Abschließender Sachstandsbericht/Final Report Leibniz-Wettbewerb/Leibniz Competition

Titel/Title: High Fertility (SOS-FERT)

Projektnummer/Project Number: K52/2017

Berichtszeitraum/Reporting Period: Mai 2018-April 2023

Federführendes Leibniz-Institut / Lead Leibniz Institute:
Forschungsinstitut für Nutztierbiologie, vormals Leibniz-Institut für Nutztierbiologie (FBN) Dummerstorf
Research Institute for Farm Animal Biology, formerly Leibniz Institute for Farm Animal Biology (FBN)

Projektleiter*in / Spokesperson:

Prof. Dr. Jennifer Schön [seit Juni 2021 Leibniz-Institut für Zoo-und Wildtierforschung (IZW) Berlin, sowie seit Januar 2022 Technische Universität (TU) Berlin]
Prof. Dr. Jennifer Schön [since June 2021 Leibniz Institute für Zoo and Wildlife Research (IZW) Berlin, and since January 2022 Technische Universität (TU) Berlin]

Executive Summary

The complex network of mechanisms determining reproductive success is largely unknown. To decipher parts of this network we used worldwide unique mouse lines created by long-term selection (>200 generations in 2023) for the integrative fertility traits “increased litter size and weight”. During the selection process, the fertility mouse lines have doubled their litter size without reduction of pup birth weight. In the SOS-FERT project, we applied next generation sequencing (NGS) and bioinformatics tools to identify distinct genomic footprints of selection, the so-called “signatures of selection”, and extracted the most likely causative alleles for the selected reproductive traits in these mouse lines. These signatures were tested for biological relevance in mice by conducting a selection experiment in the unselected control line and by searching for similar signatures in comparative animal species using NGS data from individuals with divergent litter sizes. Furthermore, we evaluated the impact of selection for a primarily female trait (high litter size) on male reproduction and assessed the significance of the identified selection signatures in males with divergent reproductive performance.

NGS data was generated from all the six FBN mouse lines, which originate from the same founder population created in the late 1960s. These include the two fertility lines (FL1/2) and the unselected control line (UCL). In addition, individuals from three lines selected for non-fertility traits (body mass, protein mass and endurance) were sequenced for contrasting purposes (in total N=150 genomes). Regions of extended homozygosity and high genetic differentiation were extracted and revealed that the FBN selection lines have low within-line genetic diversity, and are highly differentiated among lines. Patterns of homozygosity, however, were widespread across the genome, indicating that both selection and genetic drift have left genomic traces that are now difficult to distinguish in the absence of genomic data from the founder population or generations early in the selection process. Nevertheless, a stringent approach to detect line-specific regions of high genetic differentiation and structural variation allowed the identification of multiple candidate genes for the selected traits. Candidate genes in the fertility lines likely affect the reproductive process on multiple levels such as folliculogenesis, placentation, sex steroid signaling and milk composition. Both Dummerstorf fertility lines only share extremely few common regions of genomic differentiation, substantiating the complex polygenic nature of the trait “high litter size”. A genomic selection experiment in the unselected control line led to the accumulation of fertility line-specific alleles in the established UCL sublines. Biological relevance of these alleles in terms of an increase in litter size in the generated sublines could, however, not be proven.

In the comparative animal model, the pig, the consortium generated own NGS data from a maternal (relatively high litter size) and a paternal (relatively low litter size) breed (n=20 genomes). Additional published sequencing data from maternal and paternal breeds were integrated into the analysis pipeline for validation purposes. Analogous to the two fertility mouse lines, the maternal “high fertility” breeds have very few similarities in their genomic differentiation despite similar selection criteria. Nevertheless, this comparative approach, while again demonstrating the multi-dimensional genetic control of the trait “high litter size”, yielded another fertility-related candidate gene involved in trophoblast invasion, which is also amongst the candidate genes in one of the Dummerstorf mouse lines.

Male phenotype assessment revealed that breeding for the primarily female trait "high litter size" does not seem to be clearly beneficial for male reproductive performance, but it does not lead to obvious tradeoffs in the tested male fertility traits either.

The project allowed the consortium to be complemented with international young female and male scientists. The cooperation between the SOS-FERT partner institutions has withstood both the difficulties posed by the pandemic and the restructuring processes in the lead institute, as documented by several peer-reviewed articles (primarily open access) and conference contributions.

