

# Potential of F<sub>2</sub> pig crosses: perspectives from population and quantitative genomics

M.Sc. Iulia Georgiana Blaj

1. Berichterstatter: Prof. Dr. Georg Thaller

Die aktuellen Fortschritte in der Genomanalyse beim Schwein beruhen auf der Verfügbarkeit von SNP-Panels sowie auf der genomischen Information aus der Next-Generation-Sequenzierung. Die vorliegende Arbeit untersucht das Potenzial der Nutzung von SNP-Arrays und der Gesamt-Genomsequenzierung (WGS) in vier bestehenden F<sub>2</sub>-Resourcepopulationen. Damit können innovative Fragenstellungen der Populations- und quantitative Genetik bearbeitet werden. Genetische Analysen von F<sub>2</sub>-Kreuzungen bieten die Möglichkeit der Identifizierung von Genomregionen, die für die Ausprägung quantitativer Merkmale verantwortlich sind (QTL). In der vorliegenden Arbeit wurden diesbezüglich Merkmale betrachtet, die mit dem Wachstum von Schweinen sowie mit der Schlachtkörperqualität und der Fettverteilung assoziiert sind. Die untersuchten F<sub>2</sub>-Populationen sind Kreuzungen aus europäischen Rassen (Piétrain, Landrasse, Deutsches Edelschwein) einerseits sowie Kreuzungen von genetisch weiter entfernten Rassen (Piétrain, Meishan, europäisches Wildschwein). Für die Analyse der Beziehungen von phänotypischen und genetischen Informationen wurden genomweite Assoziationsstudien (GWAS) durchgeführt. In **Kapitel 1** wurde eruiert, wie die Beobachtungen der verschiedenen F<sub>2</sub>-Populationen kombiniert und gemeinsam analysiert werden können. Dafür wurde für jede Kreuzungspopulation eine separate wie auch eine gemeinsame GWAS aller Populationen durchgeführt. Weiterhin wurde eine Meta-Analyse für drei der beschriebenen F<sub>2</sub>-Designs (mit Piétrain als gemeinsame Gründerrasse) für Wachstums- und Schlachtkörpermerkmale durchgeführt. Die Ergebnisse zeigten eine bessere Kartierung durch die Kombination aller verfügbaren Daten, wodurch die Genomregionen mit kausalen Varianten präziser bestimmt werden konnten. Dabei wurde der Einfluss von Genen, die bereits in vorangegangenen Studien als mit den Merkmalen assoziiert erachtet wurden bestätigt und es konnten belastbare Hinweise auf neue Kandidatengene detektiert werden (z.B. *BMP2* bone morphogenetic protein 2). In **Kapitel 2** wurden in den F<sub>2</sub>-Populationen Dominanz- und Imprinting-Effekte mittels Varianzkomponentenschätzung sowie verschiedener GWAS Modelle untersucht. Während die Imprinting-Effekte zwischen 0 – 19% der gesamten phänotypischen Varianz erklärten, wurden für die Dominanz-Effekte Varianzanteile von bis zu 34% geschätzt. Die Ergebnisse zeigten u.a. signifikante Assoziationen für die mit Wachstum und Fettverteilung zusammenhängenden Merkmale in der *IGF2*-Region (insulin-like growth factor 2). In **Kapitel 3** wurde eine GWAS mit den vier gepoolten F<sub>2</sub>-Populationen auf Sequenzebene durchgeführt. Hierfür wurden die SNP-Genotypen der F<sub>2</sub>-Tiere auf Basis der Sequenzierungsergebnisse von ausgewählten Gründertieren und Tieren der F<sub>1</sub>-Generation auf die Sequenzebene imputiert. Die umfassenden Analysen führten zu neuen Kandidatengenen sowie zu kausalen Mutationen, die für weitere Forschung richtungsweisend sein können. Weiterhin demonstrierte die Studie einen praktikablen Ansatz zur effizienten Nutzung von bereits etablierten experimentellen Designs. In den letzten beiden Kapiteln der vorliegenden Arbeit wurden die F<sub>1</sub>- und F<sub>0</sub> Generationen näher betrachtet. **Kapitel 4** beschäftigte sich mit der Analyse von Rekombinationen, Prozesse mit Einfluss auf die genetische Variabilität und die Evolution von Genomen. Hierfür wurden Rekombinationskarten auf Basis der F<sub>1</sub>-Generationen konstruiert. Die Anzahl an Rekombinationen variierte innerhalb der Kreuzungen, zwischen Rassen und Individuen sowie zwischen Chromosomen und innerhalb von Chromosomregionen. Obwohl über die Designs hinweg eine beträchtliche Heterogenität beobachtet wurde, existierten für Kreuzungen, Geschlechter oder Chromosomen charakteristische Muster. Dabei waren diese Muster beeinflusst von der Rekombinationsrate im Zeitverlauf, vom Ausmaß der genetischen Diversität, von der Effizienz und Richtung der Selektion sowie der Genomzusammensetzung. **Kapitel 5** nutzte die Sequenzdaten von Tieren der Gründergeneration mit einem reversen genetischen Ansatz, um Befunde aus den F<sub>2</sub>-Populationen weitergehend zu erklären. Die Grundannahme dabei ist, dass ein Großteil der genetischen Variation, die in einer F<sub>2</sub>-Generation vorhanden ist, von der Gründerpopulation weitergegeben wurde. Diese explorative Analyse zeigte, wie anhand von genomweiten Daten Rückschlüsse auf die Populationsstrukturen von Gründerrassen gezogen werden können. Außerdem wurde gezeigt, wie bioinformatische Methoden und Datenbanken zu einer wissensbasierten Auswahl von Varianten für weiterführende funktionelle Studien beitragen können.