Dissection of molecular basis on a causative mutation for ear size QTL on chromosome 7 in pigs

Dissertation
to obtain the Ph. D. degree
in the International Ph. D. Program for Agricultural Sciences in Göttingen (IPAG)
at the Faculty of Agricultural Sciences,
Georg-August-University Göttingen, Germany

presented by

Yanyu Duan

born in Nei Mongol, China

Göttingen, July, 2013
D7

1. Name of referee: Prof. Dr. Dr. Bertram Brenig

2. Name of co-referee: Prof. Dr. Christoph Knorr

Date of dissertation: July 9, 2013
To my family
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>List of publication</td>
<td>1</td>
</tr>
<tr>
<td>Abstract</td>
<td>2</td>
</tr>
<tr>
<td>Zusammenfassung</td>
<td>5</td>
</tr>
<tr>
<td>Abbreviations</td>
<td>8</td>
</tr>
<tr>
<td><strong>Chapter 1</strong> General introduction</td>
<td>11</td>
</tr>
<tr>
<td>1 External ear: evolution, development, and genetics</td>
<td>12</td>
</tr>
<tr>
<td>1.1 External ear</td>
<td>12</td>
</tr>
<tr>
<td>1.2 Growth and development of the external ear</td>
<td>13</td>
</tr>
<tr>
<td>1.3 Genetic dissection of the external ear</td>
<td>14</td>
</tr>
<tr>
<td>1.3.1 Molecular genetics for the external ear in humans</td>
<td>14</td>
</tr>
<tr>
<td>1.3.2 Animal models for the external ear</td>
<td>16</td>
</tr>
<tr>
<td>1.3.3 Molecular genetics of the external ear in domestic animals</td>
<td>17</td>
</tr>
<tr>
<td>2 PPARD function</td>
<td>19</td>
</tr>
<tr>
<td>2.1 PPARD and metabolism</td>
<td>21</td>
</tr>
<tr>
<td>2.2 PPARD and cell proliferation, differentiation, and apoptosis</td>
<td>22</td>
</tr>
<tr>
<td>2.3 PPARD and inflammation</td>
<td>24</td>
</tr>
<tr>
<td>3 The function of A/B domain in PPARs</td>
<td>25</td>
</tr>
<tr>
<td>3.1 Structure of PPARD</td>
<td>25</td>
</tr>
<tr>
<td>3.2 Modification of A/B domain regulates transcription activity</td>
<td>26</td>
</tr>
<tr>
<td>4 SOX9 function</td>
<td>29</td>
</tr>
<tr>
<td><strong>Chapter 2</strong></td>
<td>40</td>
</tr>
<tr>
<td>A missense mutation in PPARD causes a major QTL effect on ear size in pigs</td>
<td>41</td>
</tr>
<tr>
<td><strong>Chapter 3</strong></td>
<td>41</td>
</tr>
<tr>
<td>The G32E functional variant reduces activity of PPARD by nuclear export and post-translational modification to promote ear size in pigs</td>
<td>42</td>
</tr>
<tr>
<td><strong>Chapter 4</strong></td>
<td>42</td>
</tr>
<tr>
<td>An 18-bp deletion in the 5'-untranslated region of the sex determining region Y-box 9 (SOX9) gene reduces RNA stability and translation efficiency</td>
<td>43</td>
</tr>
<tr>
<td><strong>Chapter 5</strong> General discussion</td>
<td>43</td>
</tr>
<tr>
<td>1 To dissect the genetic basis of ear size in pigs</td>
<td>45</td>
</tr>
<tr>
<td>Section</td>
<td>Page</td>
</tr>
<tr>
<td>------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>1.1 Current research strategy</td>
<td>45</td>
</tr>
<tr>
<td>1.2 Progress on genetic research strategy for pigs</td>
<td>46</td>
</tr>
<tr>
<td>2 To dissect biological function of causative mutations</td>
<td>49</td>
</tr>
<tr>
<td>2.1 SNP</td>
<td>49</td>
</tr>
<tr>
<td>2.2 Chromosomal rearrangement</td>
<td>50</td>
</tr>
<tr>
<td>3 The pig is a potential biomedical model for ear size</td>
<td>51</td>
</tr>
<tr>
<td>4 Future prospect</td>
<td>51</td>
</tr>
<tr>
<td><strong>Acknowledgements</strong></td>
<td>57</td>
</tr>
<tr>
<td><strong>Curriculum Vitae</strong></td>
<td>59</td>
</tr>
</tbody>
</table>
List of publication

This thesis is based on the first three papers.


Abstract

The pinna, an important component of the external ear, first appeared in mammals. It plays a vital role in collecting sounds, radiating heat and signaling mood. In humans, the most common ear defect is microtia with prevalence rates ranging from 0.83 to 17.4 per 10,000 births. However, only a minority of genetic or environmental causes have been found to date. Domestic pigs have obvious diversity in pinna size, indicating a potential animal model for pinna deformity in humans.

To identify causative mutations underlying pinna size, we performed a genome scan in a large scale White Duroc x Erhualian pig resource family, of which Erhualian has exceptionally large and floppy pinna and White Duroc has small and erect or semi-erect pinna. We mapped a 1% genome-wide QTL for pinna size to a 2 cM confidence interval around 58 cM on chromosome 7, which explained over 40% phenotypic variance. We further fine-mapped the QTL and finally identified that a missense mutation PPARD G32E could cause such major QTL effect on pinna size in pigs. This is the first time a novel function of PPARD on controlling pinna size has been found. After that, we continued to study the function of the causative mutation at the molecular level.

We showed that the G32E substitution reduced ligand-dependent transcription activity of PPARD by transient transfection analysis. We hypothesized three underlying mechanisms, i.e. the substitution might alter (i) subcellular localization, (ii) ligand binding affinity, and/or (iii) ubiquitination according to function of the A/B domain of PPARs. To check whether the substitution altered subcellular localization, we mutated glycine 32 of wild-type PPARD into glutamic acid. Using immunofluorescence analysis, we observed that wild-type PPARD completely localized in nucleus whereas G32E mutant resided partly in cytosol in PK-15 cells and primary pinna-driven chondrocytes. To elucidate the effect of the substitution on subcellular localization, we analysed A/B domain deletion mutant and G32E mutant treated by leptomycin B, an inhibitor of CRM1 export protein. The A/B domain
deletion mutant totally localized in the nucleus, whereas cytosolic localization of the G32E mutant was blocked by leptomycin B. This indicated that G32E substitution provided A/B domain a novel function to promote CRM1-mediated nuclear PPARD export. By surface plasmon resonance technology, we detected binding affinity of GW0742 towards purified His-tagged wild-type and mutated PPARD. Equilibrium dissociation constant ($K_D$) did not show significant distinction between wild-type PPARD and G32E mutant. We also found that ubiquitination at lysine (K) 16-18 in the A/B domain of PPARD can promote ligand-induced transcription activation. The G32E substitution prevented ubiquitination of PPARD and thereby reduced transcription activity comparable with the neighbouring K16-18R mutant. In summary, the G32E substitution is a loss-of-function mutation that reduces transcription activity of PPARD by promoting nuclear export and changing non-proteolytic ubiquitination. Corresponding to this, we found that overexpression of G32E mutant led to 4.1-fold reduction of the mRNA levels of β-catenin and 40% decrease of Wnt/β-catenin signaling with TCF/LEF reporter assay compared with wild-type treatment. As we known, Wnt/β-catenin pathway inhibited chondrocyte differentiation and SOX9 expression during skeletal development. Furthermore, inactivation of SOX9 in pinna chondrocyte progenitor cells - cranial neural crest (CNC) cells resulted in a complete absence of cartilages and endochondral bones derived from the CNC. It indicated a potential mechanism underlying the G32E substitution in PPARD causing an increased pinna size in pigs.

