Chapter 16

MODERN ANIMAL BREEDING TECHNOLOGIES AND MOLECULAR GENETIC MARKERS

Balatsky Viktor, Saienko Artem, Peka Mykyta
DOI https://doi.org/10.30525/978-9934-26-454-2-16

Introduction

It is impossible to envision modern selection and genetic research in animal husbandry without the use of genetic markers, which are indispensable tools for addressing numerous scientific and practical challenges. To substantiate this assertion, one need only turn to the technologies of marker-associated and genomic selection, which are extensively employed in animal husbandry. These technologies are fundamentally rooted in the use of genetic markers. Naturally, amidst the array of available markers, only those associated with productive traits of animals are utilized by these technologies, thereby reflecting the correlation between genomic variability and the corresponding phenotypic variability of economically significant parameters[1, 2].

Genetic marking of productive traits can be undertaken at various levels of biological processes manifestation within the organism. For instance, biochemical, immunological, cytogenetic, and other types of markers are discernible. However, the most informative and thus highly sought-after markers are those founded upon polymorphism of the nuclear and mitochondrial genomes. Concurrently, a crucial prerequisite for the effective utilization of such molecular genetic markers is their strong association with the selected productive traits and their co-segregation with causative mutations, which directly underlie the variability of the selected parameters[3].

[1] Hayes B., Goddard M. Genome-wide association and genomic selection in animal breeding. *Genome*. 2010. Vol. 53, № 11. P. 876–883. DOI: https://doi.org/10.1139/G10-076

[2] Dekkers M. Commercial application of marker- and gene-assisted selection in livestock: Strategies and lessons. *Journal of Animal Science.* 2004. Vol. 82. (E. Suppl.). E313–E328.

[3] Using Sequence Variants in Linkage Disequilibrium with Causative Mutations to Improve Across-Breed Prediction in Dairy Cattle: A Simulation Study / I. van den Berg et al. *G3 (Bethesda, Md.)*. 2016. Vol. 6, № 8. P. 2553–2561. DOI: https://doi.org/10.1534/g3.116.027730

Marker-associated and genomic selection entail the resolution of a complex problem, encompassing the search for candidate genes, identification of polymorphic variants therein, evaluation of the frequencies of their distribution in animal populations, exploration of potential associations with productive traits enabling certain polymorphisms to be classified as molecular genetic markers, and subsequent selection of animals based on these markers. Addressing this challenge necessitates a comprehensive approach, engaging specialists from diverse disciplines and the integrated application of methodologies from bioinformatics, biostatistics, and various laboratory research techniques. In this discourse, we underscore the present status of utilizing molecular genetic markers of animal productive traits in the realm of marker-associated and genomic selection, alongside various methodologies for identifying novel markers.

1. Marker-associated and genomic selection

Marker-associated selection (MAS). This direction represents the practical implementation of the concept of utilizing "signals" in selection[4]. This method is predicated on integrating information derived from genetic markers associated with productive traits with conventional phenotypic data. The efficacy of MAS hinges upon the availability of genetic markers whose polymorphisms are closely linked to the variability of economically significant traits[5]. A fundamental aspect of contemporary MAS methodology is the utilization of DNA polymorphisms within genomic loci controlling productive traits, referred to as quantitative trait loci (QTL), as genetic markers. Moreover, the evaluation of animals is not contingent upon factors such as age, diet, husbandry practices, or other environmental variables that may influence traditional phenotypic assessments. Additionally, MAS circumvents constraints inherent in evaluating animals and selecting for traits linked to sex, traits characterized by low heritability, traits emerging during later stages of development, or those challenging to measure without necessitating the slaughter of animals[6, 7, 8]. Indeed, there has been a paradigm

[4] Reshma, R. S., Das D. N. Chapter 9 – Molecular markers and its application in animal breeding. Advances in Animal Genomics / ed. by S. Mondal, R. L. Singh. Academic Press, 2021. P. 123–140.

[5] Johnsson M., Jungnickel M. K. Evidence for and localization of proposed causative variants in cattle and pig genomes. *Genetics Selection Evolution*. 2021. Vol. 53. 67. DOI: https://doi.org/10.1186/s12711-021-00662-x

[6] Lande R., Thompson R. Efficiency of marker-assisted selection in the improvement of quantitative traits. *Genetics*. 1990. Vol. 124, № 3. P. 743–756. DOI: https://doi.org/10.1093/genetics/124.3.743

[7] Clutter A. C. Genetic selection for lifetime reproductive performance. *Society of Reproduction and Fertility Supplement*. 2009. Vol. 66. P. 293–302.

shift from traditional breeding programs reliant solely on phenotypic information for evaluating animal quality and productivity to programs integrating genotypic information, significantly enhancing the efficiency of selection decisions.

The advancement of MAS in livestock production is intricately intertwined with progress in several scientific disciplines and molecular technologies, primarily animal molecular genetics and genomics. To employ DNA markers in selection for any productivity trait, comprehensive information on associated QTL, nucleotide sequences of genes regulating the trait, their chromosomal localization, and markers either comprising causative polymorphisms or closely linked to them is essential. DNA sequencing stands as a cornerstone method in genomics, with technological advancements enabling the complete genome sequencing of numerous domestic animal species[9]. Such data are indispensable for identifying QTLs, genes, and polymorphisms directly influencing trait expression. Ultimately, genomic investigations facilitate the translation of genetic information into corresponding phenotypic traits, forming the fundamental underpinning of MAS.

To date, more than 55,000 QTLs and regions associated with specific productive traits of different animal species have been mapped in the genome. Electronic databases such as Animal QTLdb[10] and Pig Quantitative Trait Locus (QTL) Database (Pig QTLdb)[11] house information on QTLs and genomic regions associated with productivity traits.

Genomic selection. Genomic selection represents a fundamentally distinct approach that amalgamates molecular genomic insights with breeding methodologies. The core concept of genomic selection involves not only genotyping individual genome loci, including those significantly associated with traits, but also evaluating the collective effects of numerous

[8] Khmelnychyi L. M., Ovcharenko O. O. Variability of Longevity Traits Of Ukrainian red-and-white dairy cows depending on the influence of heredity of genealogical formations. *Bulletin of Sumy National Agrarian University. The Series: Livestock.* 2023. № 3. P. 78–84. DOI: https://doi.org/10.32782/bsnau.lvst.2023.3.11

[9] When Livestock Genomes Meet Third-Generation Sequencing Technology: From Opportunities to Applications / X. Liu et al. *Genes*. 2024. Vol. 15. № 2. 245. DOI: https://doi.org/10.3390/genes15020245

[10] Animal QTLdb: an improved database tool for livestock animal QTL/association data dissemination in the post-genome era / Z. L. Hu et al. *Nucleic acids research.* 2013. Vol. 41, Database issue. D871–D879. DOI: https://doi.org/10.1093/nar/gks1150

[11] A QTL resource and comparison tool for pigs: PigQTLDB / Z. L. Hu et al. *Mammalian genome : official journal of the International Mammalian Genome Society*. 2005. Vol. 16, № 10. P. 792–800. DOI: https://doi.org/10.1007/s00335-005-0060-9

genes or loci simultaneously[12]. Genes with minor or major effects are expected to yield corresponding minor or major scores, culminating in the overall genome score of the individual, which serves as the basis for direct selection. The realization of genomic selection was made feasible through the discovery of high-density genetic markers, such as single nucleotide polymorphisms (SNPs), and the advent of microarray technology for SNP detection. Genomic selection technology, which can be considered a form of marker-associated selection on a large scale, involves determining the influence of thousands of DNA markers simultaneously. Marker effects are evaluated at both the genotype and phenotype levels, initially in reference populations, the outcomes of which are subsequently used to assess the selection level of individuals by integrating their SNP genotypes with marker effects determined in the reference population. Presently, genome sequencing data can supplant SNP genotyping, markedly enhancing individual assessment accuracy and genomic selection efficiency by encompassing causative mutations within the sequencing data.