Content

1. Achievement of objectives and implementation of milestones

The Leibniz Institute for Farm Animal Biology (FBN) has established two independent and worldwide unique fertility mouse lines (FL1, FL2), selected for increased litter size. After more than 180 generations of selection, the fertility mouse lines have doubled their litter size without reduction of pup birth weight. In the SOS-FERT project, we applied next generation sequencing and bioinformatics tools to identify the distinct genomic footprints of selection (signatures of selection; WP1), and extracted the most likely causative alleles for the selected reproductive traits in these mouse lines. These signatures were evaluated for biological relevance in mice by conducting a selection experiment in the unselected control line (UCL; WP2). Furthermore, a comparative approach was planned to search for similar signatures in other animal species (pig and lion) using NGS data from individuals with divergent litter sizes (WP4, 5, 6). We also investigated the impact of selection for a primarily female trait (high litter size) on male reproductive performance (WP3).

With this workflow we aimed at testing the following three main hypotheses:

1) High performance in the reproductive trait 'litter size' depends on selection for specific genomic regions or particular patterns of impacted genes and pathways.

The FBN members of the consortium meticulously reconstructed the breeding history of the Dummerstorf mouse lines (starting in 1969) and generated NGS data from all five FBN long-term selection mouse lines and the unselected control line. Regions of extended homozygosity and high genetic differentiation were extracted and revealed line-specific patterns of genetic variation among lines, as well as high levels of homozygosity within lines (Milestone [M] 1, table 1). This high degree of distinctiveness results from the combined effects of long-term continuous selection, genetic drift, a population bottleneck and isolation. Detection of line-specific patterns of genetic differentiation (Regions of Distinct Differentiation, RDDs) and structural variation revealed multiple candidate genes behind the improvement of the selected traits. Despite common selective pressures, the two fertility lines (selected for the integrative trait 'high litter size') showed a high degree of distinctiveness and very few common RDDs. Candidate genes in each of the fertility lines likely affect the reproductive process on multiple levels such as folliculogenesis, placentation, sex steroid signaling and milk composition.

The genomic signatures detected in the fertility mouse lines did not allow reliable deduction of trait-specific tagSNPs (M2, table 1). The consortium therefore used a commercial SNP array for the genomic selection experiment in the UCL based on the RDDs of the fertility lines. UCL-sublines were generated (M3, table 1) and subsequent genotyping revealed an accumulation of fertility line-specific alleles within the sublines. However, up to the F2 generation, the selected trait (litter size) was not affected by the genomic selection approach.

2) Long-term selection for the reproductive trait 'litter size' yielded a genotype causal for a male phenotype with elevated reproductive performance.

The consortium characterized testicular and sperm physiology parameters in the Dummerstorf mouse lines. The selection for the primarily female trait litter size did neither significantly alter sperm morphology nor their mitochondrial activity. However, in comparison to the UCL, the fertility lines showed certain differences in their testicular phenotype and sperm motility parameters after stress tests (M4, table 1). IVF trials were conducted to further evaluate sperm from both fertility lines, but did not show significant alteration of sperm fertilizing capacity in the fertility lines (M5, table 1).

3) Genomic patterns identified to be causal for male and female traits of (high) reproductive performance in the mouse lines FL1 and FL2 are also causal for (high) reproductive performance in other mammalian species

By using the established pipelines for mouse samples, NGS data from two pig breeds (Pietrain sire line with relatively low and Large White dam line with relatively high litter size) were generated and analyzed for selection signatures. However, breed-specific genes overlapping runs of homozygosity were not significantly enriched for biological functions related to reproduction. The consortium therefore carried out a comparative genomic analysis, which included own and published sequencing data of two sire and two dam lines. Consistent with the results obtained in the two fertility mouse lines, we detected only few similarities in the genomic differentiation of the maternal breeds (dam lines) despite comparable selective

pressures. Nevertheless, this comparative approach yielded one candidate gene with potential relevance for genetic control of reproduction which was also detected in one of the murine fertility lines. As no tagSNPs could be developed from the murine sequencing approach, genotyping of males with divergent sperm quality parameters as well as females with high and low litter sizes (M7, table 1) was accomplished using a commercial SNP array. Genome-wide association analyses investigating the genetic variability underlying phenotypic variations of semen quality traits in sire line boars revealed putative candidate genes with functional relationships to spermatogenesis, sperm motility and sperm resistance to environmental stress.
The phenotypic group assembly for the second comparative animal model (lion) was severely affected by the travel restrictions during the Corona pandemic and unexpected difficulties in retrieving reliable fertility data from zoos for the large sample collection already existing at the IZW. Thus, this part of the project (M8, M9, table 1) could not be carried out during the funding period.