In this thesis, we also studied an 18-bp indel in the 5’-untranslated region (UTR) of SOX9, which is an essential transcription factor in all chondrocyte development. We identified an 18-bp insertion/deletion (indel) in the 5’-UTR of SOX9 in three pig breeds—Laiwu, Bamei and Large White. The 18-bp fragment harbours a CRE half-site that can interact with a transcription factor CREB by mobility and supershift analyses. We found that the variant was located in a stable stem loop within the secondary structure of 5’-UTR using RNA-folding software. The 18-bp deletion altered the
secondary structure. By transient transfection analysis, we identified that the 18-bp was indispensable to the translation of SOX9. The deletion significantly reduced translation efficiency and caused mRNA destabilility. To determine whether the indel plays a physiological role in pinna development of pigs, we performed an association study in a White Duroc x Erhualian resource population. In this population, the indel was not associated with pinna size.
Zusammenfassung

Die Ohrmuschel stellt eine wichtige Komponente des Außenohres dar und entwickelte sich evolutionär in der Klasse der Säugetiere. Die Ohrmuschel der Tiere ist neben ihrer Funktion als Schalltrichter auch wichtig zur Wärmeableitung und als Signalgeber der optischen Kommunikation.

Die häufigste humane Ohrmuschelfehlbildung ist die Mikrotie deren Prävalenz 0,83 bis 17,4 pro 10.000 Geburten beträgt. Allerdings sind die genetischen und umweltbedingten Einflüsse auf die Ausbildung dieser Dysplasie nur unzureichend bekannt.

Die verschiedenen Hausschweinerassen zeigen eine große Diversität der Ohrmuschelgröße und bieten sich deshalb als Tiermodell für die humane Ohrmuschel-Mikrotie an.

Um kausative genetische Varianten zu identifizieren, die die Ohrmuschelgröße bei Schweinen beeinflussen, wurde zunächst eine genomweite Kopplungsanalyse an Tieren einer White Duroc x Erhualian Ressourcenfamilie durchgeführt. Während die Tiere der Rasse White Duroc kleine und aufrechte Ohren haben, sind die Ohren der chinesischen Erhualian Schweine außerordentlich groß und hängen herab. Ein Quantitative Trait Locus wurde mit einem Konfidenzbereich von 2 cM in der Region um 58 cM auf Chromosom 7 kartiert. In einer weiteren Feinkartierung konnte eine sinnverändernde Mutation im PPARD Gen identifiziert werden, die zu einem Austausch von Glycin zu Glutaminsäure an Position 32 des Proteins führt (G32E) und einen Haupteffekt auf die Entwicklung der Ohrmuschelgröße haben könnte. Ein Effekt von PPARD auf die Größenentwicklung der Ohrmuschel war jedoch zuvor nicht beschrieben, deshalb wurden weitere Experimente zur Aufklärung des Effektes der Mutation auf molekularer Ebene angestellt.

Durch Transfektionstudien konnte gezeigt werden, dass der G32E Austausch in der A/B Domäne zu einer reduzierten ligandenabhängigen Aktivität von PPARD als Transkriptionsfaktor führt. Aufgrund der Funktion der A/B Domäne der PPARs
wurden zunächst drei mögliche zugrundeliegende Mechanismen angenommen, die durch die Mutation verändert werden könnten: i) subzelluläre Lokalisation, ii) Ligandenaffinität und/oder iii) die Ubiquitierung der Domäne. Der Einfluss der Mutation auf die subzelluläre Lokalisation wurde mittels Immunfluoreszenzfärbungen von PK-15 und primären Ohrmuschelchondrozyten, die entweder das mutierte oder das Wildtypprotein exprimierten, untersucht. Es konnte gezeigt werden, dass das Wildtypprotein vollständig im Nukleus lokalisiert ist während das veränderte Protein teilweise im Zytosol zu finden war. Um den Effekt der Mutation auf die subzelluläre Lokalisation des Proteins aufzuklären, wurden die G32E exprimierenden Zellen sowie Zellen, die PPARD mit vollständig deletierter A/B Domäne exprimieren, mit Leptomycin B inkubiert. Während das Protein mit der deletierten A/B Domäne vollständig im Nukleus lokalisiert war, wurde durch die Zugabe von Leptomycin B, das das CRM1 Exportprotein blockiert, die zytosolische Lokalisation des G32E PPARD unterbunden. Der Austausch G32E führt also zu einer neuen Funktion der A/B Domäne, die zu einem durch das CRM1 vermittelten nukleären Export führt.


Zusammenfassend handelt es sich bei dem untersuchten G32E Austausch um eine Funktionsverlust-Mutation, die die Aktivität des PPARD als Transkriptionsfaktor reduziert indem sie den Export des Proteins aus dem Nukleus fördert. Dies geschieht durch die Veränderung des nicht-proteolytischen Ubiquitinierungssignals. Im

### Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>APOE</td>
<td>Apolipoprotein E</td>
</tr>
<tr>
<td>BCL6</td>
<td>B-cell CLL/lymphoma 6</td>
</tr>
<tr>
<td>BMP5</td>
<td>Bone morphogenetic protein 5</td>
</tr>
<tr>
<td>CRE</td>
<td>cAMP response element</td>
</tr>
<tr>
<td>CREB</td>
<td>cAMP response element-binding protein</td>
</tr>
<tr>
<td>CFA</td>
<td>Canis lupus familiaris chromosomes</td>
</tr>
<tr>
<td>cM</td>
<td>Centimorgan</td>
</tr>
<tr>
<td>CNS</td>
<td>Central nervous system</td>
</tr>
<tr>
<td>CRM1</td>
<td>Chromosome region maintenance protein</td>
</tr>
<tr>
<td>CD4</td>
<td>Cluster of differentiation 4</td>
</tr>
<tr>
<td>CI</td>
<td>Confidence interval</td>
</tr>
<tr>
<td>CYP7B1</td>
<td>Cytochrome P450, family 7, subfamily B, polypeptide 1</td>
</tr>
<tr>
<td>UBC9</td>
<td>Ubiquitin carrier protein 9</td>
</tr>
<tr>
<td>eQTL</td>
<td>Expression quantitative trait lous</td>
</tr>
<tr>
<td>EYA1</td>
<td>Eyes absent homolog 1</td>
</tr>
<tr>
<td>FGF</td>
<td>Fibroblast growth factor</td>
</tr>
<tr>
<td>FGFR1</td>
<td>Fibroblast growth factor receptor 1</td>
</tr>
<tr>
<td>GABPα</td>
<td>GA binding protein transcription factor, alpha</td>
</tr>
<tr>
<td>GWAS</td>
<td>Genome-wide association study</td>
</tr>
<tr>
<td>GSC</td>
<td>Goosecoid homeobox</td>
</tr>
<tr>
<td>GDF11</td>
<td>Growth differentiation factor 11</td>
</tr>
<tr>
<td>HDAC3</td>
<td>Histone deacetylase 3</td>
</tr>
<tr>
<td>HOXA</td>
<td>Homeobox A cluster</td>
</tr>
<tr>
<td>HOXB</td>
<td>Homeobox B cluster</td>
</tr>
<tr>
<td>IBD</td>
<td>Identical-by-descent</td>
</tr>
<tr>
<td>IGF2</td>
<td>Insulin-like growth factor 2</td>
</tr>
<tr>
<td>Acronym</td>
<td>Full Name</td>
</tr>
<tr>
<td>---------</td>
<td>-----------</td>
</tr>
<tr>
<td>IRF6</td>
<td>Interferon regulatory factor 6</td>
</tr>
<tr>
<td>IFNγ</td>
<td>Interferon-gamma</td>
</tr>
<tr>
<td>LCORL</td>
<td>Ligand dependent nuclear receptor corepressor-like</td>
</tr>
<tr>
<td>LFD</td>
<td>Low-fact diet</td>
</tr>
<tr>
<td>MMP9</td>
<td>Matrix metalloproteidase 9</td>
</tr>
<tr>
<td>MDM2</td>
<td>MDM2 oncogene, E3 ubiquitin protein ligase</td>
</tr>
<tr>
<td>MSRB3</td>
<td>Methionine sulfoxide reductase B3</td>
</tr>
<tr>
<td>MAPK</td>
<td>Mitogen-activated protein kinase</td>
</tr>
<tr>
<td>NCOR</td>
<td>Nuclear receptor corepressor</td>
</tr>
<tr>
<td>NR6A1</td>
<td>Nuclear receptor subfamily 6, group A, member 1</td>
</tr>
<tr>
<td>PTHLH</td>
<td>Parathyroid hormone-like hormone</td>
</tr>
<tr>
<td>PIN1</td>
<td>Peptidylprolyl cis/trans isomerase, NIMA-interacting 1</td>
</tr>
<tr>
<td>PPAR</td>
<td>Peroxisome proliferator-activated receptor</td>
</tr>
<tr>
<td>PPARA</td>
<td>Peroxisome proliferator-activated receptor alpha</td>
</tr>
<tr>
<td>PPARD</td>
<td>Peroxisome proliferator-activated receptor delta</td>
</tr>
<tr>
<td>PPARG</td>
<td>Peroxisome proliferator-activated receptor gamma</td>
</tr>
<tr>
<td>PPRE</td>
<td>PPAR response element</td>
</tr>
<tr>
<td>PTEN</td>
<td>Phosphatase and tensin homolog</td>
</tr>
<tr>
<td>PLAG1</td>
<td>Pleiomorphic adenoma gene 1</td>
</tr>
<tr>
<td>PGC2</td>
<td>PPARG coactivator2</td>
</tr>
<tr>
<td>PIAS</td>
<td>Protein inhibitor of activated STAT</td>
</tr>
<tr>
<td>PKCα</td>
<td>Protein kinase C, alpha</td>
</tr>
<tr>
<td>PDK1</td>
<td>Pyruvate dehydrogenase kinase, isozyme 1</td>
</tr>
<tr>
<td>QTG</td>
<td>Quantitative trait gene</td>
</tr>
<tr>
<td>QTL</td>
<td>Quantitative trait locus</td>
</tr>
<tr>
<td>QTN</td>
<td>Quantitative trait nucleotides</td>
</tr>
<tr>
<td>QTT</td>
<td>Quantitative trait transcript</td>
</tr>
<tr>
<td>RBCK1</td>
<td>RanBP-type and C3HC4-type zinc finger containing 1</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td>RGSs</td>
<td>Regulator of G-protein signalings</td>
</tr>
<tr>
<td>RXR</td>
<td>Retinoid X receptor</td>
</tr>
<tr>
<td>SIPL1</td>
<td>SHANK-interacting protein-like 1</td>
</tr>
<tr>
<td>SIAH2</td>
<td>Seven in absentia homolog 2</td>
</tr>
<tr>
<td>SIX5</td>
<td>Sine oculis-related homeobox 5</td>
</tr>
<tr>
<td>SUMO</td>
<td>Small ubiquitin-like modifier</td>
</tr>
<tr>
<td>SOX9</td>
<td>SRY (sex determining region Y)-box 9</td>
</tr>
<tr>
<td>SSC</td>
<td>Sus scrofa chromosomes</td>
</tr>
<tr>
<td>Timp3</td>
<td>TIMP metallopeptidase inhibitor 3</td>
</tr>
<tr>
<td>Sp1</td>
<td>Trans-acting transcription factor 1</td>
</tr>
<tr>
<td>TNFα</td>
<td>Tumor necrosis factor alpha</td>
</tr>
<tr>
<td>RAF1</td>
<td>Murine leukemia viral oncogene homolog 1</td>
</tr>
<tr>
<td>WNT5A</td>
<td>Wingless-related MMTV integration site 5A</td>
</tr>
<tr>
<td>WIF1</td>
<td>WNT inhibitory factor 1</td>
</tr>
</tbody>
</table>
Chapter 1

General introduction
1 External ear: evolution, development, and genetics

1.1 External ear

The ear is an auditory organ that can convert sound waves from the outside world into nerve impulses that are then sent to the brain. The central nervous system in the brain then translates these impulses into a spoken language, or a bird song, and so on. Besides auditory perception, the ear also holds other functions, like heat radiation, sense of balance and external signals in mood.

The ear consists of three sections: the external ear, the middle ear, and the inner ear. The external ear collects sound and directs it down the ear canal; the middle ear converts the sound waves into mechanical motion; and finally, the hair cells in the inner ear transform mechanical vibration into electrical impulses that are transmitted to the brain. The external ear includes an external auditory canal, an eardrum and a visible pinna that resides outside of the head. The eardrum, a membrane that separates the external ear from the middle ear, evolved gradually among reptiles, birds, and mammals, respectively. But the more sophisticated eardrum first appeared in mammals, which helped them to sense high-frequency vibration (reptiles, for example, can still only detect low-frequency sound waves) [1, 2]. Along with this change, the visible pinna developed as well, to capture and funnel these sounds more effectively [3]. To simplify the writing, in the following passage, the term “ear” will substitute for “pinna.”

