II типа. Достаточно значимыми в патогенезе ПЭ считают неклассические HLA-АГ: HLA-E, HLA-C HLA-, потому что они экспрессируются клетками трофобласта. Из них только HLA-C демонстрирует выраженный полиморфизм. Экспрессия HLA на имплантируемом цитотрофобласте важна, поскольку они взаимодействуют с рецепторами иммуноглобулинов клеток-киллеров (KIR), экспрессируемыми на материнских uNK (естественные киллеры матки; uterine natural killer) и цитотоксических Т-лимфоцитах, снижая их цитолитическую активность и стимулируя выработку цитокинов, необходимых для успешной плацентации. Естественные клетки-киллеры (NK) идентифицируют и уничтожают клетки, которые не экспрессируют молекулу ГКГ. Этот механизм распознавания важен для защиты живого организма от чужеродных патогенов и раковых клеток. В отличие от циркулирующих периферических NK-клеток крови, uNK-клетки, контактирующие с трофобластом, выполняют цитокинетическую, а не цитотоксическую функцию. Они активируются другими механизмами, включающими взаимодействия между их рецепторами и лигандами HLA (человеческий лейкоцитарный антиген)-C, G и E, экспрессируемыми трофобластом. Активация uNK или децидуальных NK-клеток приводит к выработке ангиогенных факторов, таких как VEGF, которые стимулируют ангиогенез, а также хемокинов IL-8 и IP-10, которые способствуют инвазии трофобласта. Нарушения взаимодействия клеток uNK с HLA-C на трофобласте изменяет децидуальный иммунный ответ, что приводит к нарушению имплантации трофобласта и недостаточному ремоделированию спиральной артерии, что связано с ПЭ.

Таким образом, плацента и солидные опухоли имеют сходство в отношении процесса клеточной инвазии и микроокружения. Действительно, опухолевые клетки и клетки трофобласта поддерживаются благоприятным микроокружением, которое допускает клеточную инвазию, ангиогенез и иммунотолерантность и, таким образом, иммунологическое ускользание от хозяина. Децидуальная оболочка и микроокружение опухоли имеют один и тот же тип иммунных клеток, таких как NK-клетки, Т-клетки, Т-регуляторные клетки, макрофаги, а также неиммунные клетки, такие как децидуальные стромальные клетки и фибробласты. В этом микроокружении как опухолевые, так и трофобластные клетки могут экспрессировать HLA-G, который взаимодействует с NK-клетками, снижая их токсичность. Эти клетки также могут секретировать иммуносупрессивные факторы, которые ингибируют активность Т-клеток, способствуя локальному привлечению Т-регуляторных клеток с их иммуносупрессивной функцией и снижают экспрессию полиморфных антигенов ГКГ класса I и II, чтобы уклониться от иммунной системы хозяина.

Во время беременности синцитиотрофобласты, которые контактируют с материнской тканью, не экспрессируют никаких антигенов МНС. Трофобласт, который взаимодействует с децидуальными клетками матки, экспрессирует только классические полиморфные антигены HLA-C класса Ia и неклассические неполоморфные антигены HLA-G и HLA-E.

HLA-G демонстрирует ограниченный полиморфизм, тем не менее, некоторые варианты гена также связаны с ПЭ. Данные варианты HLA-G, распознаваемые материнским организмом, как чужеродные, могут привести к гистонесовместимости между матерью и ребенком, что также приводит к ПЭ. HLA-G защищает клетки трофобласта от лизиса, регулирует уровень толерантности в системе взаимодействия между матерью и плодом. Один из аллельных вариантов этого гена HLA-G является инсерция/делеция 14 пар нуклеотидов в 3'-нетранслируемой области гена, которая может стать достаточно значимым маркером ПЭ. Наиболее оптимальным прогностическим инструментом в изучении генов HLA при патологии беременности, в частности ПЭ, является генотипирование материнских генов в совокупности с аллельными вариантами ДНК, как отцовского генотипа так и генотипа плода.

Таким образом, высокие показатели полиморфизма молекулы HLA-C, экспрессируемой трофобластом, и ее взаимодействие с высокополиморфным KIR (рецепторами иммуноглобулинов клеток-киллеров), экспрессируемым материнской NK-клеткой, допускает множество комбинаций и возможность несоответствия между матерью и плодом (высокодиверсифицированное взаимодействие). Напротив, неклассический ГКГ класса Ib включает предковые лиганды HLA-E, HLA-F и HLA-G и демонстрируют низкий уровень полиморфизма.

Описывая роль белков иммунной системы необходимо отметить, что некоторые провоспалительные цитокины также могут вносить известный вклад в патогенез ПЭ. В частности, установлено, что избыточная секреция фактора некроза опухоли альфа (TNF α) может принимать участие в активации эндотелия, что, в свою очередь, может способствовать возникновению симптомов ПЭ. Во время беременности TNF α оказывает влияние на синтез гормонов, архитектуру плаценты и эмбриональное развитие. Отметим, что уровни TNF α в плазме крови значительно выше у женщин с ПЭ, чем при физиологическом течении беременности. Неадекватная стимуляция секреции TNF α способствует усилению выработки активных форм кислорода, приводя к повреждению тканей, включая эндотелий. Между тем, окислительный стресс и воспалительный ответ могут играть общую роль в прогрессировании ПЭ. Понимание этого факта формирует основу для выявления генетического механизма ПЭ. Помимо этого, повышенные уровни TNF α при ПЭ негативно сказываются на функциях трофобласта, имеющих отношение к его миграционной активности, синцитиализации и, в дальнейшем, эндокринной функции. Повышенные уровни TNF α негативно влияют на взаимоотношение организмов матери и плода, изменяют секреторный профиль плацентарных иммуномодулирующих факторов, что, в свою очередь, влияет на состояние иммунной системы материнского организма.

Предполагают, что интерлейкин-10 (IL-10) также участвует в патогенезе преэклампсии, усиливая воспалительный ответ на клетки трофобласта, что приводит к снижению инвазии и ремоделированию спиральных артерий.

ГЕН HLA-G кодирует белок человеческого лейкоцитарного антигена HLA-G относится к паралограм тяжелой цепи главного комплекса гистосовместимости (ГКГ) класса I. Эта молекула класса I представляет собой гетеродимер, состоящий из тяжелой цепи и легкой цепи (бета-2 микроглобулин). Тяжелая цепь закреплена в мембране. HLA-G экспрессируется на клетках плаценты. Тяжелая цепь составляет приблизительно 45 кДа, а ее ген содержит 8 экзонов. Экзон 1 кодирует лидерный пептид, экзоны 2 и 3 кодируют домены альфа1 и альфа2, которые связывают пептид, экзон 4 кодирует домен альфа3, экзон 5 кодирует трансмембранную область, а экзон 6 кодирует цитоплазматический хвост.

Было идентифицировано семь различных транскриптов мРНК HLA-G, причем это разнообразие приписывается альтернативному сплайсингу его семи экзонов. Транскрипты HLA-G2, G3 и G4 транслируются в мембраносвязывающие изоформы; HLA-G5, G6, G7 в растворимые изоформы HLA-G (sHLA-G); а HLA-G1 в оба типа изоформ. HLA-G экспрессируется на поверхности преимплантационных эмбрионов и трофобластов. Уровень HLA-G в плазме у женщин с отслойкой плаценты был значительно снижен. Между тем, sHLA-G был обнаружен в культуральной среде эмбрионов, оплодотворенных *in vitro* (ЭКО), материнской крови, амниотической жидкости и фетальной крови. HLA-G играет важную роль в ремоделировании спиральных артерий, развитии плода и иммунной толерантности. Присутствие растворимого HLA-G в культуральной среде эмбриона значимо коррелирует с повышением частоты наступления беременности при использовании вспомогательных репродуктивных технологий (ЭКО). Кроме того, уровень растворимого HLA-G в сыворотке пациентов соотносится с исходом беременности после ЭКО.

Существует несколько однонуклеотидных полиморфизмов (SNP) в 3'-нетранслируемой области (UTR) HLA-G, которые влияют на стабильность, метаболизм и транспорт мРНК, включая +3035C/T, +3187A/G, +3196C/G, +3142C/G, +3001C/T, +3003C/T, +3010C/G, +3027C/A, +3032C/G, +3052C/T, +3092G/T, +3111A/G, +3121C/T, +3177G/T, +3183A/G и +3227A/G. Эти SNP изменяют связывание микро-РНК, что приводит к снижению экспрессии HLA-G. В свою очередь, аномальная экспрессия и полиморфизмы HLA-G связаны с неблагоприятными исходами беременности, такими как ПЭ и

привычный спонтанный аборт. Также выявлена значительная связь гаплотипов GC-ins (-725G>C SNP в промоторной области HLA-G) с бесплодием. Обсуждается роль вариантов гена rs9380142, rs1610696 и rs1707 в патогенезе ПЭ на фоне вирусных заболеваний.

ГЕН *HLA-DQB1* кодирует белок HLA-DQB1, который относится к паралогам бета-цепи HLA класса II. Эта молекула класса II представляет собой гетеродимер, состоящий из альфа (DQA) и бета-цепи (DQB), обе закреплены на мембране. Играет центральную роль в иммунной системе, представляя пептиды, полученные из внеклеточных белков. Молекулы класса II экспрессируются в антигенпрезентирующих клетках (APC: В-лимфоциты, дендритные клетки, макрофаги). Бета-цепь составляет приблизительно 26-28 кДа и содержит шесть экзонов. Экзон 1 кодирует лидерный пептид, экзоны 2 и 3 кодируют два внеклеточных домена, экзон 4 кодирует трансмембранный домен, а экзон 5 кодирует цитоплазматический хвост. В молекуле DQ, как альфа-цепь, так и бета-цепь содержат полиморфизмы, определяющие специфичность связывания пептида, что приводит к образованию до четырех различных молекул. Генотипирование этих полиморфизмов обычно проводится при трансплантации костного мозга. Альтернативный сплайсинг приводит к появлению множественных вариантов транскриптов.

По сравнению с группой женщин с физиологическим течением беременности, частота встречаемости аллеля HLA-DQB1*0502 была значительно выше в группе женщин с гипертензией, вызванной беременностью. Известно, что аллель HLA-DQB1*0503 может быть маркером высокого кровяного давления во время беременности, в то время, как женщины, несущие в своём генотипе аллель HLA-DQB1*04, подвержены повышенному риску развития ПЭ. Напротив, отсутствие аллеля HLA-DQB1*0602 является значимым фактором риска ПЭ.

ГЕН *TNF-α* кодирует многофункциональный провоспалительный цитокин, который принадлежит к суперсемейству факторов некроза опухоли (TNF). Ген расположен в области человеческого лейкоцитарного антигена (HLA) класса II главного комплекса гистосовместимости (MHC) на хромосомном участке 6p21.33. Этот цитокин в основном секретируется макрофагами. Он может связываться с рецепторами TNFRSF1A/TNFR1 и TNFRSF1B/TNFR2 и, таким образом, функционировать через них. Данный цитокин участвует в регуляции широкого спектра биологических процессов, включая пролиферацию клеток, дифференцировку, апоптоз, липидный обмен и коагуляцию. Цитокин участвует в патогенезе различных заболеваний, включая аутоиммунные заболевания, инсулинорезистентность, псориаз, ревматоидный артрит, анкилозирующий спондилит, туберкулез, аутосомно-доминантную поликистозную болезнь почек и рак. Мутации в этом гене влияют на восприимчивость к церебральной малярии, септическому шоку и болезни Альцгеймера. Исследования с нокаутом гена *TNF-α* на мышах предполагают нейтропротекторную функцию этого цитокина.

Во время беременности TNF-α может быть вовлечен в этиологию преэклампсии. Данный тезис находит подтверждение, поскольку у пациенток с преэклампсией были обнаружены повышенные уровни TNF-α в плазме крови, в то время как уровни IL-10 и IL-4 снижены. Этот дисбаланс цитокинов приводит к хроническому периферическому и плацентарному воспалению. Кроме того, есть данные, указывающие на некоторые SNP, которые действуют как транскрипционные регуляторы этого цитокина в промоторной области (-308G/A, -238G/A). В частности, полиморфизм G/A в позиции -308, который был связан с повышенным риском преэклампсии. Кроме того, женщины с преэклампсией значительно чаще являются носителями мутации, повышающей экспрессию TNF-α -308 A/G, чем женщины с нормальным артериальным давлением. Установлено, что частота встречаемости аллеля (AA) в позиции -308 значительно выше среди словацкой, финской, иранской и турецкой популяций с преэклампсией и турецких женщин с эклампсией. Привлекает внимание тот факт, что частота встречаемости аллеля AA в некоторых популяциях европейских стран была существенно выше среди пациенток с ПЭ в сочетании с задержкой внутриутробного развития (ЗВУР) по сравнению с пациентками с ПЭ, но без признаков ЗВУР.

Необходимо отметить, то полиморфизм промотора гена *TNF* в положении -308G/A (rs1800629) непосредственно влияет на экспрессию цитокина TNF-α, способствуя развитию тяжелой формы ПЭ в популяциях Европы, Азии, Ближнего Востока и Северной Америки. Возможно, варианты гена *TNF*, ассоциированные с ПЭ обладают популяционной специфичностью, поскольку аллель -238G/A (rs361525), используемая, как маркер заболевания в европейских популяциях не нашла подтверждений актуальности в прогнозе и диагностики ПЭ для популяций Юго-Восточной Азии.

ГЕН *IFNG* кодирует растворимый цитокин IFN-γ, который является членом класса интерферонов II типа. Ген локализован в хромосомном регионе 12q15. Кодированный белок секретируется клетками, как врожденной, так и адаптивной иммунной системы. Активный белок представляет собой гомодимер, связывающийся с рецептором интерферона гамма, который запускает клеточный ответ на

вирусные и микробные инфекции. Мутации в этом гене связаны с повышенной восприимчивостью к вирусным, бактериальным и паразитарным инфекциям, лейкемией и некоторыми аутоиммунными заболеваниями.

Высокие уровни IFN- γ во время беременности связаны с осложнениями беременности, такими как ПЭ, преждевременные роды и задержка внутриутробного развития. Цитокин IFN- γ , а также TNF- α , вызывают трофобластический апоптоз, ограниченную дифференцировку трофобласта, которая приводит к неполной трофобластической инвазии спиральных артерий и нарушенной имплантации плаценты, следствием чего является ПЭ. Известно, что SNP +874A/T (rs2430561), ассоциирован с преэклампсией, в котором генотип T/T в этой позиции связан с высокими уровнями продукции IFN- γ и коррелирует с повышенным риском тяжелой ПЭ у женщин бразильской и китайской популяций.

ГЕН *IL10* кодирует цитокин интерлейкин-10 (IL-10), который представляет собой цитокин, продуцируемый в основном моноцитами и в меньшей степени лимфоцитами. Этот цитокин оказывает плейотропное действие на иммунорегуляцию и воспаление. Он подавляет экспрессию цитокинов Th1, MHC класса II и костимулирующих молекул на макрофагах. Он также повышает выживаемость, пролиферацию и выработку антител В-клеток. Этот цитокин может блокировать активность NF- κ B и участвует в регуляции сигнального пути JAK-STAT. Исследования нокаута на мышах показали, что этот цитокин выполняет функцию важного иммунорегулятора в кишечном тракте. Мутации в этом гене связаны с повышенной восприимчивостью к инфекции ВИЧ-1 и к ревматоидному артриту.

При физиологическом течении беременности ИЛ-10 выполняет три основные полезные функции: способствует успешной плацентации, контролирует воспаление и регулирует сосудистую функцию. В связи с этим уровни сывороточного ИЛ-10 повышаются во время нормальной беременности и остаются высокими до родов. Кроме того, на фоне нормальной плацентарной ткани наблюдаются более высокие уровни ИЛ-10 в течение первого и второго триместров по сравнению с третьим триместром беременности, что обеспечивает важный баланс воспаления во взаимодействии организмов плода и матери.

Следует отметить, что некоторые SNP в проксимальных и дистальных (-1082A/G, -819T/C и -592A/C) областях промотора гена *IL10* оказывают угнетающее влияние на скорость транскрипции. Например, полиморфизм A/G в позиции -1082 связан с более низкой выработкой ИЛ-10. Он связан с повышенным риском преэклампсии у турецкого, бразильского и египетского населения. Установлено, что у женщин с генотипом IL-10 (-1082) AA риск развития преэклампсии в 3,38 раза выше на фоне более высоких уровней глюкозы в плазме крови и более низких уровней холестерина ЛПВП.

ГЕН *IL22* Ген является членом семейства цитокинов IL10, которые опосредуют клеточные воспалительные реакции. Кодируемый белок IL-22 играет ключевую роль в защите хозяина от внеклеточных патогенов на слизистых поверхностях, укрепляя функции эпителиального барьера, и участвует в гомеостазе тканей, восстановлении тканей и заживлении ран. Этот белок также обладает провоспалительными свойствами и играет роль в патогенезе нескольких заболеваний кишечника. Кодируемый белок является важнейшим цитокином, который регулирует иммунитет хозяина при инфекционных заболеваниях, включая COVID-19.

Известно, что концентрация IL-22 в плазме крови значительно выше в группе пациентов с ПЭ по сравнению с контрольной группой. Повышение продукции IL-22 приводит к активации иммунных реакций материнского организма. Сложившийся иммунный дисбаланс, в свою очередь, может привести к ПЭ через плацентарную ишемию. Кроме того, показано, что лабораторный анализ уровня IL-22 и соотношения концентраций IL2/IL-4 можно использовать в уточнении патологий беременности для различения преэклампсии, гестационной гипертензии и физиологической беременности.

Возможно, полиморфизм гена *IL17* следует рассматривать в качестве одной из причин повышения уровня IL-22 при ПЭ. Вместе с тем, однонуклеотидный полиморфизм (rs2227485) в гене *IL22* связан с риском преэклампсии. Установлено, что rs2227485 в *IL22* показал значительную разницу частоты встречаемости аллеля в группе с преэклампсией с ранним началом и генотипа подгрупп преэклампсии с поздним началом, а также контрольных подгрупп. Были обнаружены значительные достоверные различия между пациентами с ПЭ и контрольной группой для этого SNP с точки зрения генотипических частот. Как было установлено, патогенетический вклад rs2227485 в развитие ПЭ был оценен в рамках доминантной модели аллеля C (генотип CC/CT + TT) или рецессивной модели аллеля T (генотип TT/CC + CT).

РЕКОМЕНДОВАННАЯ ЛИТЕРАТУРА К РАЗДЕЛУ 4. ПРЕЭКЛАМПСИЯ

Генетический паспорт — основа индивидуальной и предиктивной медицины / Под ред. В. С. Баранова. — СПб.: Изд-во Н-Л, 2009. — 528 с.