Table 1: Anticipated milestones (milestones [M] that have been reached within the funding period are shown in bold):
M1. Signatures of Selection analyzed (FL1/ FL2)
M2. tagSNP-chip created (mouse, for selection experiment)
M3. UCL sub-lines assembled
M4. Male phenotypes characterized (FL1/ FL2/ UCL)
M5. Male phenotypes evaluated (FL1/ FL2/ UCL)
M6. tagSNP-chip created (pig, genotyping individuals with divergent reprod. preformance)
M7. Sampling finalized (porcine dam and sire lines, male and female)
M8. Phenotypic group assembly finalized (lion)
M9. NGS Data analyzed (lion)

Adherence to original financial planning
The initial financial plan had to be adjusted due to moderate changes in the project plan. To capture the variability still present in the selection mouse lines, the number of sequenced animals had to be more than doubled, which consequently led to an increase in sequencing costs (despite lower coverage in the additional sequencing samples and overall decreasing prices for sequencing). This was compensated for by using commercial SNP arrays instead of customized tagSNP-chips for the genotyping approaches and by partly omitting the activities planned for the second comparative animal model (lion). The remaining budget originally allotted to this WP (6) was re-transferred to the Leibniz-association. Further adjustments had to be made due to several changes in the composition of the project team. Both PhD students dropped out of the project during the Corona crisis, one because of personal reasons, and the other one because of a severe illness. The consortium acted against these setbacks by interim funding of young scientists who were involved in related projects of the cooperation partners.

2. Activities and challenges

Detection of signatures of selection in the FBN mouse lines: Biological samples of initially ten representative animals of each of the six FBN mouse lines (five selection lines and one unselected control line) were provided and whole genome sequencing (60 individuals, 30×· coverage) was carried out by the Institute of Clinical Molecular Biology (IKMB) Kiel. To better represent the remaining diversity available in the breeding populations, the consortium then decided to increase the sample size from 10 individuals per mouse line to 25 individuals by adding 15 individuals, but sequencing them at lower target coverage (5× vs. 30×). The decision to increase the number of samples for the mouse lines was supported by all involved partners and was also realized by IKMB. Thus, the full mouse data set now comprises 150 (60 initial +90 additional) genomes. Next to generating the genomic data, the FBN partners also reconstructed the breeding history of the Dummerstorf mouse lines (starting 1969) and have finally made this information available to the scientific community (DOI: 10.1186/s12915-022-01248-9). The FBN partners subsequently performed a genomic selection experiment in the unselected control line based on the regions of distinct differentiation (RDDs) detected in the

two fertility lines. Ear notch samples of 456 UCL mice from 125 litters were taken for DNA extraction. Genotyping was performed using the GGP Mouse (GIGA-MUGA) SNP array (Illumina). Although an accumulation of fertility line-specific alleles could be achieved in the established subpopulations, this did not lead to an increase in litter size in these populations (up to F2 generation). The biological relevance of the discovered alleles could thus not be conclusively demonstrated with the chosen experimental approach.

Detection of signatures of selection and tagSNPs in porcine dam and sire lines: The Bundeshybridzuchtprogramm (BHZP) provided sample material of 10 individuals each from a porcine dam and sire line. Samples were processed at FBN and sequenced (50× coverage) by IKMB Kiel. Subsequently, a similar data processing pipeline as for the mouse data was applied at FBN. Furthermore, the FBN integrated additional porcine sequencing data from other published studies to conduct a validation approach.