In order to better adapt to their environment, mammals have evolved an extraordinary variety of ear shapes and sizes. The animal with the largest ears is the African elephant. His ears developed large surface areas due to the extreme heat of his natural environment with scant amounts of shade to protect him. They act as shades, to keep the elephant’s overall body temperature regulated [4]. Similarly, animals located in hot areas around the world have much larger ears than those located in cold regions, such as the arctic fox. Some mammals, like antelopes and foxes, can turn
their ears in any direction, like aerials, for receiving sound without having to turn their heads. Some underwater and burrowing animals have very small ears as well. Large amphibious nutria has very short and rounded ear structures, for example. Marine mammals, like dolphins, whales and manatees, have no ears.

When humans began to domesticate animals, some specific preferences were gradually bred into the species, and they display greater levels of morphological and behavioural diversity. The ear belongs to one type of these traits. There are a variety of different ears in more than 350 dog breeds worldwide. According to ear shape and size, they can be roughly classified into six types, including: prick, button, semi-prick, hanging, rose, and bat. Ears also have an obvious diversity in Chinese indigenous domestic pigs; for example, Erhualian, Hetao Large-ear, Jiaxing Black and Meishan have large and floppy ears; Hang, Jiangquhai, Laiwu and Lantang have medium-sized, floppy ears; Bama Xiang, Diannan Small-ear and Tibetan have small, erect or half-flicked ears. This preference sometimes is directly correlated with the local culture. For example, owners prefer to breed Erhualian pigs with large ears as a cultural significance, because, as the saying goes, “to have huge ears is a symbol of good luck and wealth” [5].

1.2 Growth and development of the external ear

The ear is derived from the first and second pharyngeal cleft that develops into six rounded protuberances around the sixth week of gestation in a human fetus. These hillocks fold and gradually shift upwards and backwards to their final positions in the head [6]. After birth, a healthy human ear continues growing rather linearly throughout life. The ear grows rapidly until 8 or 10 years and then declines slowly from this point on. Men have ears with a significantly larger surface area than women. This sexual dimorphism appears to start during postnatal growth [7, 8]. A similar growth pattern is also observed in mice, indicating that ear size increases rapidly only throughout the first four weeks of post-birth and then soon slows down significantly.
1.3 Genetic dissection of the external ear

1.3.1 Molecular genetics for the external ear in humans

Most research on the human external ear always pays close attention to those with anomalies. According to Hunter’s standard, these anomalies are classified into three categories: (1) Grade I dysplasia: macrotia, absence or cleft in helix, lobular anomalies, and tragal anomalies. These patients have discernible ear structures and usually do not require the addition of skin or cartilage when surgical ear reconstruction is requested; (2) Grade II dysplasia: cup ear and mini ear. The external ear lacks one or more parts of the normal ear and reconstruction requires additional substitute skin and cartilage; (3) Grade III dysplasia: unilateral or bilateral microtia and anotia. In these cases, whole auricular parts of the ear are not observable, and total reconstruction is required [10].

Regarding the degree of surgical difficulty, a majority of research focused on external ear anomalies; e.g., microtia and anotia. Microtia prevalence rates range from 0.83 to 17.4 per 10,000 births in Italy, France, Sweden, Finland and United States [11]. Such birth defects have a genetic or environmental predisposition. Some medication exposure in the gestational stages can cause microtia, such as retinoic acid, thalidomide and mycophenolate mofetil. Some genetic variants, like chromosomal abnormalities and missense/frameshift mutations have also been evidenced in some microtia-associated syndromes, as summarized in Table 1.
<table>
<thead>
<tr>
<th>Mutation</th>
<th>Phenotype</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Autosomal trisomies</strong></td>
<td></td>
</tr>
<tr>
<td>Chromosome 13 [12]</td>
<td>mental retardation, developmental delay, cleft palate, and ear anomalies</td>
</tr>
<tr>
<td>Chromosome 18 [13]</td>
<td>ear deformity, neonatal hypotony, high palate, microcephaly, brachydactyly, ventricular septal defect, and 5th finger clinodactyly</td>
</tr>
<tr>
<td>Chromosome 22 [14]</td>
<td>microtia, craniofacial dysmorphism, and polygyria</td>
</tr>
<tr>
<td><strong>Autosomal translocation</strong></td>
<td></td>
</tr>
<tr>
<td>6p24 [15]</td>
<td>low set ears, developmental delay, frontal bossing, hypertelorism, flat, broad nasal bridge, and microretrognathia</td>
</tr>
<tr>
<td>2q31.1 [16]</td>
<td>malformed ears, malar and mandibular hypoplasia, lower lid coloboma, cleft palate, and conductive hearing loss</td>
</tr>
<tr>
<td><strong>Copy Number Variation</strong></td>
<td></td>
</tr>
<tr>
<td>4p16 [17]</td>
<td>microtia, eye coloboma, and imperforation of nasolacrimal duct</td>
</tr>
<tr>
<td><strong>Single gene</strong></td>
<td></td>
</tr>
<tr>
<td><em>EYA1</em> [18, 19]</td>
<td>branchio-oto-renal syndrome (microtia, branchial arch defects, hearing loss, and renal anomalies)</td>
</tr>
<tr>
<td><em>FGF3</em> [20-22]</td>
<td>hearing impairment and microtia</td>
</tr>
<tr>
<td><em>GSC</em> [23]</td>
<td>bilateral congenital concha-type microtia</td>
</tr>
<tr>
<td><em>HOXA1</em> [24]</td>
<td>horizontal gaze abnormalities, cardiac autism, profound sensorineural hearing impairment and external ear defects</td>
</tr>
<tr>
<td><em>HOXA2</em> [25]</td>
<td>bilateral microtia</td>
</tr>
<tr>
<td><em>SIX5</em> [26]</td>
<td>microtia, branchial arch defects, hearing loss, and renal anomalies</td>
</tr>
</tbody>
</table>
1.3.2 Animal models for the external ear

Mice are excellent models for dissecting the genetic basis of human ear growth and development. Twenty-seven genes were associated with external ear malformations by analyzing spontaneous mutant and conventional gene-targeting animals by Luquetti et al. [11]. Multi-functional growth factors, like Bmps and Fgfs, play a pivotal role in neuron and cartilage development and postnatal bone formation [27, 28]. Bmp5 is related to external ear growth and spontaneous mutant mice have short ears [29]. Fgfr1 null embryos have marked reduction in the posterior axis and much smaller ears by regulating Hox gene activity [30]. The external ear of Fgf8 knockout newborns is abnormal or absent in line with having much smaller or fused first and second pharyngeal arches during embryonal development [31]. Wnt proteins are a group of secreted lipid-modified signaling proteins, which are involved in neural tube patterning. Wnt5a is expressed in mesenchymal cells of the ear and a loss-of-function mutation leads to outgrowth defects of the developing ears in mice [32].