Глотов А.С., Вашукова Е.С., Глотов О.С. и соавт. Исследование популяционных частот полиморфизма генов, ассоциированных с гестозом. Экологическая генетика.- 2013.- Т.XI, № 1.-С.91-100

Цахилова С.Г. , Акуленко Л.В., Кузнецов В.М. и соавт. Генетические предикторы преэклампсии. Проблемы репродукции.-2017.-№1.-С.110-114
doi: [10.17116/repro2017231110-114](https://doi.org/10.17116/repro2017231110-114)

Kaptilnyy V.A., Reyshtat D.Y. Preeclampsia: definition, new in pathogenesis, guidelines, treatment and prevention. V.F. Snegirev Archives of Obstetrics and Gynecology, Russian journal. 2020; 7(1): 19-30. (in Russ.). DOI: <http://dx.doi.org/10.18821/2313-8726-2020-7-1-19-30>

Williams P.J., Broughton Pipkin F. The genetics of pre-eclampsia and other hypertensive disorders of pregnancy. Best Pract Res Clin Obstet Gynaecol. 2011; 25(4-4): 405–417.
doi: [10.1016/j.bpobgyn.2011.02.007](https://doi.org/10.1016/j.bpobgyn.2011.02.007)

Kleinrouweler C. E., van Uitert M., Moerland P.D. и соавт. Differentially Expressed Genes in the Pre-Eclamptic Placenta: A Systematic Review and Meta-Analysis. PLoS One. 2013; 8(7): e68991.
doi: [10.1371/journal.pone.0068991](https://doi.org/10.1371/journal.pone.0068991)

Parada-Niño L., Castillo-León L.F., Morel A. Preeclampsia, Natural History, Genes, and miRNAs Associated with the Syndrome. J Pregnancy. 2022; 2022: 3851225.
doi: [10.1155/2022/3851225](https://doi.org/10.1155/2022/3851225)

Duan W., Xia C., Wang K. и соавт. A meta-analysis of the vascular endothelial growth factor polymorphisms associated with the risk of pre-eclampsia. Biosci Rep. 2020; 40(5): BSR20190209.
doi: [10.1042/BSR20190209](https://doi.org/10.1042/BSR20190209)

Rokotyanskaya E.A., Panova I.A., Malyshkina A.I. и соавт. Technologies for Prediction of Preeclampsia. Sovrem Tekhnologii Med. 2020; 12(5): 78–84.
doi: [10.17691/stm2020.12.5.09](https://doi.org/10.17691/stm2020.12.5.09)

Sandrim V.C., Rizzatti Luizon M., Pilan E. Interaction Between *NOS3* and *HMOX1* on Antihypertensive Drug Responsiveness in Preeclampsia. Rev Bras Ginecol Obstet. 2020; 42(8): 460–467.
doi: [10.1055/s-0040-1712484](https://doi.org/10.1055/s-0040-1712484)

Lykke J.A., Bare L.A., Olsen J. и соавт. Vascular associated gene variants in patients with preeclampsia: results from the Danish National Birth Cohort. Acta Obstet Gynecol Scand. 2012;91(9):1053-1060.
doi: [10.1111/j.1600-0412.2012.01479.x](https://doi.org/10.1111/j.1600-0412.2012.01479.x)

Liu L., Zhang X., Qin K. и соавт. Characteristics of Serum Lipid Metabolism among Women Complicated with Hypertensive Disorders in Pregnancy: A Retrospective Cohort Study in Mainland China. Obstet Gynecol Int. 2024; 2024: 9070748.
doi: [10.1155/2024/9070748](https://doi.org/10.1155/2024/9070748)

Hajagos-Tóth J., Ducza E., Samavati R. и соавт. Obesity in pregnancy: a novel concept on the roles of adipokines in uterine contractility. Croat Med J. 2017; 58(2): 96–104
doi: [10.3325/cmj.2017.58.96](https://doi.org/10.3325/cmj.2017.58.96)

Howlader M., Sultana I., Akter F., Hossain M. Adiponectin gene polymorphisms associated with diabetes mellitus: A descriptive review. Heliyon. 2021; 7(8): e07851.
doi: [10.1016/j.heliyon.2021.e07851](https://doi.org/10.1016/j.heliyon.2021.e07851)

Zhang G., Zhao J., Yi J. Association Between Gene Polymorphisms on Chromosome 1 and Susceptibility to Pre-Eclampsia: An Updated Meta-Analysis. Med Sci Monit. 2016; 22: 2202–2214
doi: [10.12659/MSM.896552](https://doi.org/10.12659/MSM.896552)

Boulanger H, Bounan S., Mahdhi A. и соавт. Immunologic aspects of preeclampsia. *AJOG Glob Rep.* 2024; 4(1): 100321.
doi: [10.1016/j.xagr.2024.100321](https://doi.org/10.1016/j.xagr.2024.100321)

Xu X., Zhou Y, Wei H. Roles of HLA-G in the Maternal-Fetal Immune Microenvironment. *Front Immunol.* 2020; 11: 592010.
doi: [10.3389/fimmu.2020.592010](https://doi.org/10.3389/fimmu.2020.592010)

Zhuang B., Shang J., Yao Y. HLA-G: An Important Mediator of Maternal-Fetal Immune-Tolerance. *Front Immunol.* 2021; 12: 744324.
doi: [10.3389/fimmu.2021.744324](https://doi.org/10.3389/fimmu.2021.744324)

Wang X., Zhu H., Lei L. и соавт. Integrated Analysis of Key Genes and Pathways Involved in Fetal Growth Restriction and Their Associations With the Dysregulation of the Maternal Immune System. *Front Genet.* 2020; 11: 581789.
doi: [10.3389/fgene.2020.581789](https://doi.org/10.3389/fgene.2020.581789)

Mora-Palazuelos C., Bermúdez M., Aguilar-Medina M. и соавт. Cytokine-polymorphisms associated with Preeclampsia: A review. *Medicine (Baltimore).* 2022; 101(39): e30870.
doi: [10.1097/MD.00000000000030870](https://doi.org/10.1097/MD.00000000000030870)

Chen A., Zhao H., Wang J. и соавт. Haplotype Analysis of Candidate Genes Involved in Inflammation and Oxidative Stress and the Susceptibility to Preeclampsia. *J Immunol Res.* 2020; 2020: 4683798.
doi: [10.1155/2020/4683798](https://doi.org/10.1155/2020/4683798)

РАЗДЕЛ 5. ПОЛИГЕННЫЕ ОСНОВЫ ОНКОЛОГИЧЕСКИХ МУЛЬТИФАКТОРНЫХ ЗАБОЛЕВАНИЙ ЧЕЛОВЕКА НА ПРИМЕРЕ РАКА МОЛОЧНОЙ ЖЕЛЕЗЫ (РМЖ)

ВВЕДЕНИЕ

Рак молочной железы (РМЖ) является одним из наиболее распространенных видов рака в мире, в период с 2015 по 2020 год рак груди был диагностирован у 7,8 миллионов женщин. В 2020 году во всем мире было диагностировано 2,3 миллиона женщин, из которых 685 000 умерли от этой болезни. Рак груди встречается в каждой стране мира, причем риск развития заболевания увеличивается с возрастом. При этом, чем раньше рак груди будет обнаружен, тем больше шансов на долгосрочное выживание.

Использование моделей риска, включающих стандартные факторы риска, такие как семейный анамнез и гормональные/репродуктивные факторы, интегрированные с маммографической плотностью и ДНК-тестированием, теперь дает возможность достаточно точно оценить риск развития заболевания и на уровне популяции.

Исследования в области клеточной и молекулярной онкологии способствовали разработке генетических критериев, позволяющих определять персональный риск развития РМЖ, а также оказывать влияние на выбор схемы терапии, выработать индивидуальные рекомендации по профилактике и диагностике заболевания. Анализ генетического профиля позволяет оценить предрасположенность к РМЖ, в особенности у женщин с положительным семейным анамнезом, что дает возможность более эффективно в более раннем возрасте проводить скрининг для носителей риск-ассоциированных вариантов генов.

В настоящее время сформулированы **факторы риска развития РМЖ**, которым относятся:

- возраст: по данным Международного агентства по изучению рака заболеваемость РМЖ составляет от 1,5 случая в возрасте 20–24 лет до 421,3 случая к 75–79 годам на 100 тыс. женщин;

- гормональные препараты: прием гормональных препаратов в менопаузе и заместительная гормональная терапия у женщин в постменопаузе повышают риск развития гормон-позитивных подтипов РМЖ;

- семейный анамнез РМЖ и генетические факторы риска: эпидемиологические исследования сообщают о 1,8-кратном риске РМЖ в семьях с одним родственником первой степени с диагностированным заболеванием, 2,93-кратном риске – при наличии двух, 3,9-кратном риске – трех и более родственников;

- показатели риска РМЖ могут варьировать в зависимости от этнической принадлежности: так, представительницы субэтнической группы евреев-ашкенази обладают повышенным риском развития заболевания;

- употребление алкоголя и курение: умеренное употребление алкоголя увеличивает риск ОЗ в 1,04–1,61 раза. Курение также увеличивает риск развития РМЖ: более того, отмечена более высокая смертность от заболевания среди активных курильщиков;

- прохождение лучевой терапии на область грудной клетки в раннем возрасте также увеличивает риск развития РМЖ более чем в 4 раза.

ПОЛИГЕННАЯ ПРИРОДА ОНКОЛОГИЧЕСКИХ ЗАБОЛЕВАНИЙ. РОЛЬ ПОЛИМОРФИЗМА ГЕНОВ В РАЗВИТИИ ОНКОПАТОЛОГИЙ

Широко известна роль мутаций генов супрессоров опухоли и протоонкогенов в патогенезе онкологических заболеваний. Наряду с этим, полиморфизм генов также может вносить существенный вклад в формирование предрасположенности (повышение риска) к онкологическим заболеваниям. В данном случае речь идет о том, что специфические особенности индивидуального генетического профиля способны играть важную роль в формировании риска онкологического заболевания. Однако, достоверно выявить роль возможных вариаций аллелей генома в патогенезе заболеваний – сложная задача. Идентификацию генов предрасположенности к данной патологии необходимо проводить с учетом генетической структуры данной конкретной популяции и принадлежности испытуемых к различным этническим группам. Тем более, что генотип человека формируется из тысяч взаимодействующих полиморфных аллелей, причем каждый полиморфизм в отдельности обладает лишь весьма умеренным эффектом. В настоящее время установлены сотни генов-кандидатов, полиморфные аллели которые могут принимать участие в формировании онкологического риска.

Риск заболевания раком тесно связан с характером ответа клетки на повреждение ДНК. Т.е. эффективность репаративных процессов вполне допустимо рассматривать, как главный фактор онкорезистентности. Большое количество повреждений ДНК устраняется согласованными действиями более чем 100 генов системы контроля репарации ДНК. Кроме того, значительная часть всех ферментов

киназ (около 500 генов и более 1000 белков) участвует в формировании ответа клетки на повреждение ДНК. Например, обеспечивая активацию p53, осуществляя баланс выживаемости и гибель клетки, контролируя состояние митотического цикла. Апоптоз клеток, получивших грубые повреждения ДНК – еще один механизм онкорезистентности, широко используемый в организме.

Учитывая приведенные факты, вполне логично предположить, что изучение полиморфизма генов, регулирующих механизмы репарации ДНК, является первоочередным направлением в выявлении генов предрасположенности к онкологическим заболеваниям, включая рак молочной железы (РМЖ). Действительно, большая часть генов маркеров РМЖ представлены именно аллелями генов, определяющими эффективность репарации ДНК: ген *TP53*, гены *BRCA1* и *BRCA2*, ген *BRIP1*, ген *ATM*, ген *CHEK2*, ген *RAD51D*. Более того, данные гены функционально связаны между собой и формируют единую генную сеть, нацеленную на предотвращение генных мутаций и хромосомной нестабильности.

Наряду с этим, специфика мультифакторных заболеваний подразумевает наличие не только генетического фактора предрасположенности к заболеванию, но и фактора внешней среды, который провоцирует фенотипическое проявление патологического признака. Для многих онкологических заболеваний таким фактором среды является рацион питания. Противоопухолевый эффект строгого ограничения калорийности рациона питания хорошо известен. Также известно, что дефицит в организме человека некоторых витаминов и микроэлементов, таких, как фолиевая кислота, В6, В12, селен, железо, цинк, и витамины С и Е создает дополнительный риск развития онкологических заболеваний. Кроме того, в слоях населения, потребляющих минимум фруктов и овощей, онкологические заболевания наблюдаются вдвое чаще, по сравнению с людьми, соблюдающих диету, сбалансированную по фруктам и овощам.

Негенетическими, наиболее значимыми факторами риска заболеваемости РМЖ являются возраст (более 40 лет); высокий уровень эстрогенов в крови; прием гормональных препаратов с целью контрацепции или регуляции менструального цикла, заместительной гормональной терапии в менопаузе.

ГЕНЕТИЧЕСКИЕ ФАКТОРЫ РИСКА РМЖ

Рак молочной железы является наследственным заболеванием. Эпидемиологические исследования показывают, что около 4–5% случаев обусловлены высокой пенетрантной аутосомно-доминантной предрасположенностью. Результаты использования близнецового метода показывают, что около 27% случаев, вероятно, имеют сильный наследственный компонент. Существует много генов высокого и среднего риска, которые связаны с развитием рака молочной железы.

Положительный семейный анамнез имеют от 15 до 30% пациентов с РМЖ. В среднем от 5 до 10% случаев считаются наследственными, из которых в 4–5% присутствуют патогенные мутации в генах высокой пенетрантности, наследуемые по аутосомно-доминантному типу.

Вклад генетических факторов в развитие РМЖ характеризуется показателем пенетрантности генов, связанных с заболеванием. **В медицинской генетике понятие пенетрантности определяется долей лиц с генетической мутацией, у которых данный ген присутствует и это приводит к проявлению клинических симптомов заболевания, по отношению к лицам, у которых данный ген имеется, но клинические симптомы отсутствуют.** В частности, мутации в генах высокой пенетрантности связаны с более чем 4-кратным риском развития РМЖ по сравнению с общепопуляционным риском. В то время как патогенные варианты в генах умеренной пенетрантности свидетельствуют о 2–4-кратном риске по сравнению с общей популяцией. К категории низкопенетрантных относят гены, связанные с менее чем 2-кратным риском развития заболевания или при недостаточности сведений об их вкладе в развитие заболевания. При этом благодаря накоплению клинической информации и публикации результатов новых исследований список низкопенетрантных генов непрерывно дополняется.

Однонуклеотидные полиморфизмы (SNP) являются наиболее распространенной причиной появления новых вариантов в последовательности ДНК. В геноме одного человека встречается более 10 миллионов SNP. Наличие SNP в ДНК, с одной стороны, объясняют нормальные вариации между людьми и часто оказывают минимальное функциональное воздействие. С другой стороны, в том случае, когда SNP встречается в гене или в регуляторной области рядом с геном, функция гена может быть затронута, и это может стать причиной развития заболевания.

Многие SNP по отдельности связаны с достаточно небольшим увеличением риска РМЖ. Например, 14% женщин Великобритании, которые являются гомозиготными по rs2981582 аллели гена *FGFR2* (ген кодирует белок-рецептор фактора роста фибробластов, который взаимодействует с факторами роста фибробластов, приводя в движение каскад нисходящих сигналов, в конечном итоге влияя на митотический цикл и дифференцировку клеток) имеют повышенный (1,38 против 1,0) риск развития РМЖ по сравнению со средним популяционным показателем риска. В то же время, 38%

женщин, у которых отсутствует данный аллель, обладают пониженным риском (0,83) развития РМЖ. Выявленная закономерность может свидетельствовать о том, что продукты экспрессии разных аллелей гена могут обуславливать, как повышенную восприимчивость так и резистентность к заболеванию. Отметим, что SNP, ассоциированные с РМЖ, встречаются в популяции гораздо чаще, чем патогенные варианты, обнаруженные в генах высокого и среднего риска. На сегодняшний день идентифицировано более 300 таких однонуклеотидных полиморфизмов и риски, связанные с каждым из них, умножают онкогенные эффекты друг друга. Около 30% случаев семейного РМЖ можно объяснить известными однонуклеотидными полиморфизмами. В связи с наследуемыми формами рака молочной железы особое внимание привлекают SNP, расположенные в регуляторных зонах гена, выполняющие функции сайтов связывания факторов транскрипции.

Большинство социально значимых заболеваний, включая РМЖ, являются многофакторными, и генетическая составляющая обусловлена совместным влиянием ряда определенных мутаций в ДНК. В случае таких заболеваний выявление новых маркеров для диагностики представляется весьма сложной задачей. Один из подходов, который можно применить для ее решения, – это метод полногеномного поиска ассоциаций (Genome-wide association study, GWAS).

Метод полногеномного поиска ассоциаций (Genome-wide association study, GWAS) основан на выявлении ряда однонуклеотидных замен, которые чаще встречаются у больных людей и отсутствуют (или встречаются редко) у здоровых. Выявление набора таких мутаций осуществляют с помощью ряда статистических подходов или методами машинного анализа путем сравнения геномов двух групп (группы больных и здоровых людей). Набор однонуклеотидных замен, ассоциированных с каким-либо заболеванием, получил название polygenic risk score (PRS) – полигенный индекс риска.

Поиск генетических вариантов, приводящих к возникновению рака молочной железы проводят методом полногеномного поиска ассоциаций. Далее на основе данных этого поиска возможно определение полигенных индексов риска (PRS), разделение женщин на группы для проведения лечения и выбора мер профилактики, наиболее эффективных для каждого конкретного вида рака.

Согласно опубликованным в 2019 году результатам исследований было выделено 313 SNP, ассоциированных с РМЖ, с учетом ER-положительного и ER-отрицательного типа опухоли (эстроген-рецептор ER-положительный или ER-отрицательный). Поскольку многие варианты генов создают риски, степень которых отличается в зависимости от подтипа рака молочной железы. Это свидетельствует о том, что PRS, специфичные для подтипа опухоли, могут позволить лучше прогнозировать заболевание, специфичное для данного подтипа, включая более агрессивный ER-отрицательный рак молочной железы, и позволять выбирать женщин для профилактического лечения. Исследования проведены с использованием образцов ДНК 94 075 пациентов с РМЖ и 75 017 контрольных проб. Так была сформирована генетическая панель прогноза и диагностики РМЖ - PRS313.

При этом следует отметить, что, **во-первых**, полигенные индексы риска (PRS) являются основным компонентом достаточно точного прогнозирования риска рака груди и обладают потенциалом для улучшения стратегий скрининга и профилактики. Показатель PRS обеспечивает оценку генетической предрасположенности к заболеваниям и обычно рассчитывается на основе совокупного воздействия множественных связанных с заболеванием общих генетических вариантов или однонуклеотидных полиморфизмов (SNP), полученных из исследований ассоциаций по всему геному (GWAS). В данном случае речь идет, преимущественно, о герминальных мутациях. Индексы PRS объединяют риски от унаследованных однонуклеотидных полиморфизмов (SNP), связанных с раком груди в исследованиях ассоциаций по геному (GWAS), и объясняют более 30% наследуемости рака груди. Персонализированная оценка риска PRS, помогает выявлять женщин с более высоким риском развития рака груди и позволяет внедрять индивидуальные подходы к скринингу и профилактике. **Во-вторых**, современные методы молекулярной биологии позволяют улучшить не только качество прогноза риска развития болезни, но и способствуют пониманию молекулярной основы прогрессирования опухоли и прогнозировать её реакции на различные методы противораковой терапии. Выявление молекулярных биомаркеров, которые могут быть полезны в качестве прогностических и предиктивных маркеров, оказывает помощь врачам в принятии терапевтических решений, проведении лечения с более индивидуальным подходом и, следовательно, оптимизации терапии, а также избежания ошибочной стратегии лечения. Например, прогностические генетические маркеры могут помочь врачам предсказать агрессивность и инвазивность опухоли, что позволяет применять более оптимальную тактику лечения.

ГЕНЫ-МАРКЕРЫ РАКА ГРУДИ

Наиболее высокий риск развития РМЖ сопряжен с генами высокой пенетрантности, которая определяется более чем 5-кратным увеличением риска развития патологии по сравнению с общей популяцией. К таким генам относят *BRCA1*, *BRCA2*, *TP53*, *CDH1*, *PTEN*, *STK11 (LKB1)*.

К генам умеренной пенетрантности, ассоциированным с развитием РМЖ, относят *ATM* и *CHEK2*, участвующие в репарации двуниевых разрывов ДНК, ген-модификаторы *BRCA1* и *BRCA2* – *BRIP1 (BACH1)* и *PALB2 (FANCN)*. Мутации в генах умеренной пенетрантности встречаются в 2–5% РМЖ и могут увеличивать вероятность развития РМЖ в 2–4 раза. Тем не менее, некоторые мутации в генах умеренной пенетрантности, в частности *PALB2* и *CHEK2*, могут обуславливать настолько же высокий риск РМЖ для женщин с положительным семейным анамнезом, как и мутации в генах высокой пенетрантности.

Существующие тенденции в определении индивидуального риска развития онкологических заболеваний и, в частности РМЖ заключаются в интеграции полигенного компонента в существующие математические модели. Оценка риска, включающая определение суммы всех потенциально опасных вариантов, может обеспечить создание наиболее правильной схемы для идентификации женщин с повышенным генетическим риском развития РМЖ.

В современных методах оценки риска онкологического заболевания и диагностики широко применяются мультигенные панели на базе технологии секвенирования следующего поколения. Ниже приводятся характеристики некоторых генов, наиболее актуальных при анализе рака молочной железы.

