

Joanna GRUSZCZYŃSKA, Beata GRZEGRZÓŁKA, Agnieszka MORAWSKA

**GENOME COMPARATIVE STUDIES IN THE *Phasianidae* FAMILY
PART II. INTERSPECIFIC AMPLIFICATION IN THE FAMILY *Phasianidae*
ON THE EXAMPLE OF DOMESTIC CHICKEN (*Gallus gallus domesticus*)
AND WILD TURKEY (*Meleagris gallopavo*)**

Department of Genetics and Animal Breeding, Warsaw University of Life Sciences – SGGW, Poland

Abstract. In studies aimed at understanding the genome of turkey, the great facilitation was the use of gene sequences and markers previously used to study the chicken genome. This was possible due to the fact that domestic chicken and wild turkey are phylogenetically related species with a common ancestor. In numerous analyses using chicken sequence, positive results of amplification of the turkey sequence were obtained, which accelerated the recognition of the genome of this species. Research aimed at the sequencing of the turkey genome used methods such as DNA microarrays, expression microarrays, identification of QTLs and candidate genes, and re-sequencing.

Key word: domestic chicken, wild turkey, microsatellite sequences, QTLs, expression microarray, DNA microarray, candidate genes, re-sequencing.

INTRODUCTION

Based on the knowledge of the presence of genes / conserved sequences in terms of their distribution in the chromosome, the so-called syntenic genes, cross-species amplification tests are carried out. For this purpose, primers designed according to the nucleotide sequence of one species are used to generate microsatellite markers in another species. Success in research depends on the degree of phylogenetic relatedness of the analysed species (Gruszczyńska et al. 2018).

The aim of this article was to present domestic chicken like a model organism which sequences and genes are useful in to investigations genomes other birds species.

INTERSPECIFIC AMPLIFICATION USING MICROSATELLITE SEQUENCES

Interspecific amplification of microsatellite *loci* is extremely helpful in creating genetic maps of economically important species. Considering advances in research on the genetic map of the domestic chicken genome and the promising potential of comparative research, an attempt was made to use specific chicken microsatellites in turkey and check their informativeness. Microsatellite sequences have the highest polymorphism in the species for

which the primers were designed. The published results of previous studies with the use of different markers were ambiguous (Levin et al. 1995; Liu et al. 1996). Levin et al. (1995) in their studies investigated 48 domestic chicken markers with 92% success in the amplification of turkey genomic DNA. In addition, a significant polymorphism was observed in the subgroup of these markers, which means that the majority of chicken microsatellite markers could be used to map the turkey genome. However, studies by Liu et al. (1996), investigating 88 markers, brought completely opposite results to those published by Levin (1995). Despite positive results in the amplification of genomic turkey DNA obtained with 69% of the markers used, low allele variability and low heterozygosity were found in the studied turkey population. It was found that chicken microsatellites would provide a weak tool for constructing genetic map of a turkey.

In 1996 U.S.Poultry Genome Coordinator made available a large collection of microsatellite starters of domestic chicken. Intensive research and effort of the researchers led to the use of these markers (141 pairs of primers) for mapping turkey genome. PCR products were obtained for 78 of the 141 primer pairs, which was 55% of the positive amplification. The PCR fragments obtained in turkey were mostly (69% of cases) of similar length to the respective domestic chicken *loci*. Dinucleotide repeats (CA/TG) were determined by hybridization. A direct relationship was found between the intensity of hybridization and the length of dinucleotide repeats in the turkey in relation to chicken sequence. Amplification of the homologous loci was confirmed by direct sequencing and comparing of the turkey and domestic chicken sequences. Results indicated that the use of specific chicken microsatellite starters quickly and significantly accelerated construction of the genetic map of the turkey (Reed et al. 1999).

In subsequent interspecific amplification studies, the number of chicken microsatellite starter kits was increased to 520. As a result of PCR for 280 turkey primer pairs (54%), turkey amplification products were obtained. Most of the obtained amplification products were similar or nearly identical in size and it corresponded to the size of the fragment of the given chicken *locus*. The structure of dinucleotide repeats and flanking sequences was determined on the basis of 13 turkey fragments, amplified with the use of chicken starters. 5 chicken *loci* were amplified with turkey starters. With sequence analysis it was possible to find numerous mutations occurring in the sequences of both species, in addition to the differences found in the number of repeats. The usefulness of the amplified *loci* for mapping the genome of the turkey was estimated based on the allelic polymorphism. Despite the low success rate, yet because of the huge number of available and used markers that eventually provided information for comparative mapping of chicken and turkey markers, researchers suggested that 20% of available chicken microsatellite markers could be used for mapping the turkey genome (Reed et al. 2000).