The term "genomic selection" was coined in 1998 by Haley C. S. and Visscher P. M.[13], while Meuwissen T. in 2001 devised a methodology[14] for analytically assessing breeding value based on genome-wide markers. Genomic selection offers notable advantages over traditional selection, particularly in scenarios involving the improvement of sex-linked traits, traits with low heritability, those challenging to measure (e.g., disease resistance), traits heavily influenced by environmental factors, traits observable only under specific conditions or during later stages of development, and instances where concealed traits are harbored by carriers but manifest in subsequent generations.

Despite the advantages of genomic selection, marker-associated selection remains pertinent and holds substantial potential for enhancing the productive traits of livestock. Numerous studies are underway to delineate associations of individual loci and genes with productive traits, identify causative genes and nucleotides to formulate effective genetic markers, both for MAS and genomic selection.

---

[12] Meuwissen T., Hayes B., Goddard M. Accelerating Improvement of Livestock with Genomic Selection. *Annual Review of Animal Biosciences*. 2013. Vol. 1. P. 221–237. DOI: https://doi.org/10.1146/annurev-animal-031412-103705

[13] Haley C. S., Visscher P. M. Strategies to utilize marker-quantitative trait loci associations. *Journal of Dairy Science.* 1998. Vol. 81, № 2. P. 85–97. DOI: https://doi.org/10.3168/jds.s0022-0302(98)70157-2

[14] Meuwissen T. H., Hayes B. J., Goddard M. E. Prediction of total genetic value using genome-wide dense marker maps. *Genetics*. 2001. Vol. 157, № 4. P. 1819–1829. DOI: https://doi.org/10.1093/genetics/157.4.1819

Genetic markers. The concept of genetic markers encompasses a broad spectrum of definitions stemming from the historical development of genetics and technology, as well as their application and physical properties. In the context of current discussion, genetic markers are regarded as genes, sequences, or individual nucleotides, whose polymorphic variants are linked to the expression of productive traits in animals. DNA polymorphisms arise from mutations, among which causative mutations – those directly influencing a productive trait – are of primary interest for selection purposes. When a causative polymorphism identified through molecular analysis serves as a genetic marker, it is classified as a direct genetic marker, representing the most desirable type as it directly correlates with the trait of interest. However, the majority of markers utilized in breeding technologies belong to the category of linkage disequilibrium (LD) markers, whose association with causative mutations is characterized by linkage disequilibrium. LD markers exhibit a weaker association with traits compared to direct markers, potentially due to the disruption of their linkage with traits across several generations of animals through recombination events, as well as the observed variation in such associations among different breeds and within-breed populations.

Identifying direct markers, preferred for selection purposes, and establishing the causative nature of corresponding polymorphisms is a complex undertaking, employing various methods. Nonetheless, the general logic behind these efforts typically involves progressing from identifying a causative gene to pinpointing a causative mutation within it – the root cause of the observed polymorphism and gene variability. Establishing a causative gene primarily involves two methodological approaches or their combination. Firstly, the causative nature of a gene is inferred based on its functional relevance in governing molecular events underlying the manifestation of a productive trait. Secondly, genome-wide association analysis (GWAS)[15], in conjunction with gene mapping and genome sequencing results, is employed to identify causative genes and mutations. These approaches are often supplemented by analyses such as candidate gene expression in cell cultures[16], gene knockout experiments[17] , assessing

[15] She R., Jarosz D. F. Mapping Causal Variants with Single-Nucleotide Resolution Reveals Biochemical Drivers of Phenotypic Change. *Cell.* 2018. Vol. 172, № 3. P. 478–490.e15. DOI: https://doi.org/10.1016/j.cell.2017.12.0

[16] Genetic and functional confirmation of the causality of the DGAT1 K232A quantitative trait nucleotide in affecting milk yield and composition / B. Grisart et al. *Proceedings of the National Academy of Sciences of the United States of America.* 2004. Vol. 101, № 8. P. 2398–2403. https://doi.org/10.1073/pnas.0308518100

[17] Synaptogyrin-2 influences replication of Porcine circovirus 2 / L. R. Walker et al. *PLoS Genetics*. 2018. Vol. 14: e1007750. DOI: https://doi.org/10.1371/journal.pgen.1007750

the phenotypic effects of allelic gene variants based on the analyzed polymorphism, including the utilization of genetic engineering techniques[18]. Additional evidence of the causative nature of a mutation includes its ability to induce similar phenotypic effects across different animal species, its consistent localization in orthologous genes, strict association with a productive trait, and comparisons of nucleotide sequences with amino acid sequences encoded by the causative variant of the protein.

The aforementioned methods illustrate the diverse array of approaches employed to identify causative genes and mutations. However, they do not always yield unequivocal results, often necessitating reassessment and refinement5. Furthermore, they can be resource-intensive in terms of time and expenditure. One solution to enhance the efficiency of identifying causative mutations, which could serve as a basis for a comprehensive methodological strategy, is the use of bioinformatic analysis. Bioinformatic analysis is important for assessing the impact of specific mutations on the structural and functional properties of encoded proteins, as well as studying genome and proteome databases and developing genotyping systems.

2. Bioinformatic analysis

Over the years of extensive genetic research, a wealth of information has been accumulated about the organization of genomes across various biological species, including those of agricultural significance. Concurrently, the systematic arrangement of data within genome and proteome databases substantially simplifies the further work of genetics and animal breeding specialists. Among the most prevalent resources are those provided by the National Center for Biotechnology Information (NCBI)[19], amalgamating numerous interconnected databases. Notably, the use of NCBI resources enables researchers to access information pertaining to genome assemblies, nucleotide and amino acid sequences corresponding to specific genes and their products, taxonomy, and phylogenetics across a diverse range of biological species, etc.