ГЕН *TP53* кодирует белок-супрессор опухолей p53, содержащий домены транскрипционной активации, связывания ДНК и олигомеризации. Кодированный белок реагирует на различные клеточные стрессы, регулируя экспрессию целевых генов, тем самым вызывая остановку клеточного цикла, апоптоз, старение, репарацию ДНК или изменения метаболизма. Мутации в этом гене связаны с различными видами рака у человека, включая наследственные виды рака, такие как синдром Ли-Фраумени. Герминальные мутации *TP53* сопряжены с повышенным риском развития многих типов рака, включая РМЖ. При этом частота соматических мутаций в *TP53* выше среди базального подтипа РМЖ. Опухоли с мутациями в гене *TP53* характеризуются более агрессивным течением заболевания, резистентностью к химио- и радиотерапии. Сразу отметим, что ген *TP53* имеет большой потенциал в качестве молекулярного маркера риска РМЖ, подобно *BRCA1*. Оба гена-супрессора опухолей являются потенциальными биомаркерами для наблюдения, ранней оценки риска и предрасположенности к РМЖ, являясь терапевтическими мишенями для химиопрофилактики и таргетной терапии. При раке молочной железы чаще всего мутирует ген *TP53*, который присутствует примерно в 30–35% случаев. Интенсивность мутаций *TP53* различается в зависимости от подтипов рака молочной железы. Достигая 80% случаев мутаций при наиболее опасной формы РМЖ - трижды негативного рака молочной железы (TNBC - тип рака груди, который характеризуется отсутствием трех ключевых рецепторов: рецепторов эстрогена, прогестерона и HER2) и в 70% случаев HER2-позитивного рака молочной железы. Мутации *TP53* представляют собой важный биомаркер в клинической практике и потенциальную терапевтическую цель. Характер мутации *TP53* определяется с помощью секвенирования ДНК. Долгое время опухоли с мутациями *TP53* характеризовались резистентностью к терапии. Однако в настоящее время разрабатываются новые стратегии в области таргетной терапии, способные восстанавливать свойства дикого типа p53, предоставляя новые возможности в лечении рака. В частности, вещества производные 3-хинуклидина обладают способностью реактивировать мутантный p53, вызывая терапевтический апоптотический ответ в клетках TNBC опухоли.

ГЕНЫ *BRCA1* и *BRCA2* являются генами-супрессорами опухолей, играющими фундаментальную роль в репарации ДНК путем гомологичной рекомбинации. Герминальные мутации в этих генах связаны с повышенным риском развития рака груди и яичников. Средний кумулятивный риск рака груди в возрасте 80 лет для генов *BRCA1* и *BRCA2* составляет 72% и 69% соответственно. Потеря функции этих генов приводит к неэффективной репарации ДНК, увеличивая частоту мутаций и способствует развитию опухолей. Пациенты с патогенными или вероятно патогенными вариантами *BRCA1* имеют предрасположенность к развитию трижды негативного рака молочной железы TNBC, в то время как наличие патогенных или вероятно патогенных вариантов в *BRCA2* чаще ассоциируются с ER-позитивными опухолями. В данное время проводится интенсивное изучение ценности лабораторного анализа мутаций генов *BRCA* в составлении прогноза выживаемости пациентов и рисков развития метастазов. Опухоли, имеющие патогенные мутации *BRCA1/BRCA2* обладают более высокой чувствительностью, как к химиотерапии, так и к препаратам таргетной терапии.

Ген *PTEN* - гомолог фосфатазы и тензина, ген-супрессор опухолей. Белок, кодируемый этим геном, представляет собой фосфатидилинозитол-3,4,5-трифосфат-3-фосфатазу. Он содержит тензиноподобный домен, а также каталитический домен, аналогичный домену тирозиновых протеинфосфатаз с двойной специфичностью. В отличие от большинства тирозиновых протеинфосфатаз,

этот белок преимущественно дефосфорилирует не протеины, а фосфоинозитидные субстраты. Он отрицательно регулирует внутриклеточные уровни фосфатидинозитол-3,4,5-трифосфата в клетках и функционирует как супрессор опухолей, подавляя активность сигнального пути PI3K/Akt. Патогенные генетические варианты в гене фермента фосфатазы *PTEN* ассоциированы с риском развития многих онкологических заболеваний, включая колоректальный рак, рак почки, меланому и РМЖ. Мутации промоторного региона гена *PTEN* ассоциированы со значительным риском развития РМЖ, рака эндометрия и щитовидной железы. Данный ген является одним из наиболее часто мутирующих генов при онкологических заболеваниях, включая рак молочной железы, и его роль неразрывно связана с прогрессированием клеточного цикла, ростом клеток и выживанием. Опухолевые клетки с потерей гетерозиготности или мутациями *PTEN* значительно увеличили миграционную и инвазивную активность, способствуя пролиферации, инвазии и метастазированию. В метастатических клетках РМЖ уровни *PTEN* намного ниже, чем в локализованных раковых клетках. Потеря функции *PTEN* приводит к чрезмерной активации онкогенного сигнального пути клетки PI3K/Akt, который стимулирует рост и выживание клеток. Поэтому потеря активности *PTEN* была зарегистрирована почти в половине всех случаев РМЖ. **Инактивация *PTEN* происходит в основном в результате соматических мутаций.** Потеря *PTEN*, как правило, может свидетельствовать о резистентности опухоли к различным методам лечения.

ГЕН *CHEK2* контрольной точки киназы-2, кодирует белок серин/треонин киназу CHK2, который участвует в механизмах репарации ДНК. Он функционирует как важный ген-супрессор опухолей в регуляции клеточного цикла, ингибировании пролиферации клеток, активации восстановления ДНК и апоптоза. Аномальная экспрессия *CHEK2* может привести, в том числе, к раку груди. Установлено, что некоторые патологические варианты гена *CHEK2* с.470T>C (p.Ile157Thr) приводят к тому, что снижается связывание белка CHEK2 с белками Cdc25A, BRCA1 и p53, но не оказывает влияния на протеинкиназную активность белка CHEK2. Герминальные мутации *CHEK2* связывают с восприимчивостью к нескольким типам рака, с частотой мутации 1,08% у пациентов с РМЖ. Достаточно широко распространенный патологический вариант гена *CHEK2* (с.349A>G (p.Arg117Gly) приводит к более чем двукратному повышению риска рака молочной железы. Этот вариант был выявлен во многих случаях рака, включая семейные случаи рака молочной железы/яичников, предстательной железы, поджелудочной железы, колоректального рака, а также других типов рака. Данный патогенный вариант *CHEK2* приводит к потере активности протеинкиназы, что и обеспечивают увеличение риска развития РМЖ. У большинства пациентов с патогенными или вероятно патогенными вариантами *CHEK2* развиваются подтипы РМЖ люминальный А (наличие ER (рецептор эстрогена) и PR (рецептор прогестерона), отрицательным HER2 статусом и с низким Ki, самый часто диагностируемый подтип опухолей молочной железы) или люминальный В (присутствуют рецепторы к эстрогену, а рецепторы прогестерона могут отсутствовать, уровень Ki67 – высокий. Данный подтип встречается в 20 % случаев РМЖ).

ГЕН АТАКСИИ-ТЕЛЕАНГИЭКТАЗИИ *ATM* кодирует белок, который принадлежит к семейству PI3/PI4-киназ. Этот белок является важной киназой контрольной точки клеточного цикла. Белок функционирует как регулятор широкого спектра нижестоящих белков, включая белки-супрессоры опухолей p53 и BRCA1, киназу контрольной точки CHK2, белки контрольной точки RAD17 и RAD9 и белок репарации ДНК NBS1. Белок, кодируемый геном *ATM*, а также тесно связанная с ним киназа ATR считаются главными контроллерами сигнальных путей контрольной точки клеточного цикла, которые необходимы для реакции клеток на повреждение ДНК и для стабильности генома.

Один из наиболее распространенных генов восприимчивости к раку молочной железы — мутировавший ген *ATM*. Патогенные или вероятно патогенные варианты гена связаны с механизмом репарации двухцепочечных разрывов ДНК. Частота встречаемости мутаций данного гена у пациентов с раком молочной железы составляет 0,78%. Ген *ATM* экспрессирует белки, которые участвуют в репарации ДНК и регуляции клеточного цикла, что имеет решающее значение в ситуациях клеточного стресса и ответа на повреждение ДНК. Укороченные и миссенс-варианты *ATM* связаны с более высоким риском ER-положительного заболевания, чем ER-отрицательного. Миссенс-вариант *ATM* с.7271T>G (p.Val2424Gly) несет значительный риск РМЖ, сопоставимый с риском мутаций в гене *BRCA2*. У многих пациентов с раком молочной железы с мутациями *ATM* развивается тяжелая форма заболевания с высокой вероятностью метастазов в лимфатические узлы, более агрессивными опухолями и неблагоприятным прогнозом. Возможно, снижение уровня ATM в раковых клетках способно повысить чувствительность к таргетной терапии и sensibilizировать раковые клетки к препаратам на основе платины. Однако мутации *ATM* увеличивают риск рецидива опухоли после лучевой терапии.

ГЕН *PALB2* (партнер и локализатор *BRCA2*) является супрессором опухолей, который участвует в поддержании целостности генома. Этот белок связывается и колокализуется с белком раннего начала рака груди 2 (*BRCA2*) в нуклеоплазме и, вероятно, обеспечивает стабильную внутриядерную

локализацию и накопление BRCA2. Белок PALB2 критически важен для устранения двунитевых разрывов ДНК. Патогенные варианты приводят к риску развития РМЖ. Кумулятивный риск РМЖ у пациентов с мутацией *PALB2* зародышевой линии до 70 лет достигает 35%, а 10-летняя выживаемость ниже по сравнению с пациентами без мутаций *PALB2*. Ген *PALB2* — один из восьми генов, часто мутирующих при метастатическом раке молочной железы. Миссенс-вариант *PALB2* с.2586 + 58С>Т, был связан с 36%-ным увеличением риска РМЖ и входит в состав генетических панелей диагностики и скрининга РМЖ. У пациентов с раком молочной железы при наличии патогенных вариантов *PALB2* чаще наблюдается трижды негативный фенотип и более низкая выживаемость. В то же время, терапевтические схемы на основе платины, связанные с ингибиторами PARP, имеют большой потенциал у пациентов с вариантами зародышевой линии *PALB2*. Примерно 0,5% пациентов с РМЖ несут патологические мутации в гене *PALB2*. Хотя ранее *PALB2* считался геном риска РМЖ умеренной пенетрантности, последние данные свидетельствуют о том, что *PALB2* следует отнести к категории высокого риска. Результаты крупных популяционных исследований позволяют оценить риск развития РМЖ у носителей патологических мутаций *PALB2*, как сопоставимый с таковыми рисками у носителей патологических мутаций в гене *BRCA2*.

ГЕН *BRIP1* (взаимодействующая с раком молочной железы 1 геликаза 1) необходим для восстановления ДНК, поддержания стабильности генома. Комплекс белков BRIP1 и BRCA1 важен для нормальной функции репарации двухцепочечных разрывов. Мутировавший или сверхэкспрессированный *BRIP1* связан с началом рака молочной железы и является признанным кандидатом на прогрессирование опухоли. *BRIP1* расположен в длинном плече хромосомы 17 и кодирует белок, принадлежащий к семейству геликаз RecQ DEAN, который помогает восстанавливать поврежденную ДНК, взаимодействуя с BRCA1. Поэтому, если BRIP1 мутировал, то его белок не взаимодействует с BRCA1 и не может восстанавливать поврежденную ДНК. Таким образом, *BRIP1* играет жизненно важную роль в сохранении генетической информации клетки и действует как супрессор опухолей. Экспрессия *BRIP1* была значительно повышена в 13 типах рака, включая уротелиальную карциному мочевого пузыря (BLCA), инвазивную карциному молочной железы (BRCA), плоскоклеточную карциному шейки матки и аденокарциному эндометрия (CESC), аденокарциному толстой кишки (COAD), диффузную крупноклеточную В-клеточную лимфому лимфоидных неоплазм (DLBC), карциному пищевода (ESCA), мультиформную глиобластому (GBM), плоскоклеточную карциному легких (LUSC), аденокарциному прямой кишки (READ), аденокарциному желудка (STAD), тимому (THYM), эндометриальную карциному тела матки (UCEC) и карциносаркому матки (UCS). Ген *BRIP1* может быть геном, вовлеченным в онкогенез рака молочной железы в семьях, в которых нет мутаций *BRCA1/BRCA2*. Установлена связь редких миссенс-мутаций *BRIP1*, а также с двумя однонуклеотидными полиморфизмами *BRIP1* с раком молочной железы, что придает этому гену прогностическое значение. Сверхэкспрессия *BRIP1* связана со статусом метилирования промотора, определяя выживаемость пациентов с раком молочной железы. Следовательно, *BRIP1* может быть не только предсказательным молекулярным биомаркером развития и прогноза рака молочной железы, но также может рассматриваться в качестве терапевтической мишени.

ГЕН *CDH1* кодирует молекулу клеточной адгезии Е-кадгерин-1, которая подавляет распространение опухолевых клеток (метастазирование). Е-кадгерин-1 — это кальций-зависимая молекула межклеточной адгезии и белок-супрессор опухолей, который является единственным молекулярным дефектом зародышевой линии, связанным с наследственным диффузным раком желудка и дольковым раком молочной железы. Белок Е-кадгерин-1 играет важную роль в установлении и поддержании поляризованного и дифференцированного эпителия через межклеточные адгезивные комплексы. Эта молекула считается супрессором инвазии, и нарушение её регуляции часто обнаруживается в запущенных случаях некоторых эпителиальных карцином. Мутации в этом гене коррелируют с раком желудка, молочной железы, толстой кишки, щитовидной железы и яичников. Считается, что потеря функции этого гена способствует прогрессированию рака за счет увеличения пролиферации, инвазии и метастазирования. Эктодомен этого белка опосредует бактериальную адгезию к клеткам млекопитающих, а цитоплазматический домен необходим для интернализации. Снижение экспрессии Е-кадгерина сочетается с метастазами рака из-за потери клеточной адгезии, что приводит к увеличению подвижности клеток, что позволяет раковым клеткам пересекать базальную мембрану и проникать в близлежащие ткани. У пациентов с гиперметилированием промотора *CDH1* риск рака молочной железы повышен в 5,83 раза. Гиперметилирование промотора *CDH1* обычно повышено при HER2- и ER-отрицательном раке молочной железы, вне зависимости от статуса PR. Возможно, что *CDH1* является наиболее часто метилированным геном (90%) в случаях с метастазами в лимфатических узлах, что подтверждает связь гиперметилирования *CDH1* и метастазирования. Предполагают, что гиперметилирование *CDH1* можно использовать в качестве биомаркера

метастатического потенциала опухоли. Ген *CDH1* является потенциальной новой мишенью для таргетных препаратов, нацеленных на ингибирование метилирования ДНК.

ГЕН *BARD1* (BRCA1-ассоциированный кольцевой домен 1) представляет собой белок-партнер, связывающий BRCA, необходимый для восстановления повреждений ДНК, связанных с восприимчивостью к раку молочной железы. Взаимодействие BARD1/BRCA1 нарушается опухоленными заменами аминокислот в BRCA1. Следовательно, образование стабильного комплекса между этими белками может быть существенным фактором подавления опухоли BRCA1. Этот белок может быть целью онкогенных мутаций при раке молочной железы или яичников. Герминальные мутации в гене *BARD1* были значительно связаны с высоким риском TNBC - тип рака груди, который характеризуется отсутствием трех ключевых рецепторов: рецепторов для эстрогена, прогестерона и HER2. В исследованиях на лабораторных животных установлено, что частичная репрессия Bard1 с использованием антисмысловых РНК привела к развитию фенотипа ранних стадий злокачественности эпителиальных клеток молочной железы. Напротив, опухолевые клетки РМЖ могут экспрессировать повышенные уровни BARD1 и BRCA1, что приводит к резистентности к химиотерапии. Кроме того, более высокая экспрессия BARD1 и BRCA1 связана с худшим прогнозом для пациентов с ранней стадией РМЖ, особенно тех, кто получил лучевую терапию, что указывает на возможность использования ингибиторов PI3K для устранения химиорезистентности и радиорезистентности у ER-позитивных пациентов с РМЖ. Таким образом, ген *BARD1* может играть важную роль в патогенезе РМЖ и механизмах химиорезистентности. Предполагают, что уровень экспрессии *BARD1* при РМЖ может быть тесно связан с TNBC.

ГЕН *PIK3CA* Кодирован фосфатидилинозитол-3-киназу (PI3K), которая регулирует пролиферацию и апоптоз, а соматические мутации в *PIK3CA* могут активировать эти процессы. PI3K участвуют в нескольких клеточных процессах, таких как синтез белка, пролиферация клеток, выживание, гомеостаз глюкозы и репарация ДНК. Мутации *PIK3CA* представляют собой одно из наиболее распространенных типов мутаций при РМЖ. Активирующие мутации *PIK3CA* обнаруживаются примерно у 30–40% онкологических больных, вызывая гиперактивацию изоформы α -PI3K. Эти мутации связаны с химиорезистентностью и плохим прогнозом со сниженной общей выживаемостью. Кроме того, мутации гена *PIK3CA* при раке молочной железы являются онкогенными, играют важную роль в патогенезе и прогрессировании опухоли.

ГЕН *STK11 (LKB1)* Продукт гена представлен белком LKB1, который взаимодействует белком с p53 и регулирует специфические пути апоптоза, зависящие от p53. Белок LKB1 присутствует как в цитоплазме, так и в ядре живых клеток и транслоцируется в митохондрии во время апоптоза. Также наблюдается тесная связь функций белков LKB1 и PTEN. Некоторые точечные мутации *LKB1* нарушают взаимодействие с PTEN, что позволяет предположить, что потеря этого взаимодействия может способствовать повышению риска онкологических заболеваний. При этом, в раковых клетках наблюдается резкое снижение уровня экспрессии гена *STK11*. Снижение экспрессии гена наблюдается при многих онкологических заболеваниях: раке легкого, желудка, кишечника, поджелудочной железы и в том числе при раке груди. Установлена взаимосвязь эстроген-положительного и HER2-отрицательного РМЖ в семьях с мутациями в гене *STK11*.

ГЕН *RAD51D* принимает участие в механизмах репарации ДНК с использованием гомологичной рекомбинации. Он играет важную роль в поддержании стабильности генома и мутации гена могут быть связаны с канцерогенезом. Патогенные варианты *RAD51D* были обнаружены у пациентов с раком молочной железы с помощью тестирования генной панели. Присутствие таких патологических форм гена *RAD51D* в генотипе является причиной более чем 20% предполагаемого риска развития рака молочной железы в течение жизни. Существует корреляция между патогенными вариантами гена и трижды негативным раком молочной железы. Известно, что патологический вариант гена *RAD51D* c.270_271insTA (при нормальных формах генов *BRCA1/BRCA2*) является причиной TNBC высокой степени злокачественности с высоким риском метастазирования в лимфатические узлы. Патологическая форма *RAD51D* 270_271insTA наблюдался у женщины с анамнезом рака груди и яичников и зачастую выявляется при семейных формах рака груди и яичников. Выявлена еще одна онкогенная мутация гена (c.273del; p.Lys91fs). В настоящее время подтверждена связь между патогенными вариантами герминальных мутаций *RAD51D* и риском РМЖ. Также было показано, что патологические варианты *RAD51D* имеют выраженную связь с ER-отрицательным раком молочной железы и TNBC.