Reed and co-workers (2003) used about 800 microsatellite markers of domestic chicken in the turkey, about 10% of them in the creation of its genetic map. Most of the markers that were not used in mapping found application in research on the origin analysis (Donoghue et al. 1999) and the assessment of genetic variability of the population (Mock et al. 2002).

In total, 277 sets of primers were developed and tested on turkey DNA (Burt et al. 2003), which resulted in obtaining over 700 turkey microsatellite markers. The integrated turkey map comprised 613 *loci* arranged in 41 linkage groups, 19 markers were not linked. In total, 210 previously unmapped markers were used to create an integrated map. Overall map distance

contained within the 41 linkage groups is 3,365 cM (sex-averaged) with the largest linkage group (94 *loci*) measuring 533.1 cM. Average marker interval for the map was 7.86 cM. Significant similarity was obtained for 95.6% of the turkey sequence including about 91% of the chicken genome (Reed et al. 2007). In conclusion, it should be clearly stated that even if the number of chicken microsatellite sequences that can be amplified and polymorphic in turkeys is small, they may be important genetic markers (Reed et al. 2000).

Study by Aslam and collaborators (2010) included eighteen full sibling families (1008 birds) that were genotyped and 775 single-nucleotide polymorphisms (SNPs) were found, of which 570 were informative and used to construct turkey linkage map. The total distance between the linkage groups was 2,324 cM, while the highest observed distance between the groups (81 *loci*) was 326 cM. The average range of the marker in the 28 linkage groups was 4.6 cM.

Comparative mapping of turkey and chicken genomes revealed two intra- and 57 extra-chromosome rearrangements between the two species. The published map had more even distribution of markers as it was fully based on SNP markers. In 2010, Dalloul et al. (2010) identified 600,000 SNPs in the turkey genome. Two years later Aslam et al. (2012) found 5.49 million SNPs in 11 turkey lines, including 7 commercial lines. In the following years, researches on the genome of turkey were progressing extremely fast. Dalloul and colleagues (2014) in their research combined two Nextgeneration Sequencing Platforms (Roche / 454 and Illumina GAI) and maps of the artificial bacterial chromosome. The first map of the 2.01 turkey genome included 89% of the genome sequence. The contig sequences were assigned to 30 of 40 chromosomes, whereas about 10% of the assembled sequences corresponded to unassigned chromosomes (ChrUn) (Dalloul et al. 2014). Recent reports indicate that more than 95% of the turkey genome has already been identified GCA_000146605.3 (The National Center for Biotechnology Informations, [https://www.ncbi.nlm.nih.gov/genome/?term=txid9103\[orgn\]](https://www.ncbi.nlm.nih.gov/genome/?term=txid9103[orgn])).

QUANTITATIVE TRAITS MAPPING (QTLs)

QTL mapping is based on using genetic maps, and then the relationship between the inheritance of polymorphic DNA markers and the level of the examined feature is determined. The relation between the genotype and the phenotype may be helpful in identifying a *locus*, which is close to the genetic marker with which the linkage occurred. The final result of QTL detection is the recognition of genetic markers located near QTL, which increases the probability that a particular marker and QTL are inherited together. This phenomenon is known as linkage disequilibrium (Dekkers 2004).

With the use of QTL it is sometimes possible to define a gene or genes encoding a given trait. However, it is not easy due to diverse expression of genes, epistasis, pleiotropy or polygenic inheritance. For these reasons each QTL analysis should be checked in an independent population and precision mapping should be used to locate the gene more precisely (Spelman and Bovenhuis 1998).

Research on mapping the genes of quantitative traits was parallel to the better understanding the domestic hen (*Gallus gallus domesticus*) genome. The total chicken QTL, published on December 21, 2018 at ChickenQTLdb (Chicken QTL Data Base, <http://www.animalgenome.org/cgi-bin/QTLdb/GG/browse>), was 8,427. 6,117 QTL were

related to production traits, 1321 to habit, 712 to health, 209 to physiology and 68 to reproductive traits. They are grouped according to the chromosome including QTL data or based on the phenotypic feature to which QTL is associated. QTL mapping was applied to several production features as: growth traits (3584 QTLs), egg quality (775 QTLs), egg production (207 QTLs), meat quality (105 QTLs), disease resistance (621 QTLs), behavioural traits (320 QTLs) etc. (Chicken QTL Data Base, <http://www.animalgenome.org/cgi-bin/QTLdb/GG/browse>, 2018).