Another valuable tool is the Ensembl genome browser[20]. While Ensembl's scope of biological species is narrower compared to NCBI, its exceptional interface and tools provide additional opportunities for

[18] Genetic basis of speciation and adaptation: from loci to causative mutations / J. Kitano et al. *Philosophical Transactions of the Royal Society B.* 2022. Vol. 377. 20200503. DOI: https://doi.org/10.1098/rstb.2020.0

[19] Database resources of the national center for biotechnology information / E. W. Sayers et al. *Nucleic acids research.* 2022. Vol. 50, D1. D20–D26. DOI: https://doi.org/10.1093/nar/gkab1112

[20] Ensembl 2023 / F. J. Martin et al. *Nucleic acids research.* 2023. Vol. 51, D1. D933–D941. DOI: https://doi.org/10.1093/nar/gkac958

researchers. For instance, a database of polymorphic variants, inclusive of accompanying data such as rs ID, chromosomal location, and polymorphism type (e.g., synonymous, missense, intronic), as well as, where available, population frequencies and associations with phenotypes, is available for the entire species list. Ensembl also provides an integrated toolkit for phylogenetic analysis, identification of paralogues and orthologues, and alignment of nucleotide and amino acid sequences. For pigs, for example, there exists the capability for cross-breed comparisons based on different genomic assemblies.

In certain cases, researchers may necessitate using databases of varying specialization levels for their investigations. For instance, the UniProt proteome database[21] provides comprehensive information regarding proteins, encompassing their domain architecture, presence of specific structural motifs, expression profiles, and more. The Protein Data Bank (PDB) is the largest database of proteins with experimentally determined three-dimensional structures via X-ray crystallography, NMR spectroscopy, or electron microscopy[22].

Bioinformatics analysis frequently aids in optimizing laboratory experiment conditions. For example, polymerase chain reaction (PCR) for nucleic acid fragment amplification, is a component of various molecular genetic approaches. Determining the conditions for conducting PCR to solve a specific research problem is individual and is also tangential to bioinformatics. In particular, the precise selection of oligonucleotide primers delimiting the target fragment for amplification is important. Primers must meet several criteria, including adequate length (typically 18–30 nucleotides) to confer specificity to the target nucleic acid site, optimal melting temperature (preferably 52–58°C, not exceeding 65°C), minimal deviation in melting temperature among primers in pair (not exceeding 5°C), a GC content of approximately 45-60%, and placement of nucleotide G or C at the 3'-end. Furthermore, primers should not be self-complementary or form hairpins[23]. To ensure adherence to these requirements, a judicious approach involves leveraging specialized software for primer design. The literature delineates over a hundred such software solutions, encompassing both commercial and freely available products, including standalone and web-based applications. Among the most widely used software for primer design

[21]UniProt Consortium. UniProt: the Universal Protein Knowledgebase in 2023. *Nucleic acids research*. 2023. Vol. 51, D1. D523–D531. DOI: https://doi.org/10.1093/nar/gkac1052

[22] The Protein Data Bank / H. M. Berman et al. *Nucleic acids research*. 2000. Vol. 28, № 1. P. 235–242. DOI: https://doi.org/10.1093/nar/28.1.235

[23] Abd-Elsalam K. A. Bioinformatic tools and guideline for PCR primer design. *African Journal of Biotechnology*. 2003. Vol. 2, № 5. P. 91–95. DOI: https://doi.org/10.5897/AJB2003.000-1019

are Primer3 and its enhanced web interface Primer3Plus[24]. Among other tools, Primer-BLAST is worth mentioning[25]. The characterization and comparison of various non-commercial resources is presented in study by Guo et al.[26], in particular, the classification of software is carried out according to a possible specific target application, such as use in reverse transcription quantitative PCR, sequencing, detection of SNPs, splice variants, methylation or microsatellites. In addition, there is a number of free software provided by commercial companies, for example, PrimerQuest (Integrated DNA Technologies, Inc), OligoPerfect (Thermo Fisher Scientific Inc), GenScript Online PCR Primer Design Tool (GenScript Biotech Corporation), NEB Primer Design Tools (New England Biolabs), or PCR Primer Design Tool (Eurofins Genomics LLC).

Another type of software is designed to analyze the parameters of already existing primers. Noteworthy examples include OligoAnalyzer Tool (Integrated DNA Technologies, Inc), Oligo Analysis Tool (Eurofins Genomics LLC), OligoEvaluator (Sigma-Aldrich), Multiple Primer Analyzer (Thermo Fisher Scientific Inc).

When PCR is combined with restriction fragment length polymorphism (RFLP) analysis, designing restriction endonucleases becomes imperative. A large number of different enzymes within this category exists, including numerous isoschizomers, necessitating their machine search based on a specified target nucleotide sequence proximal to the polymorphic locus. Software products such as NEBcutter v3.0[27], GenScript Restriction Enzyme Map Analysis Tools (GenScript Biotech Corporation), and RestrictionMapper[28] are used to select restriction endonucleases that have the desired recognition sites.

Moreover, PCR-amplified nucleic acid fragments may harbor multiple additional recognition sites for the selected restriction endonuclease beyond the locus of interest. Consequently, the analysis of resulting

[24] Primer3Plus, an enhanced web interface to Primer3 / A. Untergasser et al. *Nucleic acids research*. 2007. Vol. 35, Web Server issue. W71–W74. DOI: https://doi.org/10.1093/nar/gkm306

[25] Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction / J. Ye et al. *BMC bioinformatics*. 2012. Vol. 13. 134. DOI: https://doi.org/10.1186/1471-2105-13–134

[26] Guo J., Starr D., Guo H. Classification and review of free PCR primer design software. *Bioinformatics (Oxford, England).* 2021. Vol. 36, № 22–23. P. 5263–5268. DOI: https://doi.org/10.1093/bioinformatics/btaa910

[27] Vincze T., Posfai J., Roberts R. J. NEBcutter: A program to cleave DNA with restriction enzymes. *Nucleic acids research*. 2003. Vol. 31, № 13. P. 3688–3691. DOI: https://doi.org/10.1093/nar/gkg526

[28] RestrictionMapper version 3. 2009. URL: https://restrictionmapper.org (date of access: 29.04.2024).

electropherograms necessitates accounting for these additional cleavage sites. To accurately assess the abundance of all available restriction endonuclease recognition sites within the amplified fragment, their localization, determination of restrict lengths, and computer simulation of expected electropherograms, online resources such as Restriction Analyzer[29] can be used.

A relatively new direction that has the potential to be used in the practice of marker-associated and genomic selection is the identification of causative polymorphisms via bioinformatic methods. This approach, which was previously described in several studies[30, 31], entails bioinformatic screening of polymorphisms within specific genes or chromosomal regions to assess their potential impact on the structural or functional properties of the encoded proteins. Polymorphic variants for which such an effect is most pronounced are considered to be directly influencing the phenotype and therefore potentially causative. This approach facilitates the targeted selection of promising polymorphisms to check their connections with the productive qualities of animals in associative studies, simultaneously reducing the amount of laboratory research.

The search for the most influential missense polymorphisms can be carried out in several ways. One such approach, termed sequence-oriented, leverages protein amino acid sequences from target genes as input data. Using homologous to target amino acid sequences, disparities induced by amino acid substitutions at a certain polymorphic site, as well as the evolutionary possibility of the corresponding replacement, it is possible prognosticate effects of missense polymorphisms on the function of the protein encoded, and, therefore, to predict the effect on the phenotype. A number of software and online services solve this problem, including SIFT[32], PolyPhen-2[33], PROVEAN[34], PANTHER-PSEP[35], MutPred2[36], SNAP2[37], etc.