РЕКОМЕНДОВАННАЯ ЛИТЕРАТУРА К РАЗДЕЛУ 5. РАК ГРУДИ

Золотых М.А., Билялов А.И., Нестерова А.И. и соавт. Рак молочной железы: генетические факторы персонального риска//Современная онкология. 2023; 25 (2): 190–198

Смирнова Н.В., Елмуратов А.У Перспективы применения полигенного индекса риска (PRS) для оценки риска развития рака молочной железы//Вестник СурГУ. Медицина. 2019, 42(4): 74-78

Mavaddat N, Michailidou K., Dennis J. и соавт. Polygenic Risk Scores for Prediction of Breast Cancer and Breast Cancer Subtypes. *Am J Hum Genet.* 2019; 104(1): 21–34.
doi: [10.1016/j.ajhg.2018.11.002](https://doi.org/10.1016/j.ajhg.2018.11.002)

R. Graffeo, H.Q. Rana, F. Conforti и соавт., Moderate penetrance genes complicate genetic testing for breast cancer diagnosis: ATM, CHEK2, BARD1 and RAD51D. *Breast.* 2022; 65: 32–40.
doi: [10.1016/j.breast.2022.06.003](https://doi.org/10.1016/j.breast.2022.06.003)

Roberts E., Howell S., Evans D.G. Polygenic risk scores and breast cancer risk prediction. *Breast.* 2023; 67: 71–77.
doi: [10.1016/j.breast.2023.01.003](https://doi.org/10.1016/j.breast.2023.01.003)

Lyvia Neves Rebello Alves, Débora Dummer Meira, Luiza Poppe Meriguetti и соавт. Biomarkers in Breast Cancer: An Old Story with a New End. *Genes (Basel).* 2023; 14(7): 1364.
doi: [10.3390/genes14071364](https://doi.org/10.3390/genes14071364)

Karami F., Mehdipour P. A Comprehensive Focus on Global Spectrum of *BRCA1* and *BRCA2* Mutations in Breast Cancer. *Biomed Res Int.* 2013; 2013: 928562. doi: [10.1155/2013/928562](https://doi.org/10.1155/2013/928562)

Polyak K. Breast cancer: origins and evolution. *J Clin Invest.* 2007; 117(11): 3155–3163. doi: [10.1172/JCI33295](https://doi.org/10.1172/JCI33295)

Sun Y.-S., Zhao Z., Yang Z. и соавт. Risk Factors and Preventions of Breast Cancer. *Int J Biol Sci.* 2017; 13(11): 1387–1397.
doi: [10.7150/ijbs.21635](https://doi.org/10.7150/ijbs.21635)

Schon K., Tischkowitz M. Clinical implications of germline mutations in breast cancer: *TP53*. *Breast Cancer Res Treat.* 2018; 167(2): 417–423.
doi: [10.1007/s10549-017-4531-y](https://doi.org/10.1007/s10549-017-4531-y)

Waarts M.R. и соавт. Targeting mutations in cancer. *J Clin Invest.* 2022; 132(8): e154943.
doi: [10.1172/JCI154943](https://doi.org/10.1172/JCI154943)

Brett J.O. и соавт. ESR1 mutation as an emerging clinical biomarker in metastatic hormone receptor-positive breast cancer. *Breast Cancer Res.* 2021; 23: 85.
doi: [10.1186/s13058-021-01462-3](https://doi.org/10.1186/s13058-021-01462-3)

Dustin D., Gu G, Fuqua S.A.W. *ESR1* Mutations in Breast Cancer. *Cancer.* 2019; 125(21): 3714–3728.
doi: [10.1002/cncr.32345](https://doi.org/10.1002/cncr.32345)

Silwal-Pandit L., Langerød A., Børresen-Dale A.-L. *TP53* Mutations in Breast and Ovarian Cancer. *Cold Spring Harb Perspect Med.* 2017; 7(1): a026252.
doi: [10.1101/cshperspect.a026252](https://doi.org/10.1101/cshperspect.a026252)

Kuusisto K.M, Bebel A., Vihinen M. и соавт. Screening for *BRCA1*, *BRCA2*, *CHEK2*, *PALB2*, *BRIP1*, *RAD50*, and *CDH1* mutations in high-risk Finnish *BRCA1/2*-founder mutation-negative breast and/or ovarian cancer individuals. *Breast Cancer Res.* 2011; 13(1): R20.
doi: [10.1186/bcr2832](https://doi.org/10.1186/bcr2832)

Yoshimura A., Imoto I., Iwata H. Functions of Breast Cancer Predisposition Genes: Implications for Clinical Management. *Int J Mol Sci.* 2022; 23(13): 7481.
doi: [10.3390/ijms23137481](https://doi.org/10.3390/ijms23137481)

Long G., Hu K., Zhang X. и соавт. Spectrum of *BRCA1* interacting helicase 1 aberrations and potential prognostic and therapeutic implication: a pan cancer analysis. *Sci Rep.* 2023; 13: 4435.
doi: [10.1038/s41598-023-31109-6](https://doi.org/10.1038/s41598-023-31109-6)

Mahdi K.M., Nassiri M.R., Nasiri K. Hereditary genes and SNPs associated with breast cancer. Asian Pac J Cancer Prev.2013;14(6):3403-409.
doi: 10.7314/apjcp.2013.14.6.3403.

Hasson S.P., Menes T., Sonnenblick A. Comparison of Patient Susceptibility Genes Across Breast Cancer: Implications for Prognosis and Therapeutic Outcomes. Pharmgenomics Pers Med. 2020; 13: 227–238.
doi: [10.2147/PGPM.S233485](https://doi.org/10.2147/PGPM.S233485)

Corso G., Figueiredo J., De Angelis S.P. и соавт. E-cadherin deregulation in breast cancer. J Cell Mol Med. 2020; 24(11): 5930–5936.
doi: [10.1111/jcmm.15140](https://doi.org/10.1111/jcmm.15140)

РАЗДЕЛ 6. ГЕНЕТИКА ПИТАНИЯ

НУТРИГЕНЕТИКА, НУТРИГЕНОМИКА И ПИЩЕВАЯ ЭПИГЕНЕТИКА

Нутригенетика и нутригеномика - наука о роли индивидуальных генетических вариаций в реакции организма на питательные вещества и роли диеты и пищевых биологически активных веществ в экспрессии генов соответственно.

Подчеркнем **отличие между терминами нутригеномика и нутригенетика**, поскольку, хотя эти термины тесно связаны, они не являются взаимозаменяемыми.

Нутригенетика изучает модифицирующие эффекты герминальных (или соматических в случае онкопатологии) мутаций в генах, связанных с эффективностью питания, усвоения микроэлементов, а также влияние диеты на здоровье. Нутригенетика оценивает роль вариантов генов в реакциях организма на поступление питательных веществ. Нутригенетика и нутригеномика связаны между собой, тем не менее данные термины не являются взаимозаменяемыми. Основное отличие в том, что **нутригенетика** исследует влияние наследственных генетических вариантов на усвоение и метаболизм микронутриентов. В то время, как **нутригеномика** изучает взаимосвязь между геномом и диетой с указанием на влияние рациона питания на метаболические, протеомные, транскрипционные и трансляционные изменения. Исследования нутригенетики позволили идентифицировать генетические варианты, связанные с восприимчивостью к болезням, посредством взаимодействия со специфическими диетическими факторами. В частности, варианты генов, которые принимают участие в метаболических путях, влияют на потребление и использование различных микронутриентов. В этом смысле нутригенетические исследования могут быть использованы для прогнозирования риска различных полигенных заболеваний. Поэтому используя персонализированное питание эти заболевания можно предотвратить или эффективнее корректировать. Предполагается изучение всех указанных эффектов независимо или комплексно в целях диагностики или прогноза течения заболевания. Как уже было сказано, нутригенетика охватывает проблему влияния генетических вариаций на пищевые реакции и функцию питательных веществ. Тем не менее, из числа выше указанных генетических маркеров нутригенетики актуальны **только нарушения структуры молекулы ДНК, которые можно рассматривать, как истинный маркер патологии**. К нарушениям структуры ДНК следует отнести: (1.) повреждение отдельных оснований (например, присоединение гидроксильного радикала к гуанину, вызванное окислительным стрессом); (2.) утрата нуклеотидом азотистого основания (абазические участки в последовательности ДНК); (3.) разрывы цепей ДНК; (4.) укорочение теломера; (5.) делеции или утрата хромосом; (6.) повреждение митохондриальной ДНК (делеции или повреждение оснований в кольцевой митохондриальной ДНК). Стоит отметить, что важны не только варианты первичной структуры ДНК внутри одного гена в форме SNP, но и динамические мутации, а также варианты числа копий генов. В частности выявлена связь между вариантами числа копий (CNV) для небольших участков генома и риском метаболических заболеваний. Примерами таких мутаций являются: 1) варианты числа копий гена рецептора лептина связаны с метаболическими признаками и риском сахарного диабета 2 типа; 2) меньшее число копий гена альфа 1А амилазы слюны было связано с предрасположенностью к ожирению; 3) индел пентануклеотида (СТТТА) в 3'-нетранслируемой области гена рецептора лептина была связана с риском сахарного диабета 2 типа.

Область нутригеномики охватывает несколько дисциплин и включает в себя эффекты рациона питания на стабильность генома (повреждение ДНК на молекулярном и хромосомном уровне), изменения эпигенома (метилирование ДНК), транскрипцию мРНК и микроРНК (транскриптомика), экспрессию белков (протеомика) и изменения метаболитов (метабономика). Нутригеномика новая область молекулярной биологии, в которой передовые геномные инструменты используются для анализа влияния питательных веществ на геном и экспрессию генов, а также влияния генетических вариантов на потребление питательных веществ. Термин «нутригеномика» был создан для описания взаимодействия между питательными веществами и генами. Базовые вопросы, которые призвана рассматривать нутригеномика, обусловлены тем, что, во-первых, существует богатое индивидуальное и межэтническое разнообразие, оказывающее влияние на биодоступность и метаболизм питательных веществ. Во-вторых, резкие отличия между людьми, определяющие доступность и ассортимент продуктов питания в зависимости от культурных, экономических, географических и вкусовых различий. В-третьих, несбалансированное питание (дефицит или избыток питательных веществ) порождает неблагоприятную для человека последовательность событий, определяющих нарушения экспрессии генов, нестабильность генома, интенсификацию мутационного процесса, что в финале может привести к проявлению патологического фенотипа на различных этапах жизни.

Основные принципы науки нутригеномики:

1. Питание может оказывать влияние на состояние здоровья, напрямую влияя на экспрессию генов через состояние обменных процессов в организме, оказывая влияние на интенсивность генных и хромосомных мутаций. Такие события способны изменять дозу генов и их экспрессию.

2. Воздействие потребляемых продуктов питания на здоровье человека в значительной степени определяется генетическими вариантами, которые изменяют усвоение и метаболизм питательных веществ. Этот фактор на молекулярном уровне регулируется взаимодействием ферментов в обменных процессах и кофакторами питательных веществ.

3. Обеспечить устойчивые показатели здоровья можно в том случае, если рацион питания подбирается индивидуально с учетом, как унаследованных, так и приобретенных генетических характеристик в зависимости от периода онтогенеза, пищевых предпочтений и состояния здоровья.

Рекомендуемые современной медициной суточные нормы потребления питательных веществ, в целом основаны на различных метаболических исследованиях, они не оптимизированы для генетических подгрупп. Эти подходы в выработке рекомендаций могут существенно отличаться. Конечная цель состоит в том, чтобы (1) сопоставить нутриом (совокупность потребляемых продуктов питания) с текущим статусом генома, чтобы поддержание генома, экспрессия генов, метаболизм и функция клеток могли происходить нормально в гомеостатически сбалансированном режиме. А также (2) создать предпосылки для более благоприятных тенденций здоровья популяции, на основе формирования рекомендаций по персонализированному питанию.

Пищевая эпигенетика. Нутриэпигенетика — направление, изучающее роль пищевых продуктов в эпигенетических изменениях. Эпигенетика включает в себя, как обратимые, так и наследуемые изменения экспрессии генов не связанные с изменениями кодирующей последовательности ДНК. В данном случае эпигенетические изменения могут оказывать существенное влияние на патогенез, методы коррекции и лечение целого ряда полигенных заболеваний. Например, противовоспалительные эффекты диеты рациона питания, обогащенного полиненасыщенными жирными кислотами обусловлены ингибирующим гиперметилированием промоторов провоспалительных генов ферментами ДНК-метилтрансферазами. Помимо этого, высокое потребление полиненасыщенных жирных кислот положительно регулирует экспрессию определенных микроРНК, которые ингибируют липогенные и онкогенные гены. Вещество растительного происхождения куркумин, помимо того, что он является мощным антиоксидантом, также напрямую вовлечено в эпигенетические механизмы. В частности, как эпигенетический регулятор, куркумин оказывает защитное действие в отношении сердечной недостаточности и поражения печени, через систему ферментов метилирования ДНК и модификации гистонов. Эти данные свидетельствуют о том, что введение определенных пищевых соединений в рацион человека, которые модулируют эпигенетические процессы, может быть эффективной стратегией нацеленной на снижение распространенности хронических заболеваний. Подчеркнем, что нарушение эпигенетических механизмов может быть причиной развития различных хронических заболеваний и их прогрессирования. В частности непростые взаимодействия между питательными веществами и эпигенетическими механизмами могут принимать участие в патогенезе ожирения, сахарного диабета, дислипидемии, сердечно-сосудистых заболеваний, неалкогольной жировой болезни печени и онкопатологий. В частности, рацион питания основанный на избыточном потреблении животных жиров и легко доступных углеводов, может сопровождаться аномальными процессами метилирования генов нейропептидов, которые контролируют поведенческие реакции, определяющие потребление пищи и могут быть вовлечены в развитие ожирения и СД2. Помимо этого, недостаток в организме некоторых микроэлементов и витаминов, таких как витамин А, витамины группы В, селен, калий и железо, связан с гиперметилированием генов-супрессоров опухолей, которые играют решающую роль в развитии онкологии.

ГЕНЕТИЧЕСКОЕ ОБОСНОВАНИЕ ПЕРСОНАЛИЗИРОВАННОГО ПИТАНИЯ

Персонализированное питание – принципиально новый подход в диетологии, при котором индивидуальные генетические профили определяют индивидуальные рекомендации по рациону питания. Этот подход способствует поддержанию активного долголетия, а также позволяет эффективно корректировать хронические заболевания.

Явление полиморфизма генов человека объясняется различиями между людьми на уровне первичной структуры ДНК. Эти отличия могут вносить известный вклад в разнообразие фенотипов, наблюдаемое в популяциях человека. Типы вариантов генов включают в себя SNP, вставки и делеции (индели), вариации числа копий (CNV) и более сложные структурные изменения.

Полиморфизм генов оказывает существенное влияние на метаболизм питательных веществ. Например, варианты гена метилентетрагидрофолатредуктазы (*MTHFR*) могут влиять на метаболизм

фолата. Такие изменения могут быть связаны с сердечно-сосудистыми заболеваниями и диабетом. Люди с определенными SNP в этом гене могут обладать особенностями в реакции организма на потребление фолата, которые можно корректировать индивидуальным подбором добавок фолата. В частности, у гомозигот по мутации (TT) в гене *MTHFR* (полиморфизм *MTHFR* Ala222Val, C>T) наблюдается повышенная потребность в фолате из-за нарушений метаболизма этого вещества. Таким людям может потребоваться более высокое потребление фолата, чем это предлагают существующие диетические нормы, чтобы снизить восприимчивость к заболеваниям, вызванным дефицитом фолата. Кроме того, люди с синдромом Дауна (обладают дополнительной копией хромосомы 21) демонстрируют более высокую потребность в цинке и фолате из-за повышенной экспрессии цистатионин-β-синтазы, ключевого фермента в метаболизме фолата. Наряду с этим, достаточно широко распространенные варианты генов, регулирующих метаболизм гомоцистеина, такие как *MTHFR* (метилентетрагидрофолатредуктаза) и *MTR* (метионинсинтаза), ассоциируются с повышенным риском рака молочной железы у лиц с недостаточным потреблением витамина B6, витамина B12 и фолата.

На метаболизм селена также могут влиять генетические вариации. Некоторые SNP влияют на экспрессию белка или активность селенопротеина в ответ на добавку селена. Генетические вариации в гене бета-каротиноксигеназы 1 (*BCMO1*), который имеет решающее значение для метаболизма бета-каротина, могут приводить к изменчивости уровней каротиноидов в плазме крови и могут иметь важное клиническое значение. Например, в случае развития стеатоза печени независимо от диетического витамина А. Добавление витамина А в рацион питания является основным методом коррекции дефицита данного витамина в организме. Снижения уровня распространенности гиповитаминоза А в популяциях, знание вариантов генов, связанных с метаболизмом витамина А у человека, помогает проводить лечение более эффективно и безопасно. Варианты генов, принимающие участие в метаболизме липидов: белок-переносчик эфиров холестерина, липопротеинлипаза, рецептор липопротеинов низкой плотности и аполипопротеин Е, могут способствовать риску ишемической болезни сердца и сахарному диабету 2 типа. SNP в генах, связанных с метаболизмом холина и фолиевой кислоты, могут определять диетическую потребность в холине, который имеет решающее значение для функции печени и развития плода. Приведенные выше примеры акцентируют важность учета генетических факторов при составлении рациона питания и способствуют корректной оптимизации результатов в отношении здоровья на основе индивидуальных генетических профилей.

Установлена роль полиморфизмов в гене *CYP1A2* (цитохром P450 1A2). Данный ген, который кодирует фермент занимающий центральное положение в метаболизме кофеина, что определяет его возможную роль в патогенезе сердечно-сосудистых заболеваниях. Было установлено, что высокое употребление кофеина может повысить риск сердечно-сосудистых заболеваний у людей, варианты генов которых кодируют фермент «медленного» метаболизма кофеина. Напротив, люди, имеющие генетические варианты, связанные с быстрым метаболизмом кофеина, защищены от последствий умеренного потребления кофеина.

Присутствие в генотипе замены G>A в гене *APOA1* (является основным белковым компонентом липопротеинов высокой плотности (ЛПВП) в плазме крови) приводит к тому, что у людей с аллелем А наблюдается повышенный уровень холестерина ЛПВП после повышенного потребления длинноцепочечных омега-3 полиненасыщенных жирных кислот (ПНЖК), в то время как у людей с генотипом GG этот эффект не проявляется.

Фактор риска SNP в гене сывоточной и глюкокортикоид-индуцируемой киназы 1 (*SGK1*) обусловлен тем, что данная киназа активирует калиевые, натриевые и хлоридные каналы, контролируя в том числе и почечную экскрецию натрия. Поэтому высокие уровни экспрессии этого гена могут способствовать таким состояниям, как гипертония и диабетическая нефропатия. Следовательно, полиморфизм данного гена является одной из причин повышенного кровяного давления при употреблении пищи с повышенным содержанием соли. Напротив, разработка персонализированных рационов питания в соответствии с выявленными генотипами *ACE* (ангиотензин I-превращающего фермента), могут содержать рекомендации по потреблению хлорида натрия, в сравнении со стандартными диетическими рекомендациями для данной популяции.

Однонуклеотидные полиморфизмы (SNP) в гене *VDR* (рецептор витамина D) оказывают на усвоение витамина D и могут быть связаны с предрасположенностью к остеопорозу у женщин в постменопаузе на фоне недостаточного потребления кальция.

ПРАКТИЧЕСКОЕ ПРИМЕНЕНИЕ ГЕНОВ МАРКЕРОВ В РЕКОМЕНДАЦИЯХ К ПОТРЕБЛЕНИЮ ПИЩЕВЫХ ПРОДУКТОВ

Рацион питания напрямую влияет на состояние здоровья человека. В частности, модель питания, основанная на преимущественном употреблении пищевых продуктов на основе высокой степени переработки зерновых и мяса, а также обогащенной сладостями и десертами (западная модель питания) формирует у человека генетический профиль, способствующий канцерогенезу и стимулирующий воспалительные реакции. Также, высокое потребление животного белка способствует повышенному риску рака толстой кишки. Однако, рационы питания с недостаточным содержанием белка увеличивают экспрессию генов глюконеогенеза печени, что может привести к нарушению толерантности к глюкозе. Высокие уровни потребления животных жиров формируют профиль экспрессии генов, характерный для непереносимости глюкозы, депонирования липидов в печени, индукции воспаления и повышенной экспрессии нейропептидов, что приводит к развитию ожирения. Напротив, употребление в пищу сбалансированных количеств белка, не переработанных круп, моно- и полиненасыщенных жирных кислот, фруктов и овощей препятствует выше указанным негативным эффектам. Высокие уровни потребления мононенасыщенных жирных кислот (оливковое масло) ассоциируется со снижением экспрессии генов воспаления и депонирования липидов. Диеты, основанные на высоком содержании полиненасыщенных жирных кислот, позитивно регулируют экспрессию генов нейропептидов, контролирующих энергетический гомеостаз.