Besides growth factors, some key transcription factors are required at many stages of ear development and growth, including: cranial neural crest cells proliferation, differentiation, migration into the first and second branchial arches, and chondrocytic progenitor condensation. The HOX family is a critical regulator during the development of the second branchial arch, which contributes to the formation of the external and middle ear. Hoxa1 null mice show absent or hypoplastic external ears and abnormalities of the middle and inner ear, whereas Hoxa1/Hoxb1 double mutants lack complete external ears [33]. Consistent with this, inactivating Hoxa2, Hoxa6, and Hoxb7 in mice also results in absent or hypoplastic ears [34]. Six1, Six4, and Six5 have a remarkably overlapping expression in branchial arches; they are important transcription factors in normal neural crest development and skull formation [35]. Six1/Six4 double mutants develop microtia, yet single gene null mice have a normal phenotype. The Six-interacting protein Eya1 is also involved in ear formation. Eya1 deficient mice show abnormalities in the external ear and branchial arch [36]. In
addition, two novel Eya1 interaction proteins, Sipl1 and Rbck1, enhance the function of Eya protein to act as co-activators for Six [37]. The knockdown of both Sipl1 and Rbck1 results in Eya null mice–like phenotype in zebrafish, characterized by specific abnormalities in branchial arch formation [37]. Tbx1, a T-box transcription factor, directly regulates Fgf8 expression. Tbx1 deficient mice show hypoplastic external ears, a similar phenotype compared to the Fgf8 mutants [38, 39]. Irf6 null embryo also lacks external ears accompanied by abnormal skin, limb, and craniofacial development [40].

In summary, the above-mentioned proteins related to ear growth always take part in the following steps: cranial neural crest patterning, the first and second branchial arches’ outgrowth, chondrocytic progenitor condensation, and cartilage formation. Furthermore, ear growth seems to be involved in several common signaling pathways. For example, 1) Tbx1 → Fgf → Hox pathway: the upstream transcription factor Tbx1 directly regulates growth factor Fgf expression, then Fgf controls the downstream proteins Hox’s activity. 2) Six -Eya -Sip1-Rbck1: Eya -Sip1 -Rbck1 forms a co-activator complex that regulates Six expression.

1.3.3 Molecular genetics of the external ear in domestic animals

In dogs, ear shape is a breed-defining trait. In 2010, Boyko et al. focused on this trait and identified a single region, CFA10: 10-11.6 Mb near MSRB3, which was associated with the ear shape of dogs (floppy, intermediate and prick) using 61,000 SNPs in 915 individuals from 80 domestic breeds, 83 wild canids, and 10 outbred African shelter dogs [41]. Vaysse et al. also found the same ear shape region CFA10: 9.5-11.5 Mb with a peak around 11.3-11.5 Mb, using 170,000 SNPs in 509 dogs from 46 diverse breeds [42]. This region contained 287 SNPs or small indels, which were completely fixed in floppy-ear breeds. Twenty-five SNPs near three candidate genes WIF1, HMGA2, MSRB3 showed sequence conservation and therefore were promising positional candidates.
Although strong artificial selection also impels the rapid phenotypic evolution in pigs, a distinct morphology of ears seems to be influenced by multiple genes. This is in striking contrast to the findings in dogs where a couple of genes explain most of the phenotypic variance. Guo et al. firstly dissected the genetic basis of ear shape in pigs as early as 2004 [43]. They constructed a genetic linkage map using 175 markers and a reference pedigree, including 332 offspring from 19 hybrid Pietrain x (Pietrain x Hampshire) boars and 52 hybrid Leicoma x (Large White x Landrace) sows. Only one QTL was detected on SSC6 at the 1% genome-wide level with a 28.7 cM confidence interval (CI) using the least square regression method.

In 1998, an F2 resource family derived from a cross between two divergent pig breeds—European Large White (small, erect ears) and Chinese Meishan (large, floppy ears) was created at the Roslin Institute (Edinburgh) [44]. Until 2007, Wei et al. divided ear size and erectness into three categories based on video images of 461 F2 individuals recorded at 5–6 months of age [45]; for ear size: small, medium and large; for ear erectness: floppy, medium and prick. A linkage map was constructed with 152 markers including 149 microsatellite markers and three type-I markers on 18 autosomal chromosomes, which was used for the following genome scan. Significant QTLs were detected on SSC1, 5, 7, 9, and 12 for two traits using a web-based software, QTL EXPRESS. For erectness, the loci were located on SSC1, 5, and 7 at the 1% genome-wide level; while for ear size, two 1% genome-wide significant loci were mapped on SSC5 and 7. The QTL on SSC5 has only a 4 cM CI. By comparative mapping, they suggested GDF11 as a positional candidate gene, yet no further study was performed.

From 2000 to 2006, a resource population was created by the interbreed of Chinese Erhualian (large, floppy ears) × American White Duroc (small, fully erect ears) in Nanchang (China). Ma et al. (2009) were the first to consider ear size as a quantitative trait [46]. They dissected both the left and right ears of 1029 F2 offspring at 240 days of age, drew each ear profile on a sulphate paper, scanned and then
calculated using the Leica Qwin area measurement software. Furthermore, more than 1000 F2 pigs were also evaluated for ear erectness at 210 days of age. Twenty-three genome-wide and 12 suggestive QTLs were dissected using a genome-wide scan with 194 microsatellites. The 1% genome-wide QTLs for erectness were detected on SSC1, 2, 5, 7, 12, and 18, of which three QTLs on SSC1, 5, and 7 were consistent with the findings of Wei et al. [45]. Ear size and weight, which are highly correlated, shared common QTLs on SSC1, 4, 5, 7, 8, and 16 notably. The QTLs on SSC5 and SSC7 explained more than 17% and 40% phenotypic variance, with 95% CI spanning around 10 cM and 3 cM, respectively. Both loci on SSC5 and SSC7 seemed to be the most promising targets for further fine-mapping. On this basis, Li et al. refined CI of the QTL on SSC5 to 8.7 cM by genotyping four additional markers in this F2 population [47]. The CI was narrowed down to a small region harbouring three positional candidate genes: *HMGA2*, *SOX5* and *PTHLH*. A *HMGA2* SNP (JF748727: g.2836A>G) showed the strongest association with ear size in the F2 population and a synthetic commercial Sutai pig line; this gene was also a positional candidate for ear shape in dogs [42].