[29] Restriction Analyzer. *MOLBIOTOOLS.* URL: https://molbiotools.com/restrictionanalyzer.php (date of access: 29.04.2024).

[30] Bioinformatic analysis of the effect of SNPs in the pig TERT gene on the structural and functional characteristics of the enzyme to develop new genetic markers of productivity traits / M. Peka et al. *BMC genomics.* 2023. Vol. 24, № 1. 487. DOI: https://doi.org/10.1186/s12864-023-09592-y

[31] Assessing the relationship between the in silico predicted consequences of 97 missense mutations mapping to 68 genes related to lipid metabolism and their association with porcine fatness traits / R. González-Prendes et al. *Genomics*. Vol. 115, № 2. 110589. DOI: https://doi.org/10.1016/j.ygeno.2023.110589

[32] Ng P. C., Henikoff S. SIFT: Predicting amino acid changes that affect protein function. *Nucleic acids research.* 2003. Vol. 31, № 13. P. 3812–3814. DOI: https://doi.org/10.1093/nar/gkg509

[33] Adzhubei I., Jordan D. M., Sunyaev S. R. Predicting functional effect of human missense mutations using PolyPhen-2. *Current protocols in human genetics.* 2013. Chapter 7. Unit7.20. DOI: https://doi.org/10.1002/0471142905.hg0720s76

Another, structure-oriented approach to identifying influential causative missense polymorphisms necessitates leveraging data concerning proteins' three-dimensional structures. This methodology entails evaluating the impact of amino acid substitutions at polymorphic sites on protein stability, with the change in Gibbs free energy (ΔΔG) serving as a primary indicator. Various bioinformatic resources, including mCSM[38], SDM[39], DDGun[40], PoPMuSiC[41], and PremPS[42], facilitate estimating folding free energies for individual molecules. When assessing polymorphisms' influence on protein-protein complex stability, resources such as mCSM-PPI2[43], BeAtMuSiC[44], MutaBind2[45], SAAMBE-3D[46], and BindProfX[47] are useful.

---

[34] Choi Y., Chan A. P. PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels. *Bioinformatics (Oxford, England).* 2015. Vol. 31, № 16. P. 2745–2747. DOI: https://doi.org/10.1093/bioinformatics/btv195

[35] Tang H., Thomas P. D. PANTHER-PSEP: predicting disease-causing genetic variants using position-specific evolutionary preservation. *Bioinformatics (Oxford, England)*. 2016. Vol. 32, № 14. P. 2230–2232. DOI: https://doi.org/10.1093/bioinformatics/btw222

[36] Inferring the molecular and phenotypic impact of amino acid variants with MutPred2 / V. Pejaver et al. *Nature communications*. 2020. Vol. 11, № 1. 5918. DOI: https://doi.org/10.1038/s41467-020-19669-x

[37] Hecht M., Bromberg Y., Rost B. Better prediction of functional effects for sequence variants. *BMC Genomics.* 2015. Vol. 16, Suppl. 8. S1. DOI: https://doi.org/10.1186/1471-2164-16-S8-S1

[38] Pires D. E., Ascher D. B., Blundell T. L. mCSM: predicting the effects of mutations in proteins using graph-based signatures. *Bioinformatics (Oxford, England)*. 2014. Vol. 30, № 3. P. 335–342. DOI: https://doi.org/10.1093/bioinformatics/btt691

[39] SDM: a server for predicting effects of mutations on protein stability / A. P. Pandurangan et al. *Nucleic acids research.* 2017. Vol. 45, W1. W229–W235. DOI: https://doi.org/10.1093/nar/gkx439

[40] DDGun: an untrained method for the prediction of protein stability changes upon single and multiple point variations / L. Montanucci et al. *BMC bioinformatics*. 2019. Vol. 20, Suppl. 14. 335. DOI: https://doi.org/10.1186/s12859-019-2923-1

[41] Fast and accurate predictions of protein stability changes upon mutations using statistical potentials and neural networks: PoPMuSiC-2.0 / Y. Dehouck et al. *Bioinformatics (Oxford, England)*. 2009. Vol. 25, № 19. P. 2537–2543. DOI: https://doi.org/10.1093/bioinformatics/btp445

[42] PremPS: Predicting the impact of missense mutations on protein stability / Y. Chen et al. *PLoS computational biology*. 2020. Vol. 16, № 12. e1008543. DOI: https://doi.org/10.1371/journal.pcbi.1008543

[43] mCSM-PPI2: predicting the effects of mutations on protein-protein interactions / C. H. M. Rodrigues et al. *Nucleic acids research*. 2019. Vol. 47, W1. W338–W344. DOI: https://doi.org/10.1093/nar/gkz383

[44] BeAtMuSiC: Prediction of changes in protein-protein binding affinity on mutations / Y. Dehouck et al. *Nucleic acids research*. 2013. Vol. 41, Web Server issue. W333–W339. DOI: https://doi.org/10.1093/nar/gkt4

[45] MutaBind2: Predicting the Impacts of Single and Multiple Mutations on Protein-Protein Interactions / N. Zhang et al. *iScience*. 2020. Vol. 23, № 3. 100939. DOI: https://doi.org/10.1016/j.isci.2020.100939

At the same time, determining which polymorphisms will undergo laboratory testing in association studies to explore their correlation with the productive traits of animals presents a challenging endeavor, often necessitating the preliminary implementation of a comprehensive bioinformatic analysis. In many cases, this will include work with genetic databases, phylogenetic analysis, screening of polymorphisms regarding their impact on the structural and functional properties of proteins.

3. Genotyping methods

The genotyping of numerous animal groups for a large number of polymorphisms necessitates the utilization of efficient and cost-effective technical methodologies. Since the inception of DNA analysis, various genotyping techniques have been developed, documented in different publications[48, 49].

Primarily, the leading technologies employed for determining DNA polymorphism and genotyping encompass biological microarrays for detecting SNPs, indels (InDels), copy number variations (CNVs), gene and genomic loci sequencing, and capillary electrophoresis. While these modern techniques dominate, other genotyping approaches remain relevant, particularly for individual SNP studies. Among these, allelic discrimination via the TaqMan PCR method, as well as traditional PCR-RFLP and SSCP methods, are noteworthy. It can be argued that there are no inherent challenges in discerning DNA polymorphism currently. The crucial consideration lies in the judicious selection of methodological approaches based on study objectives, method efficiency, and cost-effectiveness.