Генетические аномалии обменных процессов вызваны мутациями в соответствующих генах, которые кодируют данные ферменты. Эти патологические варианты генов качественно изменяют потребность организма в химическом составе продуктов питания. Классическими примерами таких наследственных аномалий являются непереносимость лактозы и фенилкетонурия. Вариант T>C-13910 в гене лактазы (*LCT*) приводит к неустойчивости или отсутствию фермента лактазы в зрелом возрасте. В связи с этим люди, обладающие этим патологическим вариантом гена, не в состоянии усваивать лактозу. С другой стороны, фенилкетонурия является аутомомно-рецессивным заболеванием, вызванным мутациями в гене фенилаланингидроксилазы (*PAH*), основного печеночного фермента, который отвечает за превращение фенилаланина в тирозин.

Диетические модели, которые рекомендуют крупы, овощи и фрукты, а также ограничивают потребление насыщенных жиров продемонстрировали эффективность в снижении риска ожирения у людей с высокими показателями генетической предрасположенности к ожирению, особенно у людей с аллелями риска гена *FTO* rs9939609, rs1121980 и rs1421085. Данные варианты гена *FTO* требуют рациона питания, обогащенного крупами, овощами и фруктами при минимальном содержании животных жиров. Кроме того, вариант гена *FTO* (rs1421085) предрасполагает к повышенному потреблению белка, независимо от индекса массы тела.

Следовательно, полиморфизмы гена *FTO* представляют риски для людей, страдающих ожирением. Таким лицам рекомендуется диета со сниженным потреблением животных жиров и высокая физическая активность.

Установлено, что наличие аллеля (A) варианта гена связывающего Ca²⁺ белка 39 (*CAB39*) rs6722579 сопряжено с повышенной восприимчивостью к абдоминальному ожирению у тех, кто превышает рекомендуемые нормы потребления жиров. Тогда, как у людей, несущих аллель (T) гена карбоксипептидазы Q (*CPQ*) rs59465035, наблюдают пониженную восприимчивость к абдоминальному ожирению, особенно в случаях повышенного потребления витамина C.

Для лиц с дефицитом фолиевой кислоты (полиморфизм гена *MTHFR*) рекомендуется повышение потребления продуктов питания с высоким содержанием фолиевой кислоты (цитрусовые, бананы, бобовые, зелень, яйца, печень). Такой рацион питания приводит к отчетливому снижению риска сердечно-сосудистых заболеваний.

Полиморфизм гена *APOE* относится к людям с риском сердечно-сосудистых заболеваний. В этом случае рекомендуется снижение употребления животных жиров и увеличение содержания омега-3 жирных кислот в рационе питания.

Полиморфизм генов *TCF7L2* и *PPARG* сопряжен с риском диабета и требует ограничения потребления пищи с высоким содержанием углеводов.

ГЕНЕТИЧЕСКИЕ ОСНОВЫ КОНТРОЛЯ МАССЫ ТЕЛА

Ожирение является глобальной проблемой здравоохранения. Ожирение рассматривается, как предиктор опасных хронических заболеваний людей во всём мире: сахарный диабет 2 типа, сердечно-сосудистые заболевания, метаболический синдром, нейродегенеративные заболевания и некоторые виды онкопатологий. В качестве иллюстрации можно привести данные статистики за 2017–2018 гг. по состоянию данной проблемы в США, которые показывают, что 42,4% и 9,2% взрослых граждан страны страдают соответственно ожирением и тяжелой формой ожирения.

На первый взгляд, решение проблемы ожирения чрезвычайно простое – соблюдение баланса в организме между потребленными и потраченными калориями. Тем не менее, это не так. Ключевой особенностью ожирения является нарушение самих физиологических механизмов регуляции энергетического обмена. В этих событиях центральная роль принадлежит ЦНС и эндокринной системе, выполняющим важнейшие функции в регуляции аппетита, чувства насыщения и расхода энергии, при этом гипоталамус является главным регулятором этих процессов. Аркуатное (дугообразное) ядро гипоталамуса (ARC), или латеральный гипоталамус – центральное звено управления процессами насыщения пищей. Ядро ARC содержит нейроны, которые экспрессируют функционально антагонистические пептиды. Это орексигенные пептиды, стимулирующие аппетит: нейропептид Y (кодированный геном *NPY*) и агути-родственный пептид – антагонист рецепторов меланокортина-3 и меланокортина-4 (кодированный у человека геном *AGRP*), а также анорексигенные, подавляющие аппетит: проопиомеланокортин (кодированный геном *POMC*) и кокаин и амфетамин-регулируемый транскрипт (у человека кодированный геном *CARTPT*) пептиды. Сразу отметим, что полиморфизм указанных генов у человека может являться причиной нарушений обменных процессов, включая ожирение. К периферическому звену контроля пищевого поведения следует отнести сигналы, таких гормонов, как лептин, инсулин и гормоны кишечника (холецистокинин (кодированный геном *CCK*), грелин (кодированный геном *GHRL*), пептид YY (кодированный геном *PYY*), глюкагоноподобный пептид-1 (GLP1, синтезируемый из пептида проглюкагона в ходе посттрансляционного процессинга в эндокринных L-клетках кишечника) и ряда других физиологически активных пептидов. Интеграция гормональных и пищевых сигналов от системы кровообращения и периферических и центральных нейрональных сигналов окончательно осуществляется в ARC, формируя реакцию на степень пищевого насыщения.

Состояние компонентов обоих уровней контроля энергетического обмена, потребления и расхода энергии может изменяться под влиянием полиморфизма генов, эпигенетических модификаций, состояния микробиома, образа жизни, а также социальных и психологических факторов, в финале приводя к ожирению.

Следовательно, необходимо учитывать, что наследственная составляющая является лишь одним из множества аспектов, принимаемых в расчёт при составлении рациона питания. Кроме наследственного фактора на особенности индивидуальной реакции организма на продукты питания могут оказывать влияние и целый ряд других факторов: эпигенетические, микробиомные (например, микрофлора кишечника) и экологические.

Между тем, индивидуальная изменчивость показателя массы тела имеет достаточно сильную генетическую основу. Результаты, полученные с использованием близнецового метода и в ходе анализа семейной наследуемости данного признака, оценивают степень наследуемости массы тела в диапазоне 40–70%.

В настоящее время идентифицировано около 500 генетических локусов, коррелирующих с основными признаками ожирения, такие, как соотношение талии к бедрам или индекс массы тела. Генетический локус *FTO*, связанный с массой жировой ткани и ожирением, считается наиболее актуальным генетическим маркером, влияющим на массу тела.

Основные варианты генов, вовлеченные в формирование риска ожирения и связанных с ожирением расстройств, расположены в следующих генах: *ADIPOQ*, *FTO*, *LEPR*, *LEP*, *MC4R*, *INSIG2*, *PPARG*, *PCSK1*, *ADBR3*, *ADBR2*, *PPARγ*, *APOA1*, *GHRL*, *APOA5*, *FABP2*, *LIPC*, *MTNR1B*, *TCF7L2*, *CETP*, *GIPR*, *NPY*, *IRS1* и *PCSK1*.

Например, ген *FABP2* (белок, связывающий жирные кислоты 2), экспрессируется эпителием тонкого кишечника и принимает участие в процессах всасывания жиров. Варианты данного гена могут способствовать более высокой абсорбции жира и приводить к ожирению. Еще один ген *PPARG* (рецептор, активируемый пролифератором пероксисом-гамма) экспрессируется в адипоцитах и играет важную роль в их дифференцировке. В настоящее время установлена ассоциация полиморфизма гена *PPARG* с чувствительностью к инсулину и индексом массы тела. Кроме того, известно, что ген *TMEM18* (трансмембранный белок 18) может регулировать аппетит, массу тела и способствовать ожирению. В дополнение, установлено, что моногенная форма ожирения может быть вызвана снижением экспрессии гена *MC4R* (рецептор меланокортина-4).

ВЗАИМОДЕЙСТВИЕ ВАРИАНТОВ ГЕНОВ С ПИТАНИЕМ, ПРИВОДЯЩЕЕ К ОЖИРЕНИЮ

Интерес к исследованиям взаимодействия генетических вариантов с рационом питания и другими факторами среды, как, например, физическая активность, обусловлен разработкой способов профилактики ожирения. Поэтому в разных популяциях анализируется влияние полиморфизма маркерных генов и рациона питания (с низким или высоким содержанием моносахаров, богатых животными жирами, с недостаточным содержанием белка) на риски развития ожирения. При этом

учитывается не только химический состав пищи, уровень обогащения диеты углеводами или жирами, но и суточный ритм приема пищи.

Результаты, полученные в этом направлении, например, показали, что ограничение насыщенных жиров в рационе питания более эффективны среди людей с избыточным весом и ожирением с генотипом *IRS1* rs2943641 CC. Также установлено, что люди с избыточным весом и ожирением, несущие аллель T *PPM1K* rs1440581, могут эффективней добиваться потери веса при соблюдении низкоуглеводной диеты.

В тоже время, индивидуальные особенности в пищевом предпочтении тех или иных продуктов также связаны с рисками развития ожирения. В литературе предлагается список генов, влияющих на предпочтения в еде, включает: *FTO* (rs9939609), *MC4R* (rs17782313), *DRD2* (rs1800497). В данном случае под предпочтениями в еде понимается склонность к перееданию, определяемая генетическими полиморфизмами.

Ген *FTO* кодирует белок, который участвует в энергетическом обмене, окислительных реакциях и метаболизме жирных кислот. Возможно, белок принимает участие в деметилировании ДНК. Результаты популяционных исследований позволили установить высокую степень взаимосвязи полиморфизмов данного гена с индексом массы тела, риском ожирения и сахарным диабетом 2 типа. Один из SNP данного гена (rs9939609) связан с увеличением индекса массы тела и ожирением. Высокая экспрессия *FTO* связана с более высоким потреблением пищи и последующим увеличением массы тела и жира, при условии, что уровень расхода энергии и физической активности остается неизменным.

Ген *MC4R* кодирует белок, который является мембраносвязанным рецептором и членом семейства меланокортиновых рецепторов. Кодированный белок взаимодействует с адренокортикотропными гормонами и гормонами MSH (меланоцитстимулирующие гормоны), его эффекты опосредуются G-белками. Белок играет важную роль в энергетическом гомеостазе и регуляции пищевых предпочтений. Белок, в сочетании с мелантропином, контролирует чувство насыщения. Кроме того, ген *MC4R* играет ключевую роль в регуляции гомеостаза глюкозы. Полиморфизм гена (rs17782313) является причиной моногенного ожирения, наследуемого по аутосомно-доминантному типу.

ЛИТЕРАТУРА РЕКОМЕНДОВАННАЯ К РАЗДЕЛУ 6. ГЕНЕТИКА ПИТАНИЯ

Fenech M. и соавт. Nutrigenetics and Nutrigenomics: Viewpoints on the Current Status and Applications in Nutrition Research and Practice. J Nutrigenet Nutrigenomics. 2011; 4(2): 69–89.
doi: [10.1159/000327772](https://doi.org/10.1159/000327772)

Singar S. и соавт. Personalized Nutrition: Tailoring Dietary Recommendations through Genetic Insights. Nutrients. 2024; 16(16): 2673.
doi: [10.3390/nu16162673](https://doi.org/10.3390/nu16162673)

Suzuki M., Tomita M. Genetic Variations of Vitamin A-Absorption and Storage-Related Genes, and Their Potential Contribution to Vitamin A Deficiency Risks Among Different Ethnic Groups. Front Nutr. 2022; 9: 861619.
doi: [10.3389/fnut.2022.861619](https://doi.org/10.3389/fnut.2022.861619)

Kiani A.K. и соавт. Polymorphisms, diet and nutrigenomics. J Prev Med Hyg 2022;63(suppl.3):E125-E141.
<https://doi.org/10.15167/2421-4248/jpmh2022.63.2S3.2754>

Voruganti V.S. Precision Nutrition: Recent Advances in Obesity. Physiology (Bethesda). 2023; 38(1): 42–50.
doi: [10.1152/physiol.00014.2022](https://doi.org/10.1152/physiol.00014.2022)

Softic S. и соавт. Divergent effects of glucose and fructose on hepatic lipogenesis and insulin signaling. J Clin Invest. 2017; 127(11): 4059–4074.
doi: [10.1172/JCI94585](https://doi.org/10.1172/JCI94585)

Vesnina A. и соавт. Genes and Eating Preferences, Their Roles in Personalized Nutrition. Genes (Basel). 2020; 11(4): 357.
doi: [10.3390/genes11040357](https://doi.org/10.3390/genes11040357)

РАЗДЕЛ 7. ГЕНЕТИЧЕСКАЯ КАРТА СПОРТСМЕНА

ВВЕДЕНИЕ

Вероятность наследуемости статуса спортсмена составляет около 66%, что означает, что 66% дисперсии в спортивных результатах объясняется генетическими факторами, а оставшиеся 34% объясняются факторами окружающей среды. При этом, наследуемость отдельных фенотипов, определяющих спортивные достижения, наиболее высока для состава волокон скелетных мышц: 45–99,5%, 49–56% для мышечной силы и 44–68% для кардиореспираторной составляющей. Помимо этого к уже указанным физиологическим фенотипам можно дополнить, что предполагаемая наследуемость риска спортивных травм составляет 60–80%. Эти данные свидетельствуют о том, что пока еще не выявленные генетические факторы могут играть решающую роль в отборе элитных спортсменов и прогнозировании спортивных результатов.

Картирование и исследование динамических молекулярных сетей, лежащих в основе физических упражнений человека, как правило, включает в себя подходы, сосредоточенные на генах, белках, пептидах, посттрансляционных модификациях, липидах и метаболитах в ряде образцов тканей и жидкостей организма. Эти метаболические аспекты не следует рассматривать как совершенно отдельные направления, поскольку интеграция по этим направлениям будет иметь решающее значение для разграничения молекулярных сетей и связей. Кроме того, каждое направление может быть далее разделено на отдельные области, включающие различные технологические подходы. В совокупности направления исследований геномики расширили наше понимание взаимодействия генетического фона и окружающей среды в спортивных результатах человека и реакции на физические упражнения. Технологические достижения геномики теперь предоставляют платформы для более дешевого и быстрого глубокого секвенирования человеческого генома. Предыдущие генетические исследования, в первую очередь основанные на меньших, гетерогенных когортах, оценивали гены-кандидаты и варианты. Исследования ассоциаций по всему геному (GWAS) выявили варианты генов (например, *ACTN3* и *ACE*) и однонуклеотидные полиморфизмы (SNP), потенциально влияющие на физическую работоспособность человека, пользу для здоровья от физических упражнений и предикторы реакции на физические упражнения. Молекулярные основы взаимодействия генов, лежащие в основе реакции человека на физические упражнения, сложны. Однако хорошо известно, что генетическая изменчивость влияет на взаимодействие генов и успехи в спорте. Усилия по картированию всего генома в более крупных когортах помогут раскрыть этот уровень сложности. Недавно были начаты многоцентровые исследовательские работы, такие как Консорциум проекта Athlome, чтобы помочь реализовать эту глобальную стратегию. Эти усилия существенно расширят каталог взаимодействий генов и упражнений и приведут к будущим исследованиям того, как эти гены связаны с изменениями во внешней среде, чтобы определить функциональные основы адаптации к тренировкам.

ГЕНЕТИЧЕСКИЕ ОСНОВЫ СПОРТИВНЫХ ДОСТИЖЕНИЙ

Спортивные результаты человека — это сложный фенотип, на который влияют экологические и генетические факторы, при этом большинство признаков, связанных с физическими упражнениями, являются полигенными по своей природе.

Черты физической работоспособности человека (РРТ) и адаптации к тренировкам являются результатом взаимодействия множества внутренних и внешних факторов, включая те, которые классифицируются как экологические, генетические, анатомические, физиологические и психологические.

Спортивная геномика — это научная дисциплина, изучающая организацию и функции генома у элитных спортсменов и направленная на разработку молекулярных методов, которые могут быть использованы в спортивной медицинской практике, персонализированных тренировках, назначении питания и профилактике травм и заболеваний, связанных с физическими упражнениями. Относительно недавно была сформирована новая дисциплина в этой области, известная как **кинезиогеномика**, объединяющая кинезиологию — научное изучение движения человека — и геномику. С точки зрения спортивной подготовки и физических упражнений к показателям, определяемым общими вариациями последовательности ДНК человека, относятся мышечная подготовленность, поведение во время упражнений, способность к кардиореспираторной, сердечно-сосудистой и метаболической адаптации к острым нагрузкам и восприимчивость к постоянным тренировкам. **Статус спортсмена сам по себе является в высокой степени наследуемой чертой, при этом от 66% до 70% дисперсии в статусе спортсмена объясняется аддитивными генетическими факторами**, а оставшаяся дисперсия обусловлена факторами окружающей среды, которые включают в себя тренировки, питание и

мотивацию спортсмена. Наследственность имеет очень важное значение для будущего спортсмена. Имеются доказательства того, что 60% вариации базовой аэробной емкости и сердечной функции (30–40% сердечного параметра), 70–90% вариации анаэробной мощности и 50–70% вариации мышечной силы можно объяснить генетическими факторами. Спортивным генетикам хорошо известно, что многие из основных биологических процессов, лежащих в основе спортивных результатов, находятся под генетическим влиянием, включая энергетический метаболизм скелетных мышц, митохондриальный биогенез, формирование мышц, костей и хрящей, оксигенация тканей, эритропоэз и ангиогенез. Количество исследований, направленных на выявление конкретных генов и их вариантов, способствующих выдающимся результатам, которые могли бы объяснить различия в физиологических возможностях между элитными спортсменами, в последнее время резко возросло. На основе полученных данных предложена возможность использования генетической информации для максимизации спортивных результатов и предотвращения травм, а также предлагаются практические разработки, имеющие целью конкретные диетические вмешательства, которые могут быть индивидуально адаптированы в соответствии с микробиомом спортсмена.

ГЕНЫ, ОПРЕДЕЛЯЮЩИЕ УСПЕХИ В СПОРТЕ

Эра спортивной геномики началась в начале 2000-х годов после расшифровки структуры ДНК человека и открытия первых полиморфизмов ДНК, связанных со спортивными результатами (вариации генов *ACE*, *ACTN3*, *AMPD1*, *PPARD* и *PPARGC1A*). Статус элитного спортсмена — это частично наследуемая черта, как и многие из основополагающих физиологических, антропометрических и психологических черт, которые способствуют элитным результатам. В последние годы масштабы генетического анализа конкретных генетических вариантов, которые способствуют чертам элитного спортсмена, резко возросли. Данный процесс стимулирует к попыткам использовать генетическую информацию в качестве инструмента для прогнозирования будущего статуса элитного спортсмена. Накопленные результаты оценивают степень генетической наследуемости, способствующей спортивным результатам. При этом генетические особенности, предрасполагающие к выносливости, демонстрируют диапазон 44%–68% роли наследуемого фактора. Тогда, как черты спортсмена, связанные с силой, представляют диапазон 48%–56% роли наследуемого фактора. Примером иллюстрации роли генов в эффективности физических упражнениях является ген *ACTN3*. Ген кодирует белок альфа-актинин-3, отвечающий за сильные мышечные сокращения. Вариант ген со стоп-кодоном (rs1815739; R577X) приводит к подавлению синтеза функционально активного белка (генотип XX). У людей, несущих этот генотип, отсутствует белок ACTN3. Напротив, аллель R (генотип RX или RR) отвечает за синтез функционального белка ACTN3. Известно, что спортсмены, занимающиеся силовыми видами спорта, чаще обладают функциональным белком по сравнению со спортсменами, занимающимися видами спорта на выносливость. Таким образом, аллель R был связан с видами спорта на силу (мощность), а генотип XX — с видами спорта на выносливость.

Необходимо подчеркнуть, что спортивная генетика не только открывает перспективы талантливому поколению, но и **проясняет целый ряд проблем, связанных со здоровьем популяции в целом. Следует понимать, что генетика спорта неразрывно связана с базовыми понятиями медицины.** В частности, результаты изучения генетической предрасположенности к тяжелой форме саркопении (возрастное снижение массы скелетных мышц, силы и скорости походки), показали что аллели генов, связанные с более высоким риском тяжелой саркопении тесно коррелируют с такими параметрами, как с усталость, интенсивность употребление алкоголя, курение, временем, проведенным за просмотром телевизора, и более высоким потреблением соли и обработанного мяса. Напротив, аллели, связанные с более низким риском тяжелой саркопении, были положительно связаны с уровнями сывороточного тестостерона, IGF1 и 25-гидроксивитамина D; ростом; физической активностью; а также показателями более здоровых пищевых привычек (потребление злаков, сыра, жирной рыбы, белка, фруктов и овощей).