Microsatellite sequences originating from the chicken genome were used in studies on understanding the genome of turkey (Levin et al. 1995; Huang et al. 1999; Reed et al. 1999; Reed et al. 2000; Burt et al. 2003; Dranchak et al. 2003; Reed et al. 2003; Knutson et al. 2004; Reed et al. 2005; Chaves et al. 2006).

DNA MICROARRAYS

SNP genotyping is used to study candidate genes polymorphisms and functional mutations. Current experimental methods are based on a DNA chip technology based on fluorescent labelling of genomic DNA that hybridizes to a number of specific oligonucleotides on a solid surface such as glass. A chip containing 3000 SNPs was created for the chicken genome (Muir et al. 2008).

A microarray with a capacity of 60000 SNP markers has already been created and it covers all chicken chromosomes. Results were used to increase the chicken genetic map resolution. In the fast, efficient and simultaneous determination of numerous variants of genetic markers, High-Throughput Genotyping, it was possible to use them in the Genome-Wide Marker Assisted Selection, i.e. a marker-assisted genomic selection. In case of laying hens, the ideal size of the microarrays is 60 K, or 60000 SNP markers, while for research on broilers it has been increased to 300 K (Smith et al. 2006; Gheyas and Burt 2013).

For research on the growth and development of turkeys at the level of muscle transcriptomes, a 6K oligonucleotide marker was developed. Skeletal muscle samples were taken at three critical stages of muscle development: 18 days of embryonic development, 1 day after hatching and 16 weeks of age. The oligonucleotides were designed from the sequences obtained from skeletal muscle cDNA libraries from three developmental stages. Mismatched and encoded sequences have proven useful for controlling specific hybridization for most genes (Sporer et al. 2011).

A huge advance in the development of functional genomics was possible with the use of genomic sequencing, as well as the EST (Expressed Sequence Tags) sequence. The aim of the analyses was to examine the expression of genes, which is based on microarrays known as expression microarrays. Similarly, point mutations were determined by DNA microarrays in genomic DNA.

CANDIDATE GENES

Studies on genes significantly influencing the level of production and immune traits may have practical applications in breeding. This is possible due to the use of the candidate genes analysis. Its methodology is based on selecting positional and functional genes and

establishing the relationship between genes and the phenotypic level of the traits. The gene-phenotype interaction is examined based on the function of the protein product of a given gene in metabolism or by determining the location of the gene in the area of quantitative trait loci. The analysis of candidate genes has been extended by *in silico* method (Digital Candidate Gene Approach). The genes were searched for in databases that collect information about genes and their functions in the body (Zhu and Zhao 2007). The next step was to find the gene polymorphism in the reference population, with the use, among others, genotyping SNP markers or sequencing and the last – statistical analyses of the association between genetic polymorphism and the level of the analysed quantitative trait. If a positive result of the association between a marker and a functional trait is obtained, then it can be used in Gene Assisted Selection (Kwon and Goate 2000).

CONCLUSIONS

Both domestic chicken and wild turkey belong to the same family – *Phasianidae*. Their common origin is confirmed by phylogenetic studies, which shown that both species have the same ancestor. Many analyses and research on the genome of turkey were based on databases, microsatellite sequences and methods used in research on the sequence of chicken genome. Numerous analyses confirmed the high degree of similarity between the genomes of both species. It is predicted that in the next years scientists will focus more on re-sequencing of entire genomes or parts of them and will abandon methods of genetic markers genotyping. In this way all possible polymorphisms in the genome, identified based on individual nucleotide sequences, will be used. In the future, the possibility of re-sequencing of entire genomes at the appropriate reference DNA sequence level will exist. Re-sequencing the genome is based on sequencing short fragments of genetic material, aligning them to the reference genome and then extracting genetic variants.

This technique allows to obtain information on variants stored in databases (mainly concerning coding regions), as well as on new variants and variants located outside coding regions, which are not possible to be detected by array methods or exome sequencing. The advantages of using NGS in genome re-sequencing are: various methods of results analysis (detection of variants, annotations, and personalized analysis), detection of new genetic variants, advantage over arrays methods, detection of somatic mutations, shorter testing time and lower cost compared to *de novo* sequencing. And thanks to the third generation of sequencers, it is possible to sequence DNA sections of several billion base pairs (Gbp) in a very short period of time. The bioinformatics analysis includes: basic analysis of sequencing data, alignment, assembly and coverage statistics, variants detection, annotations, and statistics. Storing such a large amount of data and computer analysis of them create further technical difficulties.