Characterization of murine and porcine male phenotypes: The FBN analyzed morphology, motility and mitochondrial activity of epididymal sperm of 20 individuals from the fertility lines (FL1, FL2) and the unselected control line (UCL). Even though the fertility line males show divergent testicular and spermatological phenotypes compared to the control line, there are neither signs for a clearly detrimental nor for a clearly beneficial effect of the selection process on sperm quality. The functional evaluation of male fertilizing capacity via competitive in vitro fertilization did also not reveal significant differences. For this part of the project, the FBN established the complete pipeline of embryo production, gamete/embryo cryopreservation and embryo transfer, which allowed the institute to establish a viable genetic reserve of their valuable mouse lines. The IFN Schoenow systematically phenotyped porcine males from a BHZP sire line (spermatological characterization of 300 ejaculates from 150 boars). Genome-wide association analyses using a Porcine SNP60k BeadChip (Illumina) revealed genomic regions that explain part of the phenotypic variation in the investigated sperm quality traits. The ejaculates were subsequently used for routine artificial insemination, and IFN and BHZP collected field fertility data (pregnancy rates, litter size) for all boars under investigation (>1500 corresponding insemination records). However, boar and semen related factors only explained 7% of the total variation in the number of live born piglets. Due to the high value of the generated and stored semen samples, for which fertility data are known, further investigations are currently carried out by IZW and IFN. They are focused on the ability of semen to cope with temperature stress, probably on the basis of a better endowment for antioxidative defense.

Identification of individuals with outstandingly high and low reproductive performance in lion populations: The Institute for Zoo-and Wildlife Research (IZW) holds a large collection of samples from male and female African lions. However, for these animals no information on reproductive performance was available. In addition, in order to avoid assignment of lionesses by chance to the group of high fertility, the search was out for females that had given birth to a larger litter (3-4 cubs) more than once. Unfortunately, retrieval of reliable information regarding the breeding success of lionesses (initially thought to be easily available) has proven more difficult than expected. Supplementary approaches of collecting new (additional) samples from low/high fertility African lionesses from zoos or using the zoo population of female Asiatic lions for the study (this population is managed by an official studbook that contains the necessary fertility data) failed because sampling was (and still is) for animal welfare reasons only possible during indispensable veterinary interventions. The frequency of such events was finally too low for the project. For the investigations on the male side, Geolifes and the IZW have collected samples of ten male African lions undergoing fertility examinations as part of breeding programs in South Africa. The pandemic, however, disrupted the original sampling plan and ultimately only one male with repeatedly low sperm quality could be identified within the tested population and period. Therefore, the consortium mutually decided to cancel the work on the second comparative animal model.

3. Results and achievements

Publications: The main results of the project have been published in four peer-reviewed articles so far. One more manuscript is currently under review and one in preparation.

Congresses, symposia and workshops: Results of the porcine male phenotyping were presented at the 53. Annual Meeting for Physiology and Pathology of Reproduction in Rostock, Germany in February 2020 (oral). In 2022, two posters summarizing the findings derived from pig NGS data analysis were presented at the symposium “Visions III: Star Gazing into the Galaxy of Animal Genetics and Genomics” in Ames, Iowa, USA (extended abstract) and at the International Congress on Animal Reproduction (ICAR 2020+2) in Bologna, Italy. The progress of the project was furthermore presented at several internal workshops. The complete consortium conducted annual symposia with presentations on the current work packages (hosts: FBN, IZW, IFN). During and after the Corona pandemic the symposia were held online.

4. Equal opportunities, career development and internationalization

For the three positions embedded within the SOS-FERT project (one Postdoc position, two PhD candidates) young researchers from Chile, Iran, and the Czech Republic could be recruited and hired. The two PhD positions were filled by female graduate researchers and the Postdoc position by a male researcher with project-relevant experiences. The PhD candidates were enrolled in the graduate School “Biomedical Sciences”, which is part of the “Dahlem Research School” of the Freie Universität Berlin.

During the first peak of the Corona crises in 2020, one of the PhD students requested the termination of her contract at very short notice for personal reasons. We complied with her request and were able to immediately recruit and hire a young female bioinformatician from Croatia to further support the NGS data analysis of murine samples. The results generated for the SOS-FERT project became an integral part of her PhD thesis, which she successfully defended in February 2023. The other PhD student working on the project developed severe health problems at the end of 2020 and was finally diagnosed with cancer in 2021. The young woman died in August 2022, a tragedy for her family, friends and colleagues, and also a severe setback for the continuation of the project. The work originally designated to this PhD candidate was partly rescued by short-term interim financing of two young scientists working on other related projects of the cooperation partners (one female from Brazil/IFN, one male from Egypt/FBN). Both the central postdoc and the PhD student who defended her thesis within SOS-FERT gathered valuable experiences and education during the funding period and have already transitioned into specialized new positions with a focus on genomics and bioinformatics.