2 PPARD function

PPARD is a member of the PPAR subfamily, which belongs to the steroid/thyroid/retinoid receptor superfamily [48]. The PPAR subfamily includes three isotypes, namely: α, β and γ, which share a similar common structure and mechanism of action as shown in Figure 2. PPAR and RXR form a functional heterodimer, which binds the PPAR response element (PPRE) in the promoter region of target genes [49]. Unliganded receptors remain inactivated; ligand binding changes conformation of the heterodimer, resulting in the release of co-repressors and the recruitment of co-activators necessary for transcription activation. Although the three isotypes share many similar characteristics, they have different tissue distribution,
ligand specificity and function. PPARG regulates fatty acid storage and glucose metabolism, is highly expressed in adipose tissue, and is a target for anti-diabetic drugs [50]. PPARA is a crucial regulator of fatty acid oxidation and is highly expressed in brown adipose tissue and liver [51]. PPARA is activated by adipose-derived fatty acids under nutrient deficient conditions. Activated PPARA up-regulates production of ketone bodies in the mitochondria of liver cells. PPARD is expressed ubiquitously and binds with endogenous ligands, such as fatty acids, triglyceride, prostacyclin, and retinoic acid [52-54]. A recent study showed that ice cream, butter, yogurt, and vegetable cream also increases PPARD activity, but not one common component was found to date [55]. PPARD plays a pivotal role in embryonic development, lipid metabolism, inflammatory response, wound healing, cancer and some basic cell activities, such as proliferation, differentiation, and survival.

Figure 2. Mechanisms of transactivation. The PPAR/RXR heterodimer binds the PPRE in the promoter of target genes through the C domain (DNA-binding domain) of PPAR and RXR. Unliganded receptors remain inactivated; ligand binding by E/F domain (ligand-binding domain) activate the PPAR/RXR heterodimer, resulting in the
recruitment of cofactors containing histone acetyl-transferase activity, modifying nucleosome structure and contacting general transcription factors [56].

2.1 PPARD and metabolism

PPARD knockout mice have a striking reduction in all types of adipose tissue, including gonadal, mesenteric, brown, and subcutaneous stores [57, 58]. However, adipose tissue-specific PPARD deficient mice do not show the similar phenotype as knockout mice [57]. Thus, reduced adiposity of knockout mice might reflect peripheral PPARD functions in systemic lipid metabolism, but not an intrinsic action within adipose tissue. In contrast, another study showed that constitutive activation of PPARD in adipose tissue improved long-chain fatty acid β-oxidation and energy dissipation, resulting in reduced adiposity [59]. Furthermore, transgenic adiposity-specific PPARD overexpressing mice were completely resistant to obesity and sensitive to insulin [59]. Activating PPARD with its ligands can reduce adiposity by controlling the expression of genes involved in fatty acid uptake, β-oxidation and energy uncoupling in adipose tissue of obese animals [60, 61]. Some lipolytic products can activate PPARD in brown adipocytes and elevate the expression of oxidative genes [62].

In addition, PPARD is also crucial in the control of fatty acid oxidation by regulating downstream target genes for fatty acid transport, β-oxidation, and mitochondrial respiration in skeletal muscle [63]. Transgenic mice overexpressing an activated form of PPARD in skeletal muscle produce more type I muscle fibres relative to wild type [64]. A high content of oxidative fibers can help mice to resist high-fat-induced obesity and glucose intolerance. Giordano et al. found that muscle-specific PPARD overexpression also resulted in increased myonuclear density, which was a characteristic response to exercise [65]. More importantly, PPARD has a potential to be a therapeutic target for Duchenne muscular dystrophy by treating its ligand [66].
It is known that adipose tissue and skeletal muscle are important in metabolism. Skeletal muscle consumes glucose and lipid [67], while adipose tissue is not only for energy-storage but also acts as a major endocrine organ. Besides, the central nervous system (CNS) also plays an important role in regulating lipid metabolism. PPARD is expressed in CNS and is specifically enriched in the hypothalamus. Neuron-specific PPARD deficient mice exhibit elevated fat mass, decreased lean mass on a low-fat diet (LFD), leptin resistance, and hypothalamic inflammation [68].

In summary, activation of PPARD can prevent obesity and meliorate insulin resistance in adipose, muscle and CNS tissues. As a result, ligands of PPARD have been promising medicines to treat metabolic syndromes, obesity, and arteriosclerosis.

PPARD is an important determinant of lipid metabolism, as described above. In pigs, over 90 QTLs for body fatness have been mapped on SSC7 containing PPARD according to pigQTLdb (http://www.animalgenome.org/). PPARD is the most promising positional and functional candidate for body fatness QTL on SSC7 by expression analysis in an F2 resource population created by crossing European Large White and Chinese Meishan [69]. An association study also showed that two haplotypes of PPARD were significantly associated with backfat thickness in F2 animals of a Mangalitsa × Pietrain cross and a commercial German Landrace population, respectively [70]. PPARD was also identified as a functional candidate by comparing the expression profile between obese German Landrace and lean Pietrain [71].

### 2.2 PPARD and cell proliferation, differentiation, and apoptosis

PPARD plays a pivotal role in cell proliferation, differentiation and apoptosis, as evidenced in animal models. PPARD null mice show abnormal placentas and smaller fetuses than controls [57, 58]. The main aberration of the placenta in PPARD deficient embryos appears in the giant cell layer; inactivation of PPARD inhibits differentiation of Rcho-1 trophoblast cells [72].
PPARD deficient mice exhibit increased keratinocyte proliferation relative to controls in wound healing [73]. Consistent with this, a hyperproliferative reaction is associated with a higher expression of cell cycle proteins. A two-stage carcinogen bioassay induced a greater tumor size and more tumor multiplicity in PPARD null mice [74]. PPARD attenuated epithelial cell proliferation and tumor formation by upregulating its direct target gene ubiquitin C. The increased number of ubiquitin C enhanced turnover of PKCa, thus preventing further Raf1 and Mek/Erk pathways from increasing cell proliferation [75]. Di-Poi et al. found that PPARD could sense cytokines releasing from injured cells, upregulate its target genes Ilk/Pdk1 and downregulated Pten, and finally suppress keratinocyte apoptosis [76]. In addition, PPARD also ameliorated adhesion/migration of mature keratinocytes by increasing MMP9 levels [76].

PPARD is associated with intestinal tumorigenesis as shown in several studies. Harman et al. found that PPARD could attenuate colon polyp formation in both Min (Multiple Intestinal Neoplasia) mutants and chemically induced mouse models [77]. Consistent with this, Apc\textsuperscript{Min} PPARD null mice are more susceptible to intestinal tumorigenesis than controls [78]. Furthermore, PPARD specific ligand treatment attenuates colon polyp multiplicity of mice [79]. In contrast, Apc\textsuperscript{Min} mice harbour increased number and size of intestinal polyps on PPARD ligand treatment [80]. Unfortunately, no clear gene network has been dissected to date.