REFERENCES

- Aslam M.L., Bastiaansen J.W., Crooijmans R.P., Vereijken A., Megens H.J., Zieloni M.A.** 2010. A SNP based linkage map of the turkey genome reveals multiple intrachromosomal rearrangements between turkey and chicken genomes. *BMC Genomics* 11, 674.

- Aslam M.L., Bastiaansen J., Elferink M., Schroeder S.G., Crooijmans R., Megens H.J., Sonstegard T., Fleischer R.C., Blomberg L., Van Tassell C., Groenen M., Long J.** 2012. Selective Sweep Analysis in Turkey Using Whole Genome Sequence Data. *Plant Anim. Gen.* 20, 14–18.
- Burt D.W., Morrice D.R., Sewalem A., Smith J., Paton I.R., Smith E.J., Bentley J., Hocking P.M.** 2003. Preliminary linkage map of the Turkey (*Meleagris gallopavo*) based on microsatellite markers. *Anim. Genet.* 34, 399–409.
- Chaves L.D., Knutson T.P., Krueth S.B., Reed K.M.** 2006. Using the chicken genome sequence in the development and mapping of genetic markers in the turkey (*Meleagris gallopavo*). *Anim. Genet.* 37, 130–138.
- Chicken QTL Data Base**, <http://www.animalgenome.org/cgi-bin/QTLdb/GG/browse>, access: 08.03.2018.
- Dalloul R.A., Long J.A., Zimin A.V., Aslam L., Beal K., Blomberg L.A., Bouffard P., Burt D.W., Crasta O., Crooijmans R.P., Cooper K., Coulombe R.A., De S., Delany M.E., Dodgson J.B., Dong J.J., Evans C., Frederickson K.M., Reed K.M.** 2010. Multi-Platform next-generation sequencing of the domestic turkey (*Meleagris gallopavo*). *Genome assembly and analysis. PLoS Biol.* 8(9), e1000475.
- Dalloul R.A., Zimin A.Z., Settlege R.E., Kim S., Reed K.M.** 2014. Next generation sequencing strategies for characterizing the turkey genome. *Poultry Sci.* 93, 479–484.
- Dekkers J.C.M.** 2004. Commercial application of marker and gene selection in livestock: strategies and lessons. *J. Anim. Sci.* 82, 313–328.
- Donoghue A.M., Sonstegard T.S., King L.M., Smith E.J., Burt D.W.** 1999. Turkey sperm mobility influences paternity in the context of competitive fertilization. *Biol. Reprod.* 61, 422–427.
- Dranchak P.K., Chaves L.D., Rowe J.A., Reed K.M.** 2003. Turkey microsatellite loci from an embryonic cDNA library. *Poultry Sci.* 82, 526–531.
- Gheyas A.A., Burt D.W.** 2013. Microarray resources for genetic and genomic studies in chicken: a review. *Genesis* 51(5), 337–356.
- Gruszczyńska J., Grzegorzóka B., Morawska A.** 2018. Genome comparative studies in the phasianidae family. Part I. Genome comparative studies in the *Phasianidae* family between domestic chicken (*Gallus gallus domesticus*) and wild turkey (*Meleagris gallopavo*). *Folia Pomer. Univ. Technol. Stetin., Agric., Aliment., Pisc. Zootech.* 345(48)4, 55–64.
- Huang H.B., Song Y.Q., Hsei M., Zahorchak R., Chiu J., Teuscher C., Smith E.J.** 1999. Development and characterization of genetic mapping resources for the turkey (*Meleagris gallopavo*). *J. Hered.* 90, 240–242.
- Knutson T.P., Chaves L.D., Hall M.K., Reed K.M.** 2004. One hundred fifty-four genetic markers for the turkey (*Meleagris gallopavo*). *Genome* 47, 1015–1028.
- Kwon J.M., Goate A.M.** 2000. The candidate gene approach. *Alcohol Res. Health* 24, 164–168.
- Levin-Cheng H.H., Baxter-Jones C., Hillel J.** 1995. Turkey microsatellite DNA loci amplified by chicken-specific primers. *Anim. Genet.* 26, 107–110.
- Liu Z., Crooijmans R.P.M.A., Poel J.J. van der, Groenen M.A.M.** 1996. Use of chicken microsatellite markers in turkey: a pessimist view. *Anim. Genet.* 27, 191–193.
- Mock K.E., Theimer T.C., Rhodes O.E.J., Greenberg D.L., Keim P.** 2002. Genetic variation across the historical range of the wild turkey (*Meleagris gallopavo*). *Mol. Ecol.* 11, 643–657.
- Muir W.M., Wong G.K., Zhang Y., Wang J., Groenen M.A., Crooijmans R.P., Megens H., Zhang H.J., Okimoto R., Vereijken A., Jungerius A., Albers G.A., Lawely C.T., Delany M.E., MacEachern S., Cheng H.H.** 2008. Genome-wide assessment of worldwide chicken SNP genetic diversity indicates significant absence of rare alleles in commercial breeds. *Proc. Natl. Acad. Sci. USA* 105, 17312–17317.
- Reed K.M., Mendoza K.M., Beattie C.W.** 1999. Utility of chicken-specific microsatellite primers for mapping the turkey genome. *Anim. Biotech.* 10, 137–141.