Restriction fragment length polymorphisms (RFLP). This method was proposed in 1978[50] and relies on detecting alterations in base pairs that either

[46] SAAMBE-3D: Predicting Effect of Mutations on Protein-Protein Interactions / S. Pahari et al. *International journal of molecular sciences.* 2020. Vol. 21, № 7. 2563. DOI: https://doi.org/10.3390/ijms21072563

[47] BindProfX: Assessing Mutation-Induced Binding Affinity Change by Protein Interface Profiles with Pseudo-Counts / P. Xiong et al. *Journal of molecular biology*. 2017. Vol. 429, № 3. 426–434. DOI: https://doi.org/10.1016/j.jmb.2016.11.022

[48] DNA Analysis of Domestic Animals / K. Kaitholia et al. *Forensic DNA Typing: Principles, Applications and Advancements* / Ed. by P. Shrivastava, H. R. Dash, J. A. Lorente, J. Imam. Singapore: Springer, 2020. P. 379–397. DOI: https://doi.org/10.1007/978-981-15-6655-4_19

[49] Jordan D., Mills D. Past, Present, and Future of DNA Typing for Analyzing Human and Non-Human Forensic Samples. *Frontiers in Ecology and Evolution.* 2021. Vol. 9. DOI: https://doi.org/10.3389/fevo.2021.646130

[50] Kan Y. W., Dozy A. M. Polymorphism of DNA sequence adjacent to human beta-globin structural gene: Relationship to sickle mutation. *Proceedings of the National Academy of Sciences of the United States of America.* 1978. Vol. 75, № 11. P. 5631–5635. DOI: https://doi.org/10.1073/pnas.75.11.5631

create or abolish a restriction site for a specific endonuclease. The resultant variation is discernible post-enzyme action, with DNA fragments detected via electrophoresis. Presently, PCR-RFLP variants are utilized, when the DNA fragment selected for analysis is amplified in site-specific PCR[51]. Despite the fact that PCR-RFLP is not characterized by high productivity, this method continues to be used successfully for locus-specific genotyping, especially in resource-constrained laboratories.

Single-Strand Conformational Polymorphism (SSCP). This technique involves denaturing PCR products into single strands and separating them on non-denaturing gels[52]. Fragment migration under these conditions is contingent upon its three-dimensional configuration, influenced in part by its sequence. While relatively straightforward, SSCP lacks the sensitivity of alternative methods and is susceptible to various factors including temperature and fragment size.

Biological microchip technology. The most powerful technique for genotyping SNPs is solid-state detection, enabling the analysis of thousands of SNPs in individual genomes overnight. This approach is based on the hybridization of the analyzed DNA with allele-specific probes (corresponding to specific SNPs), which are fixed onto microchip surfaces and separated in space. Microchip arrays facilitate mutation screening and common allele detection by evaluating hybridization signal levels. Currently, SNP microarrays are extensively employed for genotyping various animal species. For instance, according to scientific data[53], in livestock genomic selection, low-grade microarrays containing SNPs ranging from 2,900 (Illumina Golden Gate Bovine3K) to 648,875 (Affymetrix Axiom Genome-Wide BOS) are used. Microarrays with up to 600,000 SNPs are used for chicken genotyping[54]. For large-scale scanning of pig genome and genomic selection of these animals, the Illumina Porcine SNP60 v2 Genotyping BeadChip microarray of more than 64 thousand SNPs is used[55].

[51]Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction / K. Mullis et al. *Cold Spring Harbor symposia on quantitative biology*. 1986. Vol. 51, Pt. 1. P. 263–273. DOI: https://doi.org/10.1101/sqb.1986.051.01.032

[52] Detection of polymorphisms of human DNA by gel electrophoresis as single-strand conformation polymorphisms / M. Orita et al. *Proceedings of the National Academy of Sciences of the United States of America.* 1989. Vol. 86, № 8. P. 2766–2770. DOI: https://doi.org/10.1073/pnas.86.8.2766

[53]SNPchiMp: a database to disentangle the SNPchip jungle in bovine livestock / E. L. Nicolazzi et al. *BMC Genomics*. 2014. Vol. 15. 123. DOI: https://doi.org/10.1186/1471-2164-15-123

[54]Development of a high density 600K SNP genotyping array for chicken / A. Kranis et al. *BMC Genomics.* 2013. Vol. 14. 59. https://doi.org/10.1186/1471-2164-14-59

[55]Design of a high density SNP genotyping assay in the pig using SNPs identified and characterized by next generation sequencing technology / A. M. Ramos et al. *PLoS ONE*. 2009. Vol. 4, № 8. e6524. DOI: https://doi.org/10.1371/journal.pone.0006524

Additionally, specialized microchip panels have been developed: GeenSeek Genomic Profiler for Porcine LD (10,241 SNPs which are associated with productive and reproductive traits of pigs) and GeenSeek Genomic Profiler for Porcine HD (68,528 SNPs with are 43,000 most informative SNPs from PorcineSNP60 v2 and additional 25,000 SNPs that are localized in the telomeric zones of chromosomes and other areas that are not covered by PorcineSNP60 v2)[56].

TagMan allelic discrimination. TaqMan SNP genotyping provides a versatile technology for polymorphism detection in any genome. Leveraging high-performance TaqMan reagents, meticulously designed probes and primers, along with modern RealTime-PCR instruments and corresponding software, genotyping results for SNPs can be quickly obtained. The principle of TaqMan allelic discrimination is described in many sources[57, 58]. TaqMan kits find utility across various genotyping techniques, including screening, association studies, candidate site or gene analysis, and fine genetic mapping.

Sequencing. Sequencing encompasses a variety of techniques aimed at determining the sequence of nucleotides within polymer chains of nucleic acids[59]. The progression of sequencing methods from the classic Sanger sequencing (first-generation sequencing)[60] to high-performance next-generation sequencing techniques has been characterized by improvements in both quantitative and qualitative aspects of sequencing outcomes. Second-generation sequencing methods are based on clonal amplification of DNA molecule where billions of different short DNA fragments get sequenced at the same time in parallel[61]. Notable examples of second-generation

[56]Samorè A. B., Fontanesi, L. Genomic selection in pigs: state of the art and perspectives. *Italian Journal Of Animal Science*. 2016. Vol. 15, № 2. P. 211–232. DOI: https://doi.org/10.1080/1828051X.2016.1172034

[57]Mutation detection by TaqMan-allele specific amplification: application to molecular diagnosis of glycogen storage disease type Ia and medium-chain acyl-CoA dehydrogenase deficiency / K. Fujii et al. *Human Mutation*. 2000. Vol. 15, № 2. P. 189–196. DOI: https://doi.org/10.1002/(SICI)1098-1004(200002)15:2<189::AID-HUMU8>3.0.CO;2-H

[58] Allele specific Taqman-based real-time PCR assay to quantify circulating BRAFV600E mutated DNA in plasma of melanoma patients / P. Pinzani et al. *Clinical Chemistry and Laboratory Medicine*. 2010. Vol. 48, № 6. P. 669–676. DOI: https://doi.org/10.1016/j.cca.2010.05.024

[59] Heather J. M., Chain B. The sequence of sequencers: The history of sequencing DNA. *Genomics*. 2016. Vol. 107, № 1. P. 1–8. DOI: https://doi.org/10.1016/j.ygeno.2015.11.003

[60] Sanger F., Nicklen S., Coulson A. R. DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences of the United States of America.* 1977. Vol. 74, № 12. P. 5463–5467. DOI: https://doi.org/10.1073/pnas.74.12.5463

[61] Gupta N., Verma V. K. Next-Generation Sequencing and Its Application: Empowering in Public Health Beyond Reality. *Microbial Technology for the Welfare of Society.* 2019. Vol. 17. P. 313–341. DOI: https://doi.org/10.1007/978-981-13-8844-6_15

sequencing techniques include Roche 454, Illumina Genome Analyzer, and Applied Biosystems SOLiD[62]. In contrast, third-generation sequencing techniques, such as Single Molecular Real-Time (SMRT) sequencing from Pacific Biosciences or Nanopore sequencing by Oxford Nanopore Technologies, enable the sequencing of individual DNA molecules without the need for pre-amplification[63].