Исследования, изучающие генетическую основу физической работоспособности человека (PPT), изначально были инициированы с использованием близнецового и генеалогического методов. Одним из первых достижений в генетике человеческого спорта было определение индексов наследуемости индивидуальных структурных и функциональных признаков. Исследования близнецов и семейный анализ показали, что такие признаки, такие как анаэробная мощность, максимальное потребление кислорода, работоспособность и состав типа мышечных волокон, частично наследуются. В настоящее время эти подходы подвергаются эволюционным изменениям, совершая переход от исследований генетической эпидемиологии к изучению биомаркеров последовательности ДНК, в результате чего были выявлены генетические маркеры, определяющие спортивные результаты и реакцию организма на тренировки.

Идентификация генетических маркеров позволила нацелиться на генетические варианты, которые были связаны со спортивными фенотипами, чтобы предположить их функциональную роль в физической работоспособности человека на основе влияния каждого варианта на определенные физиологические процессы. Например, варианты гена ангиотензинпревращающего фермента (*ACE*), который является известным регулятором артериального давления, как было показано многочисленными исследованиями, являются детерминантой потенциала выносливости у спортсменов. Другие примеры генов-кандидатов, которые были выделены и которые впоследствии были связаны с фенотипами, лежащими в основе спринтерских, силовых или выносливостных показателей, включают скелетно-мышечную изоформу α -актинина 3 (*ACTN3*), аденозинмонофосфатдеаминазу (*AMPD1*), мышечную креатинкиназу (*CKM*), инсулиноподобный фактор роста (*IGF1*) и семейство генов рецепторов, активируемых пролифераторами пероксисом (*PPAR*).

Из общего числа полиморфизмов ДНК, связанных со спортом, обнаруженных с 1998 по 2020 год, наиболее перспективными генетическими маркерами являются *HFE* rs1799945, *MCT1* rs1049434, *MYBPC3* rs1052373, *NFIA-AS2* rs1572312, *PPARA* rs4253778, *PPARGC1A* rs8192678 и *VEGFR2* rs1870377 **маркеры выносливости**; *ACTN3* rs1815739, *AMPD1* rs17602729, *CPNE5* rs3213537, *CKM* rs8111989 и *NOS3* rs2070744 **маркеры мощности**;

LRPPRC rs10186876, *MMS22L* rs9320823, *PHACTR1* rs6905419 и *PPARG* rs1801282 **маркеры прочности**; *COL1A1* rs1800012, *COL5A1* rs12722, *COL12A1* rs970547, *MMP1* rs1799750, *MMP3* rs679620 и *TIMP2* rs4789932 **маркеры повреждений мягких тканей**.

Благодаря открытиям генетических профилей элитных спортсменов были выявлены вариации последовательностей ДНК, которые связаны с предрасположенностью спортсменов к достижению успеха в различных видах спорта, их внутренней уязвимостью к травмам, связанным со спортом, и их индивидуальными потребностями в питании.

Поиск генов-кандидатов, который составляет большую часть исследований спортивной геномики на сегодняшний день, часто исследует варианты, первоначально выявленные в исследованиях на животных с использованием таких моделей, как нокаут гена, нокадаун гена и трансгенные мыши. Эти подходы выявили множество генов-кандидатов, впоследствии проанализированных на предмет их потенциального вклада в фенотипы работоспособности человека. Например, мутации в гене миостатина (*MSTN*) приводят к экстремальному росту скелетных мышц, что указывает на то, что миостатин является ключевым регулятором мышечной массы.

ВЫЯВЛЕНИЕ ГЕНОВ-КАНДИДАТОВ В СПОРТИВНОЙ ГЕНОМИКЕ

Исследования в области спортивной геномики базируются на выявлении генов-кандидатов. Т.е., осуществляется селекция генов и их вариантов, которые, связаны с интересующим параметром физической работоспособности человека (РРТ). Для этого проводят исследования ассоциации случай-контроль, чтобы определить, является ли генотип или аллель более распространенным в когорте спортсменов, чем в контрольной группе практически здоровых лиц неспортсменов из общей популяции. Подход на основе ассоциаций по всему геному считается наиболее эффективным дизайном исследования на сегодняшний день для выявления генетических маркеров, связанных со спортивными характеристиками. Действительно, применение этого подхода позволило обнаружить сотни однонуклеотидных полиморфизмов (SNP), напрямую или косвенно связанных с физическими упражнениями и спортом, таких как рост (12 111 SNP), безжировая масса конечностей (1059 SNP), уровень тестостерона (855 SNP), сила рукопожатия (170 SNP), саркопения (78 SNP) и быстрая ходьба (70 SNP). При этом важно подчеркнуть, что открытие 12 111 независимых общегеномных значимых SNP, связанных с ростом потребовало исследования 5,4 миллионов человек. Это также предполагает, что для обнаружения значительного количества генетических маркеров, связанных со статусом спортсмена, необходимы очень большие выборки спортсменов и неспортивных контрольных лиц, при этом некоторые предполагают, что эти маркеры можно использовать как часть стратегий выявления спортивных талантов.

Основная проблема спортивной генетики в том, что национальные выборки элитных спортсменов относительно невелики и чтобы преодолеть проблему небольших выборок элитных спортсменов, было предложено использовать общегеномные значимые маркеры фенотипов, связанных с физическими упражнениями (коррелирующие со спортивными чертами), обнаруженные в больших когортах нетренированных субъектов. Такой подход может значительно снизить риск получения ложноположительных результатов при обнаружении маркеров, связанных со статусом спортсмена, что признано ключевым ограничением исследований спортивной геномики, проводимых на ограниченных размерах выборки.

Другой подход спортивной геномики, доказавший свою эффективность в устранении возможности ложноположительных результатов, заключается в проведении исследований в двух или более независимых спортивных когортах (даже с небольшими или средними размерами выборки) с последующим метаанализом для количественной оценки общего эффекта полиморфизма на статус спортсмена или отдельную характеристику, связанную со спортом и физическими упражнениями. Однако в некоторых случаях экстраполяция результатов невозможна из-за уникальности полиморфизма для определенных популяций на основе их географического происхождения. Например, полиморфизм rs671 G/A гена альдегиддегидрогеназы 2 (*ALDH2*) был связан с силой у спортсменов и неспортсменов из японской популяции. Привлекает внимание, что неблагоприятный (связанный со сниженной силой) аллель rs671 A не присутствует у европейцев (частота встречаемости = 0%), но распространен в популяциях китайцев, японцев и вьетнамцев (15–25%).

Тем не менее, исследования случай-контроль остаются наиболее распространенным дизайном исследований в спортивной геномике. Как правило, такие исследования включают определение того, является ли один аллель последовательности ДНК (ген или некодирующая область ДНК) более распространенным в группе элитных спортсменов, чем в общей популяции. Тем самым подразумевается, что искомая аллель повышает работоспособность спортсмена.

Выявляемые таким образом полиморфизмы ДНК, в целом можно классифицировать как генетические маркеры, связанные с выносливостью (марафонцы, пловцы на 800–1500 м и т. д.), силой (тяжелотлеты, пауэрлифтеры и т. д.), мощностью (спринтеры, метатели, прыгуны, борцы, гимнасты и т. д.) или со статусом спортсмена в командных видах спорта. Связь генетического маркера с определенной спортивной дисциплиной можно объяснить с помощью разных механизмов. Например, спринтеры не только быстрее, но и выше, чем общая популяция и спортсмены, демонстрирующие выносливость. Поэтому гены, связанные как со скоростью, так и с ростом, можно также рассматривать как гены статуса спринтера. В частности, полиморфизмы генов *ACTN3* и *CPNE5* можно рассматривать, как наиболее значимые генетические маркеры, связанные со статусом спринтера. Ген *ACTN3* кодирует протеин семейства генов альфа-актин-связывающего белка. Кодированный белок в первую очередь экспрессируется в скелетных мышцах и функционирует как структурный компонент саркомерной Z-линии. Этот белок участвует в сшивании актина, содержащего тонкие нити. Аллельный полиморфизм в этом гене приводит как к кодирующим, так и к некодирующим вариантам; референсный генотип представляет собой кодирующий аллель. Привлекает внимание, что именно нефункциональный аллель этого гена связан со статусом элитного спортсмена. Ген *CPNE5* кодирует белок, относящийся к кальций-зависимым мембраносвязывающим белкам, которые могут регулировать молекулярные события на границе клеточной мембраны и цитоплазмы. Ген *CPNE5* является одним из нескольких генов, которые кодируют кальций-зависимый белок, содержащий два N-концевых домена C2 типа II и последовательность, подобную домену интегрина A, на C-конце. Для этого гена было обнаружено несколько альтернативно сплайсированных вариантов транскриптов, кодирующих различные изоформы.

Существует еще много связанных с физическими упражнениями черт, которые генетически детерминированы и важны для спортивных результатов, включая интеллект, время реакции, черты личности, мышечную массу, гибкость, уровень тестостерона и лактата, иммунологический статус, эффективность сжигания жира, сердечно-сосудистые и метаболические реакции и другие. Например, бойцы смешанных боевых искусств (MMA), в сравнении с неспортивными представителями контрольной группы, обладают генотипом воина (*COMT* rs4680 GG; Val/Val) подразумевающего высокую активность фермента катехол-О-метилтрансферазы на фоне низкого уровня дофамина. Такие генетические особенности обеспечивают преимущество, связанное с устойчивостью к стрессу.

Обычно гены-кандидаты идентифицируются на основе известных метаболических путей или экспериментов по экспрессии генов, с последующим анализом случай-контроль, выполняемым на одном генетическом маркере. Однако PPT является полигенным показателем. То есть на физическую работоспособность человека влияют многочисленные гены, многие из которых либо плохо охарактеризованы, либо неизвестны. Это, в свою очередь, указывает на то, что многие варианты генов-кандидатов еще не обнаружены. Данное конкретное ограничение можно преодолеть, анализируя целые геномы, а не отдельные варианты генов. Следовательно, полногеномные ассоциативные исследования (GWAS) стали высокопроизводительной альтернативой, которая позволяет обойти некоторые ограничения подхода в выявлении генов-кандидатов.

Несмотря на очевидные преимущества подхода GWAS, существенным ограничением этих исследований является то, что в некоторых странах и в определенных спортивных дисциплинах присутствует ограниченное количество когорт элитных спортсменов (недостаточность гомогенности спортивных когорт). С целью преодоления данной проблемы в 2015 году был создан Консорциум проекта ATHLOME, состоящий из 15 научных центров по всему миру, которые коллективно изучают данные генотипа и фенотипа элитных спортсменов для продвижения современных знаний о

молекулярной основе спортивных результатов. Основное внимание ATHLOME уделяется пониманию того, как спортсмены адаптируются к тренировкам. Проводится анализ основных причин травм, связанных с физическими упражнениями, и эпигенетических изменений, которые влияют на спортивные результаты и адаптацию. В конечном счете, цель Консорциума заключается в создании руководящих принципов, которые могут информировать о стратегиях для персонализированных тренировок, профилактики травм и обнаружения допинга.

Один из 15-ти центров изначально занимался анализом генетических данных спортсменов и лиц контрольной группы Восточной Европы (Беларусь, Литва, Россия и Украина; всего $n=8228$ спортсменов и $n=4121$ контрольная группа). Спортсмены сгруппированы в международные (включая участников Олимпиад и чемпионатов мира), национальные, региональные или местные категории. К ним относились биатлонисты, бегуны на длинные дистанции, велосипедисты, триатлонисты, байдарочники, гребцы, каноисты, современные пятиборцы, ориентировщики, лыжники, конькобежцы, шорт-трекеры, тяжелоатлеты, бодибилдеры, пауэрлифтеры, силачи, бегуны на короткие дистанции (≤ 400 м), пловцы на короткие дистанции (50–100 м), десятиборцы, семиборцы, единоборцы, легкоатлеты, бобслеисты, художественная и спортивная гимнастика, фигуристы, фехтовальщики и игроки в командные виды спорта с мячом.

Следует отметить, что полигенные профили спортсменов могут быть затронуты явлением, известным как эпистаз (или взаимодействия SNP–SNP). Эти взаимодействия происходят, когда комбинация двух или более SNP влияет на фенотип в большей степени или иным образом, чем эффект, наблюдаемый с отдельным геном. Например, эпистатический эффект двух SNP (*SNAP25* rs362584 и *ACTN3* rs1815739) был описан у гимнастов.

ГЕНЕТИКА И ИНДИВИДУАЛЬНЫЕ РАЗЛИЧИЯ В РЕАКЦИИ НА ТРЕНИРОВКИ

Анализ спортсменов «высокореагирующих» и «низкореагирующих» на стандартизированную тренировочную нагрузку может дать полезные сведения о механизмах адаптации к тренировкам и способствует разработке индивидуального графика тренировок.

Действительно, точные знания генетики атлета способствуют более ясному пониманию индивидуальной реакции организма на тренировки. К таким реакциям следует отнести: изменение индекса массы тела, химический состав тела, оптимизация активности инсулина для усвоения глюкозы во время нагрузок, благоприятный ответ на тренировки на выносливость для диастолического артериального давления (ДАД) и среднего артериального давления (САД) и улучшение VO_{2max} после тренировок на выносливость.

Важен также показатель восстановления между тренировками у выносливых спортсменов, который, возможно, определяется наследственностью. Например, предполагается, что полиморфизмы гена цитохром P450 2D6 (*CYP2D6*) и генов, кодирующих семейство G-глутатионтрансфераз, скорее всего, свидетельствуют о способности восстановления организма за счет удаления свободных радикалов или нейтрализации pH. Установлено, что частота встречаемости «благоприятных» вариантов в гене цитохрома P450 (т. е. с.506-1G>A полиморфизм в *CYP2D6*; ранее 1846G>A: rs3892097) и семейство глутатион-S-трансфераз (полиморфизм *GSTP* p.Ile105Val (rs1695)), была разной среди спортсменов с высокой выносливостью по сравнению с контрольной популяцией.

Дезадаптация спортсменов из-за дисбаланса между общим стрессом и восстановлением, может привести к развитию хронической усталости. В связи с этим необходимо учитывать, что реакция на тренировку может модулироваться временем и составом рациона питания, и, следовательно, факторы питания, как факторы среды, также могут потенциально способствовать этим индивидуальным вариациям.

СПОРТИВНАЯ НУТРИГЕНОМИКА

Питательные вещества один из важнейших факторов внешней среды, определяющий успехи в спорте. Фундаментальным аспектом, который может повлиять на связь между спортивными результатами и генетикой, является область нутригеномики и нутригенетики, которые представляют собой экспериментальные подходы, использующие технологии генетического тестирования для изучения роли индивидуальных генетических различий в адаптации реакции спортсмена на пищевые продукты. Генетические различия влияют на индивидуальные способности организма спортсмена абсорбировать, метаболизировать, использовать и выводить вещества, полученные с пищей. Нутригеномика и нутригенетика — это экспериментальные подходы, которые используют генетическую информацию и технологии генетического тестирования для изучения роли индивидуальных

генетических различий в изменении реакции спортсмена на питательные вещества и другие компоненты пищи. Люди по-разному реагируют на одни и те же потребляемые продукты, питательные вещества и добавки. Это утверждение справедливо для разных возрастов, этнических групп и уровней спортивного мастерства.

Персонализированное питание, основанное на знании генотипа человека, не является принципиально новым подходом к решению генетических особенностей человека. Существует несколько примеров редких (фенилкетонурия) и распространенных (непереносимость лактозы) наследственных патологий, которые нуждаются в коррекции рациона питания. Хотя генетические тесты хорошо зарекомендовали себя в клинических условиях, растет количество возможностей улучшить качество здорового образа жизни, повысить спортивные результаты при помощи тестирования генома, ориентированного на рацион питания.

Еще одной иллюстрацией такого подхода является чувствительность человека к кофеину. Кофеин, присутствующий в таких продуктах, как кофе, чай, какао и т.д. широко используется в спорте в качестве средства, повышающего работоспособность или повышающего эргогенность, часто в форме кофеинсодержащих таблеток и ли других формах. Установлено, что однонуклеотидный полиморфизм (SNP) гена *CYP1A2* -163A>C (rs762551) изменяет активность фермента *CYP1A2* и используется для идентификации людей как с «быстрым» или «медленным» метаболизмом кофеина. Например, в SNP *CYP1A2* rs726551 люди с генотипом AA (быстрые метаболизаторы) предрасположены к положительному ответу (повышение работоспособности) на кофеин. Люди с генотипом *CYP1A2* AC или CC, соответственно, не испытывают никакого эффекта или испытывают ухудшение физической работоспособности при употреблении кофеина.

Витамин А — жирорастворимый витамин, который присутствует в организме в разных биологически активных формах (ретинале и ретиноевой кислоте) и выполняет ключевую функцию, как в зрительном анализаторе, так и в иммунных реакциях организма. Витамин А выполняет разнообразные иммуномодулирующие функции; следовательно, дефицит витамина А связан как с иммунными дисфункциями в кишечнике, так и с несколькими системными иммунными расстройствами. Кроме того, витамин А рассматривается, как мощный антиоксидант, который обеспечивает защиту органов зрения от заболеваний и помогающим поддерживать зрение. Элитные спортсмены, скорее всего, обладают превосходными зрительными возможностями, основанными на их различных зрительных навыках, таких как контрастная чувствительность, динамическая острота, и стереоострота, что востребовано во многих видах спорта. Медленное время зрительно-моторной реакции было связано с риском травм опорно-двигательного аппарата в спортивных ситуациях, если существуют большие проблемы с эффективностью зрительного стимула и выполнением двигательного ответа. Эти зрительно-моторные навыки являются ключевыми факторами достижения спортивных результатов и, соответственно, требуют исключительного здоровья глаз. После абсорбции каротиноиды провитамина А легко преобразуются в витамин А ферментом *BCMO1*, экспрессируемым в энтероцитах кишечника. Один из вариантов гена *BCMO1* rs11645428 оказывает влияние на уровень циркулирующих каротиноидов в плазме, влияя на преобразование каротиноидов диетического провитамина А в активные формы витамина А в тонком кишечнике. По этой причине у людей с генотипом GG данные процессы неэффективны в этом преобразовании. Поэтому люди с таким генотипом могут подвергаться более высокому риску дефицита витамина А. Предполагают, что такие люди неэффективно реагируют на пищевой β-каротин, поэтому употребление необходимого количества уже готового витамина А с пищей может помочь обеспечить достаточный уровень циркулирующего активного витамина А для поддержания зрения и иммунитета.

Распространенность анемии, обусловленной нарушениями обмена железа, а также низкими уровнями фолиевой кислоты и витамина В12, по-видимому, выше у элитных спортсменов, чем в популяции в целом. Дефицит витамина В12 может иметь значительные негативные последствия для спортивных результатов. Наиболее распространенными симптомами этого расстройства являются усталость, слабость, одышка или сердцебиение. Железо входит в состав гемоглобина и миоглобина — белков, отвечающих за транспортировку и хранение кислорода. Витамин В₁₂ связан с эритропоэзом и аэробными процессами в организме. Мегалобластная анемия возникает из-за дефицита витамина В₁₂ и связана с повышенным уровнем гомоцистеина, а также приводит к общему чувству усталости и слабости. Гипергомоцистемия может являться результатом низкого потребления фолиевой кислоты и витамина В12. Эти обстоятельства могут способствовать увеличению риска нарушения функции скелетных мышц, включая мышечную слабость и регенерацию мышц. Метилентетрагидрофолатредуктаза (*MTHFR*) является ферментом, лимитирующим скорость в метильном цикле, и кодируется геном *MTHFR*. Полиморфизм C677T (rs1801133) в гене *MTHFR* был связан с низким уровнем фолата в сыворотке и эритроцитах, а также повышенным уровнем гомоцистеина в плазме, что является независимым фактором риска сердечно-сосудистых заболеваний.