5. Structures and cooperation

The SOS-FERT consortium remained stable throughout the prolonged funding period and was characterized by straightforward, loyal cooperation. After a kick-off meeting in February 2018 we met in small project sub-groups every 3-6 months. Organizational and administrative challenges for the progress of the project arose, however, due to the demanding and extensive restructuring process of the FBN following its exclusion from the Leibniz Association in 2020 and the transfer of the project spokesperson (Jennifer Schön) from the FBN Dummerstorf to the IZW Berlin / TU Berlin in 2021.

6. Quality management

General measures of quality management: All participating project partners have installed measures to secure the quality of work processes and scientific results in research based on the “Principles for Safeguarding Good Scientific Practice” of the DFG. All employees of the project received an obligatory individual training on the principles of “Good Scientific Practice”

with a refreshment training for all scientists once a year. The obligatory FBN standard operation procedures (SOPs) on topics such as “keeping a good laboratory notebook”, “digital storing of research data”, “handing over data and important materials”, “anti-plagiarism” etc. are familiar to all members of the working group. The PIs regularly monitored and controlled the laboratory work as well as the documentation of the analog and digital research data (SOP "Leading a working group according to the rules of good scientific practice") and all results were critically discussed in the regular meetings of the consortium. Primary and metadata are stored in folders with daily backups at a second location

Data availability and open access: The data generated in the SOS-FERT project have been presented at national and international conferences and are or will be published in peer-reviewed journals. The datasets generated and/or analyzed during the study are included in the published articles or they are available on reasonable request. All codes to reproduce the analysis from the raw sequencing reads (the most primal state of the analysis pipeline) have been made public along with the publications and can be found on the „GitHub“-platform. Furthermore, sequencing reads already have been and will continue to be made available for the scientific community along with the corresponding publications.

Animal experiments and animal protection: Animal experiments have been carried out for the WPs 1, 2, and 3 and included mice from the FBN’s own breeding program (long-term selection lines and unselected control line). The killing of the animals for scientific purposes according to § 4 paragraph 3 TierSchG was reported to the local animal protection authorities. The housing and experimental conditions in the model animal laboratory of the Leibniz Institute for Farm Animal Biology meet the requirements of the German Animal Welfare Act and the Animal Welfare Ordinance (TierSchVersV). The permission to breed and keep the Dummerstorf long-term selection lines was granted by the district administrator of the Rostock district in June 2014 with the permission according to § 11 para. 1 No. 1a Animal Protection Act for the "Laboratory for innovative farm animal models (LIN)". The animals are kept in a specific-pathogen-free husbandry system (SPF) with a defined hygiene status according to FELASA recommendations, which corresponds to the international standards for the species-appropriate housing of laboratory mice. The cage stocking density complies with the legal requirements of the European Laboratory Animal Regulation (Directive 2010/63/EU).

7. Additional own resources

The project partners invested personnel costs not covered by the project budget, which sums up to approximately 330000€. Additionally, the FBN financed a project-relevant computer assisted sperm analysis system that was initially included in the project budget, but had to be excluded due to the general budget reduction of 3.4% for all projects in 2017.

8. Outlook

The whole-genome sequencing data obtained for the worldwide unique FBN mouse lines are a valuable resource which will be and is already used to further explore the genetic and functional diversity of these exceptional animal models and their line-specific phenotypes (e.g. 10.1038/s42003-022-03339-3). As it proved particularly difficult to disentangle differentiation caused by selection from that caused by genetic drift in the long-term selected mice, other strategies (e.g. inter-crossing of the fertility lines) are considered to verify the RDDs associated with the fertility traits. The data obtained from the pig as a comparative animal model will be used further to ultimately enable the implementation of fertility traits in breeding programs, as far as this is achievable. It will be of utmost importance to evaluate how the identified genetic features can affect gene expression and its regulation in the reproductive organs via genetic (regulatory sequences), epigenetic or post-transcriptional mechanisms. This is part of the worldwide efforts to functionally annotate DNA sequence information to understand tissue-/organ-specific cellular traits.