Experimental data also supported that PPARD could regulate cell cycle and differentiation in different cell types [81]. PPARD can stimulate adipocyte differentiation and lipid accumulation by upregulating PPAR\textsubscript{G} in adipose tissue [82-84], inhibit the differentiation of paneth cells by Indian hedgehog signal [85] and reduce proliferation of fibroblast by increasing expression of cell cycle inhibitor genes [86].
2.3 PPARD and inflammation

Inflammation is part of a complex biological response of vascular tissues to injury, microbial infections, and altered physiological conditions. PPARD is expressed in several inflammatory cells and closely connected with atherosclerosis [87]. PPARD ligands significantly repress atherosclerotic lesion formation in apoE−/− mice through an HDL-raising effect and anti-inflammatory activity [88]. Activated PPARD serves as a molecular brake in the inflammatory response by increasing the expression of IL1 and Bcl6, reduces chemokine receptor signaling by upregulating RGSs (G protein-coupled signalings), and blocks cell migration and Tnfo shedding through raising Timp3 [88]. Yasunori et al. obtained a similar result, showing that the PPARD ligand increased the expression of Bcl6 and RGSs in an AngII -accelerated atherosclerosis model, characterized by increased vascular inflammatory response by repressing Bcl6 and RGSs [89]. PPARD also protects against dextran sodium sulfate-induced colitis by a ligand-independent mechanism [90].

Furthermore, PPARD represses immune inflammation in the central neural system [91]. In an experimental autoimmune encephalomyelitis model, PPARD deficient mice developed stronger inflammatory response as characterized by higher CD4+, IFNγ, and IL17 in the spinal cord than controls.

Interestingly, macrophage activation also depends on PPARD. Th2 cytokines in adipocytes can activate its expression, which switches macrophage towards activation state M2 to prevent uncontrolled inflammation in white adipose tissue and liver [92]. ArgI, an M2 marker gene, seems to be a target of activated PPARD and PPARG in macrophages [93]. Consistent with these observations, bone marrow-specific PPARD deficient mice show insulin resistance, severe hepatosteatosis and increased adipocyte lipolysis. As a result, PPARD links lipid metabolism and immunity together.

Besides macrophage polarization, PPARD exhibits another anti-inflammation role. The activated macrophages would rapidly engulf apoptotic cells in order to protect the body against autoimmune responses and the release of noxious cellular
contents. The process of uptaking apoptotic cells is delayed in PPARD null mice compared with wild type under ligand treatment [94]. This is partly due to a factor of the classical complement activation pathway, C1q, which is a direct target of PPARD.

3 The function of A/B domain in PPARs

3.1 Structure of PPARD

Nuclear receptors are a group of transcription factors, which regulate the expression of downstream target genes. These factors share similar structures: the N-terminal A/B domain contains a ligand-independent activation function, the C-domain is responsible for DNA binding; the D-domain is a hinge region, and the E/F-domain takes charge of ligand-binding and ligand-independent activation function (Figure 2). Like most nuclear receptors, PPARs totally follow this structure.

Among the four domains, the A/B domain is poorly conserved. Up to now, the crystallization of the A/B domain is impossible, demonstrating that it is intrinsically flexible [95]. Mutation analyses show that the region harbors conserved secondary structures [96]. The helix content in the A/B domain is important to PPARA transactivation as demonstrated by CD spectrum analysis [96]. The C/E domains are highly homologous among the three PPAR subtypes. The C domain contains two zinc finger motifs, while the E domain harbors several α-helices and β-sheets that form a large ligand-binding pocket [97]. The size and amino acid composition of the pocket mainly determines the ligand specificity [98]. Ligand binding alters the conformation of PPARs and switches them towards activation state. The F domain, namely AF-2, is also a part of the ligand binding domain (LBD). The AF-2 domain is an activation surface of PPARs, which contains a conserved helix 12 and additional parts of the LBD [99]. The helix 12 folds up against the core and forms a lid over the ligand-binding pocket. Its conformation is important for the receptor activation.
3.2 Modification of A/B domain regulates transcription activity

PPARs are phosphoproteins whose function can be turned on or off by phosphorylation of serine, threonine or tyrosine residues. The phosphorylation and dephosphorylation of proteins are important molecular switches, which enable cells to give an expeditious response to environmental change. As mentioned above, PPARG is an important regulator of fatty acid storage and glucose metabolism with a high sensitivity to insulin. Insulin is an activator of MAPK, which can phosphorylate the A/B domain of human PPARG (hPPARG) at serine 112 [100]. Phosphorylated PPARG has a reduced transcription activity to downstream target genes that stimulate lipid uptake and adipogenesis. If phosphorylation of hPPARG is artificially prevented by transiting serine into alanine, mice do not show increased adiposity, but are protected from insulin resistance on high-fat diets [101]. Furthermore, insulin treatment also stimulates hPPARA phosphorylation at serine 12 and serine 21 in the A/B domain and increases its transcription activity [102]. Cellular stress in the heart can also activate phosphorylation of hPPARA along with phosphosites by MAPK pathway at the same time [103]. Several additional PPAR kinases are known that directly act on these receptors as summarized by Katherine et al. [104]. PPARs’ phosphorylation in the A/B domain also seems to be a regulator for subcellular localization. PPARG phosphorylation at serine 16 and serine 21 by CK-II can stimulate CRM1/Ran/RanBP3-mediated PPARG export from the nucleus [105].

Ubiquitination is an additional biochemical mechanism that regulates PPAR activity. The process links ubiquitins to some specific lysine residues of the target proteins mediated by a series of enzymes. There are three types of ubiquitination patterns, including polyubiquitination linked via lysine 48 or lysine 63 of ubiquitin, monoubiquitination. The canonical ubiquitin chains, linked by their own lysine 48, are a degradation signal of substrates by 26S proteasome. The other two play wider roles in cellular functions such as translation, cell cycle regulation and protein-protein interactions [106, 107]. An E3 ubiquitin ligase, MDM2, can interact with PPARA
through the A/B domain [108]. MDM2 overexpression increases hPPARA ubiquitination and thereby decreases its protein level. MDM2 can also regulate transcription activity of PPARA. For example, if the ratio of MDM2 to PPARA is < 1, then PPARA activity is increased while a ratio > 1 inhibits its activity in response to the ligand. Regardless of the amount of ligands, the degradation of PPARA usually can be rescued by bound ligands [109]. However, unlike PPARA, hPPARG protein levels are significantly reduced in adipocytes and fibroblasts in the presence of specific ligands [110]. Interestingly, the degradation of PPARG is essential for its transcription activity mediated by ligands in adipose cells. In 2012, a mysterious ubiquitin ligase was identified as SIAH2, regulating ligand-dependent activation by linking polyubiquitin chains with hPPARG [111]. Phosphorylated A/B domain of hPPARG leads to reduced transcription activity, which might be partly due to a decreased degradation by the ubiquitin proteasome system. Fujimoto et al. confirmed the hypothesis that Pin1 inhibits binding of ubiquitin ligases by directly interacting with the phosphorylated A/B domain [112]. To date, ubiquitination might be the only posttranslational modification of PPARD. Like in PPARA, ligands also prevent ubiquitination and subsequent degradation of PPARD in U2OS cells [113]. However, another study showed that protein levels of PPARD in mouse fibroblasts are a determiner to ligand-dependent ubiquitination and degradation pattern [114].