Несколько исследований среди спортивных и неспортивных популяций показали, что люди с генотипом СТ или ТТ подвержены повышенному риску низкого уровня циркулирующего фолата. По сравнению с лицами с генотипом СС rs1801133, лица с генотипом ТТ и, возможно, генотипом СТ могут подвергаться большему риску гипергомоцистеинемии, хотя это может не быть причиной более низкой физической работоспособности. С другой стороны, футболисты и люди, ведущие малоподвижный образ жизни с генотипом СС, имеют более благоприятный прогноз и показатели работоспособности, по сравнению с носителями аллеля Т. Также аллель гена *FUT2* (rs602662) оказывает значительное влияние на уровень В₁₂ в плазме крови, при этом люди с генотипами GG или GA имеют наибольший риск низкого уровня витамина В₁₂ в плазме крови. Железодефицитная анемия - наиболее распространенный тип анемии среди спортсменов, которые имеют более высокую потребность в железе из-за повышенной потребности в эритропоэтине, за счет более высокой интенсивности и объемов тренировок. Спортсменки подвергаются особому риску дефицита железа из-за менструации и, как правило, более низкого общего потребления энергии или пищи по сравнению с мужчинами. На генетическом уровне содержание железа в плазме крови контролируется несколькими генами, такими, как *HFE*, *TMPRSS6*, *TFR2* и *TF*. Например, ген *HFE* принимает участие в регуляции усвоения железа в кишечнике. Известно, что полиморфизм этого гена может быть связан с повышением риска гемохроматоза (перегрузки организма железом). Избыток железа может быть токсичным для тканей и клеток, поскольку «свободное» железо реагирует с активными формами кислорода (ROS), такими как супероксид и перекиси водорода, или перекиси липидов, с образованием свободных радикалов. Свободные радикалы, в свою очередь, могут вызывать повреждение клеток и тканей (включая мышцы) и, в конечном итоге, приводить к гибели клеток. Биомаркеры железа, такие как ферритин и трансферрин, чаще встречаются у тех, кто генетически предрасположен к перегрузке железом на основе варианта гена *HFE*. Привлекает внимание тот факт, что спортсмены с редким генотипом *HFE* (rs1800562) AA, который связан с повышенным риском гемохроматоза, могут иметь генетическое преимущество для достижения успеха в спорте, если уровень железа находится на верхней границе нормы, но не является чрезмерным, чтобы вызвать повреждение тканей. Следует отметить, что риск перегрузки железом, чаще встречаются у спортсменов элитного уровня по сравнению с общей популяцией людей. Поэтому два SNP в гене *HFE* (rs1800562 и rs1799945), с одной стороны можно использовать в практической медицине, в целях выявления риска наследственного гемохроматоза, а с другой стороны генетический риск перегрузки железом может оказывать благоприятное влияние на выносливость у спортсменов, которым, тем не менее, необходимо избегать приема пищевых добавок железа, поскольку это может привести к неблагоприятным последствиям для здоровья. Еще три важных полиморфизма SNP: *TMPRSS6* (rs4820268), *TFR2* (rs7385804), *TF* (rs3811647) могут быть использованы для оценки генетического риска недостаточного уровня железа в организме. В данном случае указанные варианты генов принимают участие в регуляции экспрессии гепсидина, пептидного гормона, который контролирует усвоение железа. Анемия, вызванная дефицитом железа, ухудшает работоспособность за счет снижения способности гем-содержащих протеинов переносить кислород.

Витамин С выполняет функции водорастворимого антиоксиданта и кофермента при синтезе некоторых белков. Тем не менее, в медицине спорта пищевое потребление аскорбиновой кислоты, прежде всего, обусловлено ее свойствами антиоксиданта, которые помогают снизить образование свободных радикалов, вызванное интенсивной физической нагрузкой. Суммарное пищевое поступление витамина С, до 250 мг в день, вероятно, достаточно для снижения окислительного стресса, не оказывая отрицательного влияния на работоспособность. Наряду с этим принимается во внимание тот факт, что витамин С необходим для биосинтеза коллагена. Это говорит о том, что витамин С может играть роль в росте и восстановлении мышц. Поэтому присутствие витамина С в рационе питания во время тренировочного цикла улучшает синтез коллагена и может играть полезную роль в профилактике травм и ускорении восстановления опорно-двигательного аппарата, связок и/или сухожилий. Связь между диетическим витамином С и циркулирующим уровнем аскорбиновой кислоты зависит от генотипов глутатион-S-трансферазы *GSTT1* и транспортерами витамина С, *SVCT1* и *SVCT2*, которые кодируются *SLC23A1* и *SLC23A2*. В частности, у людей с генотипом *GSTT1 Del /Del*, риск развития гиповитаминоза С существенно выше, чем у людей с аллелем Ins.

В литературе имеются указания на роль витамина D, которые связаны с состоянием костной ткани, иммунитетом и восстановлением после тренировок. Генетические детерминанты циркулирующего 25-гидроксивитамина D (25(OH)D) могут влиять на каждый из этих факторов, тем самым влияя на работоспособность спортсмена. Только два гена *CYP2R1* (rs10741657) и *GC* (rs2282679) были значительно связаны со статусом витамина D. В частности, у спортсменов с генотипом GG или GA *CYP2R1* (rs10741657) почти в четыре раза повышается склонность к недостаточному уровню витамина D. Те, у кого был генотип GG гена *GC* (rs2282679), были значительно более склонны к низкому уровню витамина D по сравнению с теми, у кого был генотип ТТ.

Приведенный краткий обзор показывает, что понимание индивидуальной генетической изменчивости необходимо в определении потребления спортсменами витаминных добавок и микроэлементов.

ВАРИАНТЫ ГЕНОВ, ОТВЕЧАЮЩИЕ ЗА ВЫНОСЛИВОСТЬ СПОРТСМЕНА

Выносливость человека определяется многими факторами, включая типологию мышечных волокон, массу гемоглобина, митохондриальный биогенез, максимальный сердечный выброс и максимальную скорость потребления кислорода (VO_{2max}) и т. д. Способность человека выполнять физические упражнения на выносливость обусловлена состоянием аэробного метаболизма, показателями эффективности использования кислорода для производства энергии. Данные показатели в значительной степени поддерживаются улучшенной функцией митохондрий, экспрессией генов и активностью ферментов у элитных спортсменов. Выносливость элитных спортсменов подкрепляется постоянным мышечным сокращением в течение длительного периода времени, используя высокий VO_{2max} и другие резервы выносливости, сопряженные с митохондриальной дыхательной цепью. Способность выполнять упражнения на выносливость зависит от нескольких центральных факторов, связанных с мышечной и сердечно-сосудистой функцией. К ним относятся доля медленно сокращающихся волокон в скелетных мышцах и такие факторы, как максимальный сердечный выброс, который лежит в основе показателя максимальной скорости потребления кислорода (VO_{2max}).

Есть доказательства того, что на фенотипическое проявление указанных признаков оказывают существенное влияние генетические факторы. Возможно, именно генетические факторы объясняют до 70% изменчивости признаков, связанных с выносливостью. Обычно генетические маркеры, связанные со статусом выносливости спортсмена, определяются путем сравнения аллельных частот между выносливыми спортсменами (например, биатлонистами, шоссейными велогонщиками и т. д.) и контрольной группой людей, не занимающихся профессиональным спортом. Помимо этого, уровень восстановления между тренировками у выносливых спортсменов связан с генами, определяющих уровень метаболических процессов в печени. Такими как цитохром P450 2D6 (*CYP2D6*) и генами, кодирующими семейство G-глутатионтрансфераз.

Регулярные тренировки на выносливость способствуют приросту адаптации скелетной мускулатуры, обеспечивая метаболическую адаптацию скелетных мышц к упражнениям на выносливость. В качестве примера можно привести замедление темпов использования мышечного гликогена и глюкозы крови, более высокий вклад окисления жиров в энергетический обмен мышц во время физической нагрузки. Которые наблюдаются на фоне снижения образования лактата при высокой интенсивности упражнений, увеличение объема миокарда и увеличение циркулирующего объема крови. Все это способствует большему наполнению желудочков и, как следствие, большему ударному объему и сердечному выбросу. Все это в целом играет важную роль в значительном увеличении способности выполнять длительные напряженные упражнения на выносливость. Тем не менее, не все названные физиологические параметры могут быть определены генетикой.

Поэтому, для подтверждения наследственного характера полученных результатов исследований по схеме «случай-контроль», нацеленных на выявление генетической природы выносливости проводят дополнительно функциональные лабораторные исследования для определения связи между генотипами и физиологическими показателями. Такие измерения, используемые для дополнения геномных исследований, включают (но не ограничиваются) оценки VO_{2max} , объем форсированного выдоха за одну секунду (FEV1), долю медленно сокращающихся мышечных волокон, скорость восстановления, производительность бега на длинные дистанции, экономичность бега, порог лактата, уровни эритропоэтина и гемоглобина, количество эритроцитов, плотность капилляров, плотность митохондрий, жировой обмен и устойчивость к усталости.

Получил широкое признание тот факт, что большинство фенотипов, связанных с выносливостью, находятся под сильным генетическим влиянием. Некоторые из 40–50% дисперсии в доле медленно сокращающихся волокон в мышцах человека, по-видимому, генетически детерминированы. Кроме того, VO_{2max} и аэробная мощность также имеют высокую наследуемость. Возможно, на 50% изменения VO_{2max} при тренировках связано с врожденными факторами. Генетический компонент наблюдаемой изменчивости типов мышечных волокон у людей составляет ~ 40–50%, что указывает на то, что состав типа мышечных волокон в равной степени определяется генотипом и средой.

Подчеркивая связь спортивной медицины и практического здравоохранения, заметим, что VO_{2max} является эффективным прогностическим фактором заболеваемости и смертности по многим причинам и, в частности, от сердечно-сосудистых заболеваний. В связи с этим было проведено множество исследований генетических ассоциаций для других факторов риска сердечно-сосудистых

заболеваний, таких как неправильное питание, высокое кровяное давление и холестерин, стресс, курение и ожирение, что выявило новые терапевтические цели. Тем не менее, показатель $\text{VO}_{2\text{max}}$ имеет весомый генетический компонент, но исследования ассоциаций со здоровьем человека все еще отсутствуют.

Возможно, статус выносливого спортсмена остается наиболее изученной чертой, поскольку с генетической предрасположенностью в настоящее время этот показатель связывают, примерно, со 100 генетическими вариантами. Основанный на этих принципах поиск показал, что, по крайней мере, 41 из 114 зарегистрированных генетических маркеров может быть связан со статусом выносливости спортсмена на основе указанных критериев. В частности, показано, что элитные спортсмены с высокими показателями выносливости, обладатели генотипа TT в гене *ACTN3* или генота DD в гене *ACE* встречаются в когорте элитных спортсменов чаще, чем в контрольных популяциях нетренированных людей. При этом, наиболее перспективными генетическими маркерами для определения статуса выносливости у спортсменов являются аллели *AMPD1* rs17602729 C, *CDKN1A* rs236448 A, *HFE* rs1799945 G, *MYBPC3* rs1052373 G, *NFIA-AS2* rs1572312 C, *PPARA* rs4253778 G и *PPARGC1A* rs8192678 G.

ВАРИАНТЫ ГЕНОВ, ОТВЕЧАЮЩИЕ ЗА СТАТУС СПОРТСМЕНА В СИЛОВЫХ ВИДАХ СПОРТА

Силовые показатели спортсменов коррелируют с такими физиологическими параметрами, как уровень тестостерона, доля быстро сокращающихся мышечных волокон, мышечная масса, высота тела и пяточной кости, длина мышечного пучка, время реакции и другие признаки. Вероятно, наследуемость фенотипов, связанных с мощностью, колеблется в диапазоне от 49% до 86% в различных фенотипах. Генетический фактор оказывает существенное влияние на способность скелетных мышц производить силу с высокой скоростью, что имеет решающее значение для успеха в силовых и спринтерских видах спорта. Результаты в силовых видах спорта основаны на множестве факторов, которые, вносят значительный вклад в фенотип силы. Указанные генетические факторы выбраны на основе целого ряда физиологических параметров, определяющих успешность спортсмена-силовика: гипертрофию скелетных мышц (размер мышечных волокон), преобладание быстро сокращающихся мышечных волокон, больший угол перистости мышечных пучков, улучшенную неврологическую адаптацию, высокую гликолитическую емкость и повышенный уровень циркулирующего тестостерона. Имеются доказательства того, что спортсмены-силовики демонстрируют достаточно специфические транскриптомные, биохимические, антропометрические, физиологические и биомеханические параметры по сравнению со спортсменами, ориентированными на выносливость, а также в сравнении с контрольной группой испытуемых. Указанные межгрупповые отличия можно пояснить, с одной стороны, наличием факторов окружающей среды (тренировки, питание и т. д.), так и, с другой стороны, эффектами генетических факторов. Результаты проведенных исследований показывают, что существует сильная наследуемость признаков, связанных с мощностью и силой, где генетические факторы объясняют до 85% вариации максимальной изометрической, изотонической и изокинетической силы.

Наследуемость фенотипов, связанных с силой, составляет приблизительно от 49 до 86%. При этом, несколько силовых характеристик положительно связаны с показателями, включая циркулирующие уровни тестостерона, процент и площадь поперечного сечения быстро сокращающихся мышечных волокон, мышечную массу и силу, высоту тела и пяточной кости, длину мышечного пучка и время реакции, среди прочего.

Как правило, гены-маркеры, связанные со статусом спортсмена, определяются путем сравнения частоты встречаемости аллелей между спортсменами (например, бегунами на 100 м, толкателями ядра, армрестлерами и т. д.) и нетренированными людьми. Для подтверждения результатов исследований случай-контроль проводится анализ генотипа-фенотипа, основанный на измерении времени спринта, прыжковой производительности, размера мышечных волокон, типологии мышечных волокон, максимальной силы, скорости развития силы и уровня анаболических гормонов в крови, таких как тестостерон. По меньшей мере, 45 из 95 генетических маркеров, связанных со статусом силового атлета, соответствовали истинным критериям. Наиболее актуальными, из числа возможных генетических маркеров, связанных со статусом спортсмена, в настоящее время являются аллели *ACTN3* rs1815739 C, *AMPD1* rs17602729 C, *CDKN1A* rs236448 C, *CPNE5* rs3213537 G, *GALNTL6* rs558129 T, *IGF2* rs680 G, *IGSF3* rs699785 A, *NOS3* rs2070744 T и *TRHR* rs7832552 T.

Привлекает внимание тот факт, что фенотипы мышечной силы в общей популяции могут быть менее выражены, чем у спортсменов-силовиков. В этом случае спортсмены-силовики могут рассматриваться, как идеальная популяция для выявления геномных вариантов, которые определяют объем скелетных мышц. С другой стороны, знания, полученные в этой выборке спортсменов, расширяют наши представления о таких патологических явлениях, как саркопения, способствуя выработке

стратегии на снижение негативного влияния возрастного снижения мышечной массы. Как правило, маркерные гены спортсмена-силовика, можно выявить путем сравнения частоты встречаемости аллелей между спортсменами-силовиками и контрольной группой населения в целом. В целях верификации данных результатов проводятся исследования генотипа и фенотипа, измеряя силу хвата и изокINETической силы, производительность в пауэрлифтинге/тяжелой атлетике, а также оценивая острые и хронические реакции на силовые тренировки.

В настоящее время верифицированы 23 полиморфизма ДНК, которые отражают силу хвата руки и связаны со статусом силового спортсмена. Из этих SNP полиморфизмы *LRPPRC* rs10186876, *MMS22L* rs9320823 и *PHACTR1* rs6905419 также коррелируют с высокими результатами в соревнованиях по тяжелой атлетике. Возможно, 42 генетических маркера могут объективно отражать статус силового спортсмена. При этом, наиболее перспективными генетическими маркерами для определения статуса спортсмена-силовика являются аллели *ACTN3* rs1815739 C, *LRPPRC* rs10186876 A, *MMS22L* rs9320823 T, *PHACTR1* rs6905419 C и *PPARG* rs1801282 G, включая $AR \geq 21$ CAG-повторы.

ГЕНЕТИКА СПОРТИВНОЙ ТРАВМЫ

В современном спорте элитные спортсмены подвергаются высокой физической и эмоциональной нагрузке, обусловленной очень плотным графиком спортивных соревнований и неуклонным повышением объемов тренировок. Очевидно, что по мере повышения интенсивности соревновательного графика и объема тренировок усиливается риск развития травмы. Следует отметить, что вероятность травмы выше во время соревнований (72,2%), чем во время тренировок (21,8%). Как правило, около 80% спортивных травм затрагивают опорно-двигательный аппарат. При этом характер травмы может определяться видом состязаний, в которых принимает участие атлет. Например, риск травмы дельтовидной мышцы в год составляют от 12% до 19% у бейсболистов и от 23% до 38% у пловцов. У стайеров частота травм мышц нижних конечностей, связанных с тренировочным процессом, оценивается в 19–58%. В теннисе нередко регистрируются нарушения в суставах (29,5%), включая голеностопный сустав (20%), а также запястье (15,8%).

Идея о том, что генетические факторы частично объясняют восприимчивость к спортивным травмам, была предложена десять лет назад (Коллинз М., 2010), и было высказано предположение, что генетические полиморфизмы могут быть биомаркерами спортивных травм. Недавние исследования показали, что определенные генетические полиморфизмы связаны со спортивными травмами: *IGF2* (rs3213221), *CCL2* (rs2857656), *ELN* (rs2289360), *COL1A1* (rs1107946), *COL5A1* (rs12722), *COL22A1* (rs11784270) и rs6577958), *VDR* (rs7975232), *MCT1* (rs1049434), *ACTN3* (rs1815739), *HGF* (rs5745697 и rs1011694), *SOX15* (rs4227), *ESR1* (rs2234693), *CYP19A1* (rs936306), *ACE* (rs4341), *MMP3* (rs679620), *TN C* (rs2104772), *IL6* (rs1800795), *NOS3* (rs1799983), *HIF1A* (rs11549465) рассматриваются, как маркеры возможной мышечной травмы;

ELN (rs2289360), *COL1A1* (rs1800012), *COL5A1* (rs12722, rs71746744, rs16399, rs4919510), *MIR608* (rs4919510) возможной травмы связок или сухожилий; и *RANK* (rs3018362), *RNAKL* (rs1021188), *OPG* (rs4355801), *P2X7R* (rs1718119 и rs3751143), *VDR* (rs10735810, rs731236), *SOST* (rs1877632), *CYP19A1* (rs936306), *COL1A1* (rs1107946) возможного стрессового перелома.

Следовательно, эти генетические полиморфизмы могут способствовать восприимчивости к спортивным травмам. Однако качество доказательств может быть недостаточным из-за небольшого размера выборки элитных спортсменов.

В тоже время, использование методов молекулярной биологии позволило выявить генетические маркеры, связанные с риском определенных спортивных травм. Результаты, полученные в данном направлении исследований, позволяют надеяться, что выявленные генетические маркеры могут быть использованы спортсменами в целях персонализации их тренировок и диетических режимов.

К началу 2020-ых годов с помощью технологии НЖС было идентифицировано порядка 124 однонуклеотидных полиморфизмов (SNP), позволяющих прогнозировать возможный риск спортивной травмы, связанный с анатомическими или физиологическими особенностями организма данного атлета. Наиболее часто прогнозируемые риски связаны с: возможным разрывом передней крестообразной связки, травмой ахиллова сухожилия, низкой минеральной плотностью костей и стрессовым переломом, остеоартритом, дефицитом витаминов/минералов и признаком серповидноклеточной анемии. Следует особо подчеркнуть, что 91% из числа выявленных SNP ранее не привлекали внимание специалистов в области спортивной генетики.

Первая категория травм, разрыв передней крестообразной связки (ПКС), является одной из самых тяжелых травм, полученных при выполнении упражнений с многократным приземлением и резкими поворотами. Исследования генов-кандидатов, проведенные на спортсменах, обнаружили однонуклеотидные полиморфизмы в генах *COL1A1*, *COL5A1*, *COL12A1* и *MMP12*, которые связаны с разрывами ПКС. С другой стороны, известно, что динамические программы разминки, направленные на правильную технику приземления и резких движений, регулирование силы корпуса и подколенных сухожилий, эффективны для снижения риска травм передней крестообразной связки в том случае, если спортсмен предупрежден о существующем риске.