The enhancement in sequencing efficiency in recent years has culminated in the completion of whole-genome sequencing for key agricultural species. In many cases there is more than one genome assembly for a certain species, which in particular allows tracing the differences between breeds. Whole-genome sequencing serves as a potent means for identifying numerous polymorphic variants within animal genomes, i.e., it is actually the main source of information about SNPs that can potentially be associated with the productive traits of animals. These SNPs can be used to develop species-specific SNP microarrays and be incorporated into genomic breeding practices[64].

Consequently, a number laboratory methodologies are available for the purposes of animal breeding, enabling the identification and analysis of polymorphisms within animal genomes. The aforementioned compilation encompasses the most prevalent molecular genetic approaches and may be expanded. Each approach possesses distinctive characteristics, advantages, and limitations regarding practical utility, thus necessitating judicious selection within the context of specific breeding investigations.

CONCLUSIONS

In our study, we have presented information concerning the use of molecular genetic markers based on DNA polymorphism within modern breeding methodologies, namely marker-associated and genomic selection. We delve into the exploration of methodological approaches aimed at the development of the most efficacious DNA markers. While not exhaustive, our work encompasses a broad spectrum of issues pertinent to this research area. The integrated use of bioinformatic and laboratory techniques facilitates the examination of polymorphisms across various animal genes, thereby enabling the identification of pertinent correlations with productive traits through associative analysis. Thus, the issue of developing novel

[62] Genomic sequencing in clinical trials / K. K. Mestan et al. *Journal of translational medicine.* 2011. Vol. 9. 222. DOI: https://doi.org/10.1186/1479-5876-9-222

[63] Next generation sequencing in animal science – A review / A. Dunisławska et al. *Animal Science Papers and Reports.* 2017. Vol. 35, № 3. P. 205–224.

[64]Next generation sequencing in livestock species – A Review / A. Sharma et al. *Journal of Animal Breeding and Genomics.* 2017. Vol. 1, № 1. P. 23–30. DOI: https://doi.org/ 10.12972/jabng.20170003

genetic markers that can be used in selective breeding practices is being addressed.

We particularly wish to direct the reader's attention to a novel approach in the pursuit of optimal DNA markers, namely the use of bioinformatic analysis. Such an analysis provides a fundamentally new tool for identifying and proving the causality of polymorphisms of candidate genes in relation to productive traits. Significantly, the novelty lies in the predictive computational bioinformatic evaluation of the impact of DNA polymorphisms on the structural and functional features of the protein encoded by the gene, and, accordingly, on the trait in which such protein participates in its formation. Causative polymorphisms, as noted above, are in themselves the most effective genetic markers. This bioinformatic approach is illustrated in several publications where missense polymorphisms were analyzed. However, with the advent of new computational algorithms and corresponding evaluation capabilities, we expect the evolution of approaches that consider polymorphisms not only affecting protein structure, but also influencing gene expression, its kinetics, specificity, alternative splicing, and other molecular mechanisms.

SUMMARY

Modern breeding and selection of farm animals are closely linked to genetic research. This study delineates two main directions: marker-associated and genomic selection, both based on the use of genetic markers associated with various productive traits of animals. The characteristics of potential applications of bioinformatic approaches for addressing challenges related to the analysis of genetic data and optimizing laboratory conditions for animal genotyping are outlined. Particular emphasis is placed on a novel application of bioinformatic analysis in animal breeding: the screening of polymorphic variants in animal genomes and selecting those with the greatest potential to influence the structural and functional properties of the proteins. These variants act as causative polymorphisms crucial for marker-associated and genomic selection. Additionally, the work offers an overview of common laboratory methods for genotyping animals and identifying polymorphisms. Thus, the search for genetic markers and their integration into breeding practices is characterized as a complex task necessitating the integrated use of various bioinformatic and laboratory techniques.

Bibliography

1. Hayes B., Goddard M. Genome-wide association and genomic selection in animal breeding. *Genome*. 2010. Vol. 53, № 11. P. 876–883. DOI: https://doi.org/10.1139/G10-076

2. Dekkers M. Commercial application of marker- and gene-assisted selection in livestock: Strategies and lessons. *Journal of Animal Science*. 2004. Vol. 82. (E. Suppl.). E313–E328.

3. Using Sequence Variants in Linkage Disequilibrium with Causative Mutations to Improve Across-Breed Prediction in Dairy Cattle: A Simulation Study / I. van den Berg et al. *G3 (Bethesda, Md.)*. 2016. Vol. 6, № 8. P. 2553–2561. DOI: https://doi.org/10.1534/g3.116.027730

4. Reshma, R. S., Das D. N. Chapter 9 – Molecular markers and its application in animal breeding. Advances in Animal Genomics / ed. by S. Mondal, R. L. Singh. Academic Press, 2021. P. 123–140.

5. Johnsson M., Jungnickel M. K. Evidence for and localization of proposed causative variants in cattle and pig genomes. *Genetics Selection Evolution*. 2021. Vol. 53. 67. DOI: https://doi.org/10.1186/s12711-021-00662-x

6. Lande R., Thompson R. Efficiency of marker-assisted selection in the improvement of quantitative traits. *Genetics*. 1990. Vol. 124, № 3. P. 743–756. DOI: https://doi.org/10.1093/genetics/124.3.743

7. Clutter A. C. Genetic selection for lifetime reproductive performance. *Society of Reproduction and Fertility Supplement*. 2009. Vol. 66. P. 293–302.

8. Khmelnychyi L. M., Ovcharenko O. O. Variability of Longevity Traits Of Ukrainian red-and-white dairy cows depending on the influence of heredity of genealogical formations. *Bulletin of Sumy National Agrarian University. The Series: Livestock.* 2023. № 3. P. 78–84. DOI: https://doi.org/10.32782/bsnau.lvst.2023.3.11

9. When Livestock Genomes Meet Third-Generation Sequencing Technology: From Opportunities to Applications / X. Liu et al. *Genes*. 2024. Vol. 15. № 2. 245. DOI: https://doi.org/10.3390/genes15020245

10. Animal QTLdb: an improved database tool for livestock animal QTL/association data dissemination in the post-genome era / Z. L. Hu et al. *Nucleic acids research.* 2013. Vol. 41, Database issue. D871–D879. DOI: https://doi.org/10.1093/nar/gks1150

11. A QTL resource and comparison tool for pigs: PigQTLDB / Z. L. Hu et al. *Mammalian genome : official journal of the International Mammalian Genome Society*. 2005. Vol. 16, № 10. P. 792–800. DOI: https://doi.org/10.1007/s00335-005-0060-9

12. Meuwissen T., Hayes B., Goddard M. Accelerating Improvement of Livestock with Genomic Selection. *Annual Review of Animal Biosciences*. 2013. Vol. 1. P. 221–237. DOI: https://doi.org/10.1146/annurev-animal-031412-103705