- Reed K.M., Mendoza K.M., Beattie C.W.** 2000. Comparative analysis of microsatellite loci in chicken and turkey. *Genome* 43, 796–802.
- Reed K.M., Harry D.E., Zaitlin D., Marini P.J.** 2003. A first-generation map of the turkey genome. *Genome* 46, 914–24.
- Reed K.M., Chaves L.D., Hall M.K., Knutson T.P., Harry D.E.** 2005. A comparative genetic map of the turkey genome. *Cytogen. Genome Res.* 111, 118–127.
- Reed K.M., Chaves L.D., Mendoza K.M.** 2007. An integrated and comparative genetic map of the turkey genome. *Cytogen. Genome Res.* 119, 113–126.
- Smith J., Speed D., Hocking P.M., Talbot R.T., Degen W.G., Schijns V.E., Glass E.J., Burt D.W.** 2006. Development of a chicken 5 K microarray target towards immune function. *BMC Genomics* 7, 49.
- Spelman R.J., Bovenhuis H.** 1998. Moving from QTL experimental results to the utilization QTL in breeding Programmes. *Anim. Genet.* 29, 77–84.
- Sporer K.R., Chiang W., Tempelman R.J., Ernst C.W., Reed K.M., Velleman S.G., Strasburg G.M.** 2011. Characterization of a 6 K oligonucleotide turkey skeletal muscle microarray. *Anim. Genet.* 42, 75–82.
- The National Center for Biotechnology Information**, [https://www.ncbi.nlm.nih.gov/genome/?term=txid9103\[orgn\]](https://www.ncbi.nlm.nih.gov/genome/?term=txid9103[orgn]), access: 08.03.2018.
- Zhu M., Zhao S.** 2007. Candidate gene identification approach: progress and challenges. *Int. J. Biol. Sci.* 3, 420–427.

BADANIA PORÓWNAWCZE GENOMÓW W RODZINIE *Phasianidae*. CZĘŚĆ II. AMPLIFIKACJA MIĘDZYGATUNKOWA W RODZINIE *Phasianidae* NA PRZYKŁADZIE KURY DOMOWEJ (*Gallus gallus domesticus*) I INDYKA ZWYCZAJNEGO (*Meleagris gallopavo*)

Streszczenie. W badaniach, których celem było poznanie genomu indyka, wielkim ułatwieniem było zastosowanie mikrosatelitarnych markerów specyficznych dla kury. Takie rozwiązanie było możliwe ze względu na fakt, iż kura domowa i indyk zwyczajny są gatunkami spokrewnionymi filogenetycznie, posiadającymi wspólnego przodka. W wielu analizach z wykorzystaniem mikrosatelitarnych markerów kury uzyskano pozytywne wyniki amplifikacji sekwencji indyka, co przyspieszyło poznanie genomu tego gatunku. Poznanie genomu indyka nastąpiło poprzez zastosowanie: sekwencjonowania, mikromacierzy DNA, mikromacierzy ekspresyjnych, identyfikacji QTL i genów kandydujących, a także poprzez resekwencjonowanie.

Słowa kluczowe: kura domowa, indyk zwyczajny, sekwencje mikrosatelitarne, QTL, mikromacierz ekspresyjna, mikromacierz DNA, geny kandydujące, resekwencjonowanie.