Вторая категория травм, травма ахиллова сухожилия, включает дегенеративные и болезненные состояния, которые поражают спортсменов в широком спектре видов спорта, включая до 18,5% бегунов. Исследования генов-кандидатов обнаружили варианты в генах *MMP3*, *COL5A1*, *COL1A1*, *CASP-8* и *GDF5*, которые связаны с патологиями ахиллова сухожилия. Вариант гена rs1800012 расположен в области промотора гена *COL1A1*, который кодирует коллаген I типа, основной белковый компонент сухожилий и связок. Редкий аллель Т увеличивает экспрессию гена, тогда как аллель G снижает экспрессию и связан с повышенным риском разрыва ахиллова сухожилия. Эффективные меры профилактики данного вида травмы основаны на применении растяжки для повышения гибкости, а также укреплении икроножной и камбаловидной мышц.

Третья категория травм, стрессовый перелом, является распространенной травмой, вызванной перенапряжением, и поражает до 20% женщин-спортсменок и начинающих спортсменов. В отличие от перелома, вызванного одним сильным ударом, стрессовые переломы — результат накопившихся микротравм от циклических нагрузок, например, при беге или прыжках. Минеральная плотность костной ткани - МПК является ведущим фактором, определяющим способность кости выдерживать нагрузку, а низкая МПК является важным фактором риска стрессовых переломов у спортсменов. При этом, наследственность определяет 50–85% дисперсии МПК. Активные молодые женщины, с семейным анамнезом остеопороза и низкой МПК, были почти в два раза более подвержены риску развития стрессового перелома.

К четвертой категории риска относится остеоартрит, который, с одной стороны, является основной причиной боли и инвалидности у пожилых людей, а с другой стороны, является наиболее распространенной формой артрита во всем мире. Высокая интенсивность спортивных тренировок и состязаний приводят к поражениям хрящей, которые в конечном итоге являются причиной раннего остеоартрита. В настоящее время идентифицированы пять значимых локусов по всему геному, ассоциированных с высоким риском остеоартрита и три локуса чуть ниже этого порога. Самая сильная ассоциация была выявлена на хромосоме 3 для rs6976, который находится в неравновесном сцеплении с rs11177. Этот SNP кодирует миссенс-полиморфизм в гене *GNL3*, кодирующем нуклеостемин. Уровни нуклеостемина были повышены в хондроцитах у пациентов с остеоартритом. Другие значимые локусы были на хромосоме 9 близко к *ASTN2*, хромосоме 6 между *FILIP1* и *SENP6*, хромосоме 12 близко к *KLHDC5* и *PTHLH* и в другом регионе хромосомы 12 близко к *CHST11*. Один из сигналов, близких к общегеномному значению, находился в гене *FTO*, который участвует в регуляции массы тела — сильного фактора риска остеоартрита. Данные полиморфизмы имеют аллельные коэффициенты риска между 1,1 и 1,2, тем не менее, все семь локусов имеют аддитивный аллельный эффект.

Пятая категория включает SNP, связанных с дефицитом в организме атлета витаминов и минералов. Интенсивная физическая нагрузка приводит к биохимическим адаптациям в мышцах, которые подвергают метаболические пути стрессу и увеличивают потребность в микроэлементах. Регулярные тренировки увеличивают потерю микроэлементов, тем самым требуя более интенсивного восполнения микроэлементов для поддержания и восстановления химического состава тела. Оптимальные уровни кальция и витамина D важны для поддержания здоровья костей, в то время как железо, цинк и комплекс витаминов В важны для стабильности гематологических показателей.

Завершающая шестая категория, серповидноклеточная анемия (гемоглобин S, rs334), подразумевает атлетов-гетерозигот по мутантному аллелю β-цепи гемоглобина. В этом случае под влиянием целого ряда неблагоприятных факторов: большая высота, сильная жара, обезвоживание и истощение эритроциты могут принимать патологический серповидный фенотип. В свою очередь, при повышенной физической нагрузке это может привести к таким опасным последствиям, как макрогематурии, инфаркту селезенки и рабдомиолизу, который может быть фатальным. Современные рекомендации предписывают спортсменам гетерозиготам по гену серповидноклеточной анемией придерживаться медленных и постепенных режимов подготовки, отдыхать и полноценно восстанавливаться в тренировочных циклах, немедленно прекращать физические нагрузки при появлении симптомов, поддерживать оптимальный уровень жидкости в организме, использовать дополнительный кислород по мере необходимости и немедленно обращаться за медицинской помощью при возникновении недомогания.

Кроме этого, ряд авторов выделяют в самостоятельную группу повреждение мышечной ткани. Повреждения мышц - серьезная проблема для элитных спортсменов, на этот вид спортивных травм приходится не менее 30% случаев и они являются причиной временной утраты физической активности спортсмена. Наибольшая доля мышечных травм происходит во время спортивных состязаний и составляет 10–55% всех травм. Повреждение мышц, как правило, происходит посредством двух основных механизмов: мышца подвергается внезапной прямой сжимающей силе, что приводит к ушибу, или мышца подвергается чрезмерной растягивающей силе, что приводит к повреждению мышечных волокон и возможному разрыву. К группе мышц, наиболее часто повреждаемых во время занятий спортом, относятся четырехглавая и икроножная мышцы. Эти мышцы проходят через два сустава и наиболее восприимчивы к силам ускорения и замедления. Повреждение мышц сопровождается временным ультраструктурным разрушением миофибрилл, потерей мышечной силы, отсроченной мышечной болезненностью, ограниченной подвижностью пораженной конечности и отеком. Нельзя исключить того факта, что существует высокая индивидуальная изменчивость спортсменов в ответе на повреждение мышц, вызванное физическими упражнениями. Появляется больше сообщений о том, что определенные варианты генов могут быть связаны с повреждением мышц, вызванным физическими упражнениями. У людей, страдающих от повреждения мышц в результате физических упражнений, были обнаружены полиморфизмы следующих генов: *ACTN3* (R577X, rs1815739), *TNF* (-308 G>A, rs1800629), *IL6* (-174 G>C, rs180795) и *IGF2* (G>A, rs680). Также, возможно, в этих процессах принимают участие и эпигенетические механизмы. В частности, приводятся данные о том, что острые физические нагрузки могут влиять на состояние процессов метилирования молекулы ДНК, приводя к гипометилированию промоторных областей генов *TFAM*, *PGC-1α*, *PDK4* и *MEF2A*.

МИТОХОНДРИАЛЬНАЯ ДНК И СПОРТИВНЫЕ ДОСТИЖЕНИЯ

Митохондриальная ДНК (мтДНК), двухцепочечная кольцевая ДНК, состоит из 16 569 пар оснований и содержит тринадцать генов, кодирующих белки, две рибосомальные РНК (рРНК) и двадцать две транспортные РНК (тРНК). Одной из важнейших функций митохондрий является выработка энергии. Аденозинтрифосфат (АТФ), необходимый для сокращения скелетных мышц во время упражнений, наиболее эффективно вырабатывается системой окислительного фосфорилирования (ОХРНОС) в митохондриях. Из-за своей роли в выработке клеточной энергии митохондрии были тщательно изучены специалистами медицины спорта. Известно, что тренировки на выносливость увеличивают содержание митохондрий в скелетных мышцах и приводят к повышению выносливости при выполнении упражнений за счет улучшения выработки АТФ.

Митохондриальная ДНК (мтДНК) была открыта в 1963 году. Она располагается в митохондриальном матриксе в виде двухцепочечной кольцевой молекулы. В одной митохондрии, как правило, присутствует от 2 до 10 копий мтДНК.

В мтДНК выявлено 37 генов, кодирующих 13 белковых субъединиц дыхательной цепи, 22 тРНК, а также две рРНК (12 S рРНК и 16 S рРНК).

Важной особенностью митохондриального генома является его уязвость к различным повреждениям. Поэтому риск мутаций мтДНК почти в десять раз выше, чем у ядерных генов. Причины высокой степени риска мутаций мтДНК обусловлены несколькими причинами. Во-первых, мтДНК не содержит защитные белки гистоны, во-вторых, эффективность процессов репарации несколько снижены, по сравнению с ядерной ДНК, в-третьих, мтДНК более уязвима для мутагенных эффектов свободных радикалов, образующихся в митохондриях при окислительном фосфорилировании, в-четвертых, мтДНК полностью лишена интронов и содержит очень мало избыточности. В результате делеция или точечная мутация в мтДНК может легко и быстро спровоцировать биохимический эффект.

Как уже отмечалось, тринадцать белков, кодируемых мтДНК непосредственно являются элементами ОХРНОС, и они необходимы для полноценного функционирования ОХРНОС. Поэтому вполне вероятно, что митохондриальные однонуклеотидные полиморфизмы (mtSNP) могут оказывать влияние на спортивные достижения.

Действительно, современные данные говорят о том, mtSNP m.8701A>G, m.10873T>C, m.12705C>T и m.15043G>A, достоверно связаны с реакцией организма на тренировки, оцениваемой по изменению порога лактата. Помимо этого, частота встречаемости mtSNP m.185G>A, m.228G>A, m.295C>T, m.462C>T, m.489T>C, m.4215T>C, m.10397A>G, m.11250A>G, m.13707G>A, m.14797T>C, m.15451C>A, m.16068C>T и m.16125T>C выше у испытуемых с наименьшими показателями $\dot{V}O_{2\max}$, в сравнении с людьми с максимально высокой реакцией.

В частности, m.152T>C, m.4343A>G, m.11215C>T, m.15518C>T и m.15874A>G были связаны со статусом элитного спортсмена в видах спорта на выносливость. Тогда как m.151C>T, m.204T>C,

m.4833A>G, m.5108T>C, m.5601C>T, m.7600G>A, m.9377A>G, m.13563A>G, m.14200T>C, m.14569G>A, m.15314G>A и m.16278C>T были связаны со статусом элитного спортсмена спринтера.

ЭПИГЕНЕТИКА В СПОРТИВНОЙ МЕДИЦИНЕ

Эпигенетика — быстро развивающаяся область в молекулярной физиологии физических упражнений, которая изучает изменения экспрессии генов при отсутствии изменений в последовательности ДНК. Эпигенетические модификации (например, метилирование ДНК, ацетилирование гистонов, некодирующие РНК) регулируют экспрессию генов (путем активации или «молчания» генов) и часто происходят в ответ на экологические или внешние стимулы, такие как специфический тренировочный стимул или диета, и вызывают специфические изменения в транскрипционном ответе. Эпигенетика, своего рода, стыковочный узел между индивидуальным генетическим статусом человека и внешней окружающей средой. Эпигенетические биомаркеры в спортивной медицине, прежде всего, следует понимать, как предикторы адаптивного ответа организма на упражнения и тренировки. Например, было показано, что тренировки могут приводить к гипометилированию ДНК в промоторных областях ключевых метаболических генов, регулируемых физическими упражнениями (например, *PPARGC1A*, *PPARD*), что сопровождается сопутствующей сверхэкспрессией этих генов. Таким образом, изменчивость фенотипической адаптации, вызванная физической активностью, может быть частично объяснена эпигенетическими факторами. Однако конкретные эпигенетические изменения, вызванные упражнениями, недостаточно изучены. Также было показано, что эпигенетические маркеры адаптации к тренировкам сохраняются в скелетных мышцах даже после периодов детренированности, что позволяет предположить, что эпигенетические механизмы лежат в основе внутренней способности человеческого организма «запоминать» адаптации к тренировкам. Недавно новый класс эпигенетических маркеров, известный как сигнатура конформации хромосомы, дал многообещающие результаты в медицинских исследованиях и может также использоваться для изучения того, участвуют ли изменения в конформации генома (организация генома в трехмерном ядерном пространстве) в регуляции реакции на упражнения.

Важной задачей для ученых остается определение типа, интенсивности и частоты упражнений, необходимых для поддержания эпигенетической памяти (например, на уровне метилирования ДНК) адаптации к упражнениям, которая может сохраняться в течение длительного периода времени. Если такие успехи будут достигнуты, возможно, эта информация может быть использована для оптимизации программ спортивных тренировок или улучшения процесса восстановления после травм. Физическая активность оказывает значительное влияние на паттерны метилирования ДНК на уровне структур мозга, таких как гипоталамус, гиппокамп и кора, подчеркивая роль упражнений и физической активности как эпигенетических регуляторов пластичности мозга и когнитивных процессов, противодействующих старению и нейродегенерации.

Еще одной важной составляющей эпигенетики спорта является вопрос взаимосвязи между нутригеномикой и нутригенетикой, а также спортивными результатами и генетикой. Нутригеномика и нутригенетика представляют собой экспериментальные подходы, использующие технологии генетического тестирования для изучения роли индивидуальных генетических различий в адаптации реакции спортсмена на пищевые факторы. Учет данных факторов, безусловно, важен при оценке влияния генетики на спортивные результаты, поскольку диетические и пищевые стратегии человека, как факторы среды, могут влиять на спортивные результаты. Персонализированное питание у элитных спортсменов на основе генетического профиля направлено на оптимизацию здоровья, течения обменных процессов и спортивных результатов. Спортивные диетологи успешно применяют углубленную проверку диетических рекомендаций во всех общих рекомендациях по питанию для населения, чтобы учесть различные спортивные группы.

Изменчивость фенотипической адаптации спортсмена, вызванная физической активностью, может быть частично объяснена эпигенетическими факторами. Однако конкретные эпигенетические изменения, вызванные упражнениями, недостаточно изучены. Важной задачей для ученых остается определение типа, интенсивности и частоты упражнений, необходимых для поддержания эпигенетической памяти (например, на уровне метилирования ДНК) адаптации к упражнениям, которая может сохраняться в течение длительного периода времени. Если такие успехи будут достигнуты, возможно, эта информация может быть использована для оптимизации программ спортивных тренировок или улучшения процесса восстановления после травм.

РЕКОМЕНДОВАННАЯ ЛИТЕРАТУРА К РАЗДЕЛУ 7. ГЕНЕТИЧЕСКАЯ КАРТА СПОРТСМЕНА

Генетический паспорт — основа индивидуальной и предиктивной медицины / Под ред. В. С. Баранова. — СПб.: Изд-во Н-Л, 2009. — 528 с.

Ahmetov I.I., Fedotovskaya O.N. Current Progress in Sports Genomics. *Adv Clin Chem*. 2015;70:247-314
doi: [10.1016/bs.acc.2015.03.003](https://doi.org/10.1016/bs.acc.2015.03.003)

Ahmetov I. и соавт. Chapter Five - Advances in sports genomics. *Advances in Clinical Chemistry*. 2022;107: 215-263

Ginevičienė V., Utkus A., Pranckevičienė E. и соавт. Perspectives in Sports Genomics. *Biomedicines*. 2022; 10(2): 298.
doi: [10.3390/biomedicines10020298](https://doi.org/10.3390/biomedicines10020298)

Tanisawa K., Wang G., Seto J. и соавт. Sport and exercise genomics: the FIMS 2019 consensus statement update. *Br J Sports Med*. 2020; 54(16): 969–975.
doi: [10.1136/bjsports-2019-101532](https://doi.org/10.1136/bjsports-2019-101532)

Hoffman N.J. Omics and Exercise: Global Approaches for Mapping Exercise Biological Networks. *Cold Spring Harb Perspect Med*. 2017; 7(10): a029884.
doi: [10.1101/cshperspect.a029884](https://doi.org/10.1101/cshperspect.a029884)

Konopka M.J. и соавт. Genetics and athletic performance: a systematic SWOT analysis of non-systematic reviews. *Front Genet*. 2023;14:1232987.
doi: [10.3389/fgene.2023.1232987](https://doi.org/10.3389/fgene.2023.1232987)

Pitsiladis Y.P., Tanaka M., Eynon N. и соавт. Athlome Project Consortium: a concerted effort to discover genomic and other “omic” markers of athletic performance. *Physiol Genomics*. 2016; 48(3): 183–190.
doi: [10.1152/physiolgenomics.00105.2015](https://doi.org/10.1152/physiolgenomics.00105.2015)

Semenova E.A., Hall E.C.R., Ahmetov I.I. Genes and Athletic Performance: The 2023 Update. *Genes (Basel)*. 2023; 14(6): 1235.
doi: [10.3390/genes14061235](https://doi.org/10.3390/genes14061235)

Varillas-Delgado D., Del Coso J., Gutiérrez-Hellín J. и соавт. Genetics and sports performance: the present and future in the identification of talent for sports based on DNA testing. *Eur J Appl Physiol*. 2022; 122(8): 1811–1830.
doi: [10.1007/s00421-022-04945-z](https://doi.org/10.1007/s00421-022-04945-z)

Kraemer W.J., Ratamess N.A., Hymer W.C. и соавт. Growth Hormone(s), Testosterone, Insulin-Like Growth Factors, and Cortisol: Roles and Integration for Cellular Development and Growth With Exercise. *Front Endocrinol (Lausanne)*. 2020; 11: 33.
doi: [10.3389/fendo.2020.00033](https://doi.org/10.3389/fendo.2020.00033)

Guest N.S. и соавт. Sport Nutrigenomics: Personalized Nutrition for Athletic Performance. *Front Nutr*. 2019; 6: 8.
doi: [10.3389/fnut.2019.00008](https://doi.org/10.3389/fnut.2019.00008)

Martín-Rodríguez A. и соавт. Advances in Understanding the Interplay between Dietary Practices, Body Composition, and Sports Performance in Athletes. *Nutrients*. 2024; 16(4): 571.
doi: [10.3390/nu16040571](https://doi.org/10.3390/nu16040571)

Clemente-Suárez V.J. и соавт. Personalizing Nutrition Strategies: Bridging Research and Public Health. *J Pers Med*. 2024;14(3):305.
doi: [10.3390/jpm14030305](https://doi.org/10.3390/jpm14030305)

Pickering C., Kiely J., Grgic J. и соавт. Can Genetic Testing Identify Talent for Sport? *Genes (Basel)*. 2019; 10(12): 972.
doi: [10.3390/genes10120972](https://doi.org/10.3390/genes10120972)

Rajesh Kumar и соавт. The Role of Mitochondrial Genes in Neurodegenerative Disorders. *Curr Neuroparmacol*. 2022;20(5):824–835 doi: [10.2174/1570159X19666210908163839](https://doi.org/10.2174/1570159X19666210908163839)

Sellam M. и соавт. Molecular Big Data in Sports Sciences: State-of-Art and Future Prospects of OMICS-Based Sports Sciences. *Front Mol Biosci*. 2021; 8: 815410.
doi: [10.3389/fmolb.2021.815410](https://doi.org/10.3389/fmolb.2021.815410)

Giraldo-Vallejo J.E. и соавт. Nutritional Strategies in the Rehabilitation of Musculoskeletal Injuries in Athletes: A Systematic Integrative Review. *Nutrients*. 2023; 15(4): 819.
doi: [10.3390/nu15040819](https://doi.org/10.3390/nu15040819)

Goodlin G.T., Roos A.K., Roos T.R. и соавт. Applying Personal Genetic Data to Injury Risk Assessment in Athletes. *PLoS One*. 2015; 10(4): e0122676.
doi: [10.1371/journal.pone.0122676](https://doi.org/10.1371/journal.pone.0122676)

arcOGEN Consortium and arcOGEN Collaborators Identification of new susceptibility loci for osteoarthritis (arcOGEN): a genome-wide association study. *Lancet*. 2012; 380(9844): 815–823.
doi: [10.1016/S0140-6736\(12\)60681-3](https://doi.org/10.1016/S0140-6736(12)60681-3)

Maffulli N. и соавт. The genetics of sports injuries and athletic performance. *Muscles Ligaments Tendons J*. 2013; 3(3): 173–189

Tarnowski M. и соавт. Epigenetic Alterations in Sports-Related Injuries. *Genes (Basel)*. 2022; 13(8): 1471.
doi: [10.3390/genes13081471](https://doi.org/10.3390/genes13081471)

Kumagai H. и соавт. Novel Insights into Mitochondrial DNA: Mitochondrial Microproteins and mtDNA Variants Modulate Athletic Performance and Age-Related Diseases. *Genes (Basel)*. 2023; 14(2): 286.
doi: [10.3390/genes14020286](https://doi.org/10.3390/genes14020286)