13. Haley C. S., Visscher P. M. Strategies to utilize marker-quantitative trait loci associations. *Journal of Dairy Science.* 1998. Vol. 81, № 2. P. 85–97. DOI: https://doi.org/10.3168/jds.s0022-0302(98)70157-2

14. Meuwissen T. H., Hayes B. J., Goddard M. E. Prediction of total genetic value using genome-wide dense marker maps. *Genetics*. 2001. Vol. 157, № 4. P. 1819–1829. DOI: https://doi.org/10.1093/genetics/157.4.1819

15. She R., Jarosz D. F. Mapping Causal Variants with Single-Nucleotide Resolution Reveals Biochemical Drivers of Phenotypic Change. *Cell.* 2018. Vol. 172, № 3. P. 478–490.e15. DOI: https://doi.org/10.1016/j.cell.2017.12.015

16. Genetic and functional confirmation of the causality of the DGAT1 K232A quantitative trait nucleotide in affecting milk yield and composition / B. Grisart et al. *Proceedings of the National Academy of Sciences of the United States of America.* 2004. Vol. 101, № 8. P. 2398–2403. https://doi.org/10.1073/pnas.0308518100

17. Synaptogyrin-2 influences replication of Porcine circovirus 2 / L. R. Walker et al. *PLoS Genetics*. 2018. Vol. 14: e1007750. DOI: https://doi.org/10.1371/journal.pgen.1007750

18. Genetic basis of speciation and adaptation: from loci to causative mutations / J. Kitano et al. *Philosophical Transactions of the Royal Society B.* 2022. Vol. 377. 20200503. DOI: https://doi.org/10.1098/rstb.2020.0503

19. Database resources of the national center for biotechnology information / E. W. Sayers et al. *Nucleic acids research.* 2022. Vol. 50, D1. D20–D26. DOI: https://doi.org/10.1093/nar/gkab1112

20. Ensembl 2023 / F. J. Martin et al. *Nucleic acids research.* 2023. Vol. 51, D1. D933–D941. DOI: https://doi.org/10.1093/nar/gkac958

21. UniProt Consortium. UniProt: the Universal Protein Knowledgebase in 2023. *Nucleic acids research*. 2023. Vol. 51, D1. D523–D531. DOI: https://doi.org/10.1093/nar/gkac1052

22. The Protein Data Bank / H. M. Berman et al. *Nucleic acids research*. 2000. Vol. 28, № 1. P. 235–242. DOI: https://doi.org/10.1093/nar/28.1.235

23. Abd-Elsalam K. A. Bioinformatic tools and guideline for PCR primer design. *African Journal of Biotechnology*. 2003. Vol. 2, № 5. P. 91–95. DOI: https://doi.org/10.5897/AJB2003.000-1019

24. Primer3Plus, an enhanced web interface to Primer3 / A. Untergasser et al. *Nucleic acids research*. 2007. Vol. 35, Web Server issue. W71–W74. DOI: https://doi.org/10.1093/nar/gkm306

25. Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction / J. Ye et al. *BMC bioinformatics*. 2012. Vol. 13. 134. DOI: https://doi.org/10.1186/1471-2105-13-134

26. Guo J., Starr D., Guo H. Classification and review of free PCR primer design software. *Bioinformatics (Oxford, England).* 2021. Vol. 36, № 22–23. P. 5263–5268. DOI: https://doi.org/10.1093/bioinformatics/btaa910

27. Vincze T., Posfai J., Roberts R. J. NEBcutter: A program to cleave DNA with restriction enzymes. *Nucleic acids research*. 2003. Vol. 31, № 13. P. 3688–3691. DOI: https://doi.org/10.1093/nar/gkg526

28. RestrictionMapper version 3. 2009. URL: https://restrictionmapper.org (date of access: 29.04.2024).

29. Restriction Analyzer. *MOLBIOTOOLS.* URL: https://molbiotools.com/restrictionanalyzer.php (date of access: 29.04.2024).

30. Bioinformatic analysis of the effect of SNPs in the pig TERT gene on the structural and functional characteristics of the enzyme to develop new genetic markers of productivity traits / M. Peka et al. *BMC genomics*. 2023. Vol. 24, № 1. 487. DOI: https://doi.org/10.1186/s12864-023-09592-y

31. Assessing the relationship between the in silico predicted consequences of 97 missense mutations mapping to 68 genes related to lipid metabolism and their association with porcine fatness traits / R. González-Prendes et al. *Genomics*. Vol. 115, № 2. 110589. DOI: https://doi.org/10.1016/j.ygeno.2023.110589

32. Ng P. C., Henikoff S. SIFT: Predicting amino acid changes that affect protein function. *Nucleic acids research.* 2003. Vol. 31, № 13. P. 3812–3814. DOI: https://doi.org/10.1093/nar/gkg509

33. Adzhubei I., Jordan D. M., Sunyaev S. R. Predicting functional effect of human missense mutations using PolyPhen-2. *Current protocols in human genetics.* 2013. Chapter 7. Unit7.20. DOI: https://doi.org/10.1002/0471142905.hg0720s76

34. Choi Y., Chan A. P. PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels. *Bioinformatics (Oxford, England).* 2015. Vol. 31, № 16. P. 2745–2747. DOI: https://doi.org/10.1093/bioinformatics/btv195

35. Tang H., Thomas P. D. PANTHER-PSEP: predicting disease-causing genetic variants using position-specific evolutionary

preservation. *Bioinformatics (Oxford, England)*. 2016. Vol. 32, № 14. P. 2230–2232. DOI: https://doi.org/10.1093/bioinformatics/btw222

36. Inferring the molecular and phenotypic impact of amino acid variants with MutPred2 / V. Pejaver et al. *Nature communications*. 2020. Vol. 11, № 1. 5918. DOI: https://doi.org/10.1038/s41467-020-19669-x

37. Hecht M., Bromberg Y., Rost B. Better prediction of functional effects for sequence variants. *BMC Genomics.* 2015. Vol. 16, Suppl. 8. S1. DOI: https://doi.org/10.1186/1471-2164-16-S8-S1

38. Pires D. E., Ascher D. B., Blundell T. L. mCSM: predicting the effects of mutations in proteins using graph-based signatures. *Bioinformatics (Oxford, England)*. 2014. Vol. 30, № 3. P. 335–342. DOI: https://doi.org/10.1093/bioinformatics/btt691

39. SDM: a server for predicting effects of mutations on protein stability / A. P. Pandurangan et al. *Nucleic acids research.* 2017. Vol. 45, W1. W229–W235. DOI: https://doi.org/10.1093/nar/gkx439

40. DDGun: an untrained method for the prediction of protein stability changes upon single and multiple point variations / L. Montanucci et al. *BMC bioinformatics*. 2019. Vol. 20, Suppl. 14. 335. DOI: https://doi.org/10.1186/s12859-019-2923-1

41. Fast and accurate predictions of protein stability changes upon mutations using statistical potentials and neural networks: PoPMuSiC-2.0 / Y. Dehouck et al. *Bioinformatics (Oxford, England)*. 2009. Vol. 25, № 19. P. 2537–2543. DOI: https://doi.org/10.1093/bioinformatics/btp445

42. PremPS: Predicting the impact of missense mutations on protein stability / Y. Chen et al. *PLoS computational biology*. 2020. Vol. 16, № 12. e1008543. DOI: https://doi.org/10.1371/journal.pcbi.1008543

43. mCSM-PPI2: predicting the effects of mutations on protein-protein interactions / C. H. M. Rodrigues et al. *Nucleic acids research*. 2019. Vol. 47, W1. W338–W344. DOI: https://doi.org/10.1093/nar/gkz383

44. BeAtMuSiC: Prediction of changes in protein-protein binding affinity on mutations / Y. Dehouck et al. *Nucleic acids research*. 2013. Vol. 41, Web Server issue. W333–W339. DOI: https://doi.org/10.1093/nar/gkt450

45. MutaBind2: Predicting the Impacts of Single and Multiple Mutations on Protein-Protein Interactions / N. Zhang et al. *iScience*. 2020. Vol. 23, № 3. 100939. DOI: https://doi.org/10.1016/j.isci.2020.100939

46. SAAMBE-3D: Predicting Effect of Mutations on Protein-Protein Interactions / S. Pahari et al. *International journal of molecular*

sciences. 2020. Vol. 21, № 7. 2563. DOI: https://doi.org/10.3390/ijms21072563

47. BindProfX: Assessing Mutation-Induced Binding Affinity Change by Protein Interface Profiles with Pseudo-Counts / P. Xiong et al. *Journal of molecular biology*. 2017. Vol. 429, № 3. 426–434. DOI: https://doi.org/10.1016/j.jmb.2016.11.022

48. DNA Analysis of Domestic Animals / K. Kaitholia et al. *Forensic DNA Typing: Principles, Applications and Advancements* / ed. by P. Shrivastava, H. R. Dash, J. A. Lorente, J. Imam. Singapore: Springer, 2020. P. 379–397. DOI: https://doi.org/10.1007/978-981-15-6655-4_19

49. Jordan D., Mills D. Past, Present, and Future of DNA Typing for Analyzing Human and Non-Human Forensic Samples. *Frontiers in Ecology and Evolution.* 2021. Vol. 9. DOI: https://doi.org/10.3389/fevo.2021.646130

50. Kan Y. W., Dozy A. M. Polymorphism of DNA sequence adjacent to human beta-globin structural gene: Relationship to sickle mutation. *Proceedings of the National Academy of Sciences of the United States of America.* 1978. Vol. 75, № 11. P. 5631–5635. DOI: https://doi.org/10.1073/pnas.75.11.5631

51. Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction / K. Mullis et al. *Cold Spring Harbor symposia on quantitative biology*. 1986. Vol. 51, Pt. 1. P. 263–273. DOI: https://doi.org/10.1101/sqb.1986.051.01.032

52. Detection of polymorphisms of human DNA by gel electrophoresis as single-strand conformation polymorphisms / M. Orita et al. *Proceedings of the National Academy of Sciences of the United States of America.* 1989. Vol. 86, № 8. P. 2766–2770. DOI: https://doi.org/10.1073/pnas.86.8.2766

53. SNPchiMp: a database to disentangle the SNPchip jungle in bovine livestock / E. L. Nicolazzi et al. *BMC Genomics*. 2014. Vol. 15. 123. DOI: https://doi.org/10.1186/1471-2164-15-123

54. Development of a high density 600K SNP genotyping array for chicken / A. Kranis et al. *BMC Genomics.* 2013. Vol. 14. 59. https://doi.org/10.1186/1471-2164-14-59

55. Design of a high density SNP genotyping assay in the pig using SNPs identified and characterized by next generation sequencing technology / A. M. Ramos et al. *PLoS ONE*. 2009. Vol. 4, № 8. e6524. DOI: https://doi.org/10.1371/journal.pone.0006524

56. Samorè A. B., Fontanesi, L. Genomic selection in pigs: state of the art and perspectives. *Italian Journal Of Animal Science*. 2016.

Vol. 15, № 2. P. 211–232. DOI: https://doi.org/10.1080/1828051X.2016.1172034

57. Mutation detection by TaqMan-allele specific amplification: application to molecular diagnosis of glycogen storage disease type Ia and medium-chain acyl-CoA dehydrogenase deficiency / K. Fujii et al. *Human Mutation*. 2000. Vol. 15, № 2. P. 189–196. DOI: https://doi.org/10.1002/(SICI)1098-1004(200002)15:2<189::AID-HUMU8>3.0.CO;2-H

58. Allele specific Taqman-based real-time PCR assay to quantify circulating BRAFV600E mutated DNA in plasma of melanoma patients / P. Pinzani et al. *Clinical Chemistry and Laboratory Medicine*. 2010. Vol. 48, № 6. P. 669-676. DOI: https://doi.org/10.1016/j.cca.2010.05.024

59. Heather J. M., Chain B. The sequence of sequencers: The history of sequencing DNA. *Genomics*. 2016. Vol. 107, № 1. P. 1–8. DOI: https://doi.org/10.1016/j.ygeno.2015.11.003

60. Sanger F., Nicklen S., Coulson A. R. DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences of the United States of America*. 1977. Vol. 74, № 12. P. 5463–5467. DOI: https://doi.org/10.1073/pnas.74.12.5463

61. Gupta N., Verma V. K. Next-Generation Sequencing and Its Application: Empowering in Public Health Beyond Reality. *Microbial Technology for the Welfare of Society*. 2019. Vol. 17. P. 313–341. DOI: https://doi.org/10.1007/978-981-13-8844-6_15

62. Genomic sequencing in clinical trials / K. K. Mestan et al. *Journal of translational medicine*. 2011. Vol. 9. 222. DOI: https://doi.org/10.1186/1479-5876-9-222

63. Next generation sequencing in animal science – A review / A. Dunisławska et al. *Animal Science Papers and Reports*. 2017. Vol. 35, № 3. P. 205–224.

64. Next generation sequencing in livestock species – A Review / A. Sharma et al. *Journal of Animal Breeding and Genomics*. 2017. Vol. 1, № 1. P. 23-–30. DOI: https://doi.org/10.12972/jabng.20170003

Information about the authors:
Balatsky Viktor Mykolaiovych,
Doctor of Agricultural Sciences, Professor,
Principal researcher at the Genetics laboratory,
Institute of Pig Breeding and Agroindustrial Production
of the National Academy of Agrarian Sciences of Ukraine
1, Shvedska Mohyla str., Poltava, 36013, Ukraine

Saienko Artem Mykhailovych,
Candidate of Agricultural Sciences, Senior researcher,
Head of the Genetics laboratory,
Institute of Pig Breeding and Agroindustrial Production
of the National Academy of Agrarian Sciences of Ukraine
1, Shvedska Mohyla str., Poltava, 36013, Ukraine

Peka Mykyta Yuriiovych,
Researcher at the Genetics laboratory,
Institute of Pig Breeding and Agroindustrial Production
of the National Academy of Agrarian Sciences of Ukraine
1, Shvedska Mohyla str., Poltava, 36013, Ukraine