Besides the above modifications, SUMOylation is also a posttranslational regulator of PPAR activity. Like ubiquitin, small ubiquitin-like modifier (SUMO) proteins are also covalently attached to K residues in the consensus site ψKX(D/E) of target proteins by a series of enzymes in cells to modulate their functions. SUMO is involved in several cellular processes, such as protein stability, transcription regulation, and nuclear-cytosolic transport [115]. It has been shown that SUMOylation of PPARA can repress gene expression at cellular and physiological levels. hPPARA can link a SUMO at K185 mediated by E2-conjugating enzyme UBC9 and SUMO E3-ligase PIASy in COS-7 cells and human hepatoma cells,
thereby downregulating its transcription activity through specific recruitment of corepressor NCoR [116]. Posttranslational SUMOylation of PPARA also plays a key role in repressing genes expression involved in hepatic steroid synthesis and complement system of mouse liver in a sex-dependent manner [117]. In females, PPARA with a SUMO at K358 in the E/F domain acts as a transrepressor, interacting with a transcription factor GABPα bound to CYP7B1 gene promoter [117, 118]. This complex triggers recruitment of histone deacetylase and histone methylases, thereby methylating adjacent Sp1 and histones. These events block CYP7B1 expression, which protects females against estrogen-induced intrahepatic cholestasis, a common hepatic disease during pregnancy [117]. Interestingly, the K358-SUMOylated PPARA does not bind with its functional element PPRE, but forms a complex with GABPα [117]. The transcription activity of PPARG can also be repressed by SUMOylation. PPARG1 can be modified by SUMO-1 at K107 in the A/B domain [119]. A site-mutation assay showed that the loss of SUMOylation enhanced target gene expression in HEK293 cells [119]. Consistent results have been obtained with a PPARG1 K107R mutant exhibiting an increased transcription activity of the PPRE promoter region compared with a wild type in NIH3T3 cells using a luciferase reporter assay [119, 120]. Floyd et al. unveiled the potential mechanism on SUMOylation-dependent repression [121]. They found that the covalently bound SUMO peptide affected stability and transcription activity of PPARG1 instead of its nuclear localization [121]. PPARG also harbors another SUMO site K365, which is a determinant to transrepression of inflammatory responses in mouse macrophages [122]. The initial step of the response pathway is SUMOylation at K365 in the E/F domain in the presence of ligands [122]. This modification turns off the expression of inflammatory response genes due to the new conformation preventing the interaction between PPARG and 19S proteasome [122]. In contrast, the activation function of the ligand has been blocked by this SUMO peptide that promotes recruitment of corepressor NCOR-HDAC3 complexes [122].
4 SOX9 function

SOX9 is a transcription factor, which binds to the motif CCTTGAG together with other members of the HMG-box class DNA binding proteins. It plays a crucial role in chondrogenesis, sex determination, cancer and organ maintenance [123-126]. Here I will focus on its function in chondrogenesis.

Chondrogenesis is a multistep process that begins with the condensation of mesenchymal progenitor cells. These cells differentiate into chondrocytes, which proliferate and produce a cartilage extracellular matrix to prefigure the future shape of endochondral bones. At the initial differentiation stage, highly proliferative round chondrocytes are present, which are followed by slowly proliferative flat chondrocytes. Then the chondrocytes gradually exit the cell cycle and differentiate into prehypertrophic and subsequently hypertrophic chondrocytes. Finally, mature hypertrophic chondrocytes go through apoptosis [127].

SOX9 is an evolutionarily conserved transcription regulator in chondrogenesis. Its expression level is high in mesenchymal progenitors and proliferating chondrocytes, and declines abruptly in prehypertrophic chondrocytes. Such expression patterns are strict in every stage of bone development as studied in mice: (i) in mesenchymal progenitors, the miss-expression of Sox9 results in disordered mesenchymal condensation, ectopic chondrogenesis and final limb bud development [128]; (ii) in proliferating chondrocytes, inactivation causes chondrocyte apoptosis and terminal maturation, skipping from the intermediate hypertrophy stage [127]; (iii) in hypertrophic chondrocytes, overexpression retards vascularization, bone marrow formation, and endochondral ossification [129].

Mutations in and around SOX9 also result in skeletal malformations in humans. Several missense and frameshift mutations cause campomelic dysplasia, characterized by the bowing of the femur and tibia [130]. Some regulatory mutations result in the
disruption of normal bone growth and development; i.e., patients with Cooks Syndrome carry microduplications upstream \textit{SOX9} overlapping a critical 1.2 Mb interval [131]. In Pierre Robin sequence, translocation breakpoints flanking \textit{SOX9} have been described [132].

References

early vertebrate development. Endocr Rev 26, 63-77.


92. K. Kang, S. M. Reilly, V. Karabacak, M. R. Gangl, K. Fitzgerald et al., (2008) Adipocyte-derived Th2 cytokines and myeloid PPARdelta regulate...


Chapter 2

A missense mutation in PPARD causes a major QTL effect on ear size in pigs
Chapter 3

The G32E functional variant reduces activity of PPARD by nuclear export and post-translational modification to promote ear size in pigs
Chapter 4

An 18-bp deletion in the 5’-untranslated region of the sex determining region Y-box 9 (SOX9) gene reduces RNA stability and translation efficiency
Chapter 5

General discussion
The pinna refers to the visible section of an animal’s ear, and serves several functions, including: collecting sounds, radiating heat and signaling mood. The structure itself is important in identification as well: for example, the size of the pinna is a significant conformation that distinguishes certain pig breeds. In humans, defects in the pinna are quite common; for instance, microtia, which is characterized by a small and abnormally shaped pinna, has a high prevalence, ranging from 0.83 to 17.4 per 10,000 births internationally. However, only a minority of genetic or environmental causes have been found for this prevalence to date. Long selective breeding has created obvious diversities in pinna size among domesticated pigs, indicating a potential model to explain pinna deformity in humans. However, because of a complex genetic background behind quantitative traits, to definitively identify the causative mutation in pigs has not had much success in the past several decades.

To identify the genes that control pinna size, Ma et al. performed a genome scan in a large White Duroc x Erhualian pig resource family and mapped a 1% genome-wide QTL to a 2 cM confidence interval on chromosome 7 [1]. In this thesis, my all studies focused on dissection of the molecular basis on the QTL. In chapter 2, we further fine-mapped the QTL and finally identified that a missense mutation PPARD G32E could in fact cause such major QTL effects on the pinna size in pigs on a genetic level [2]. This is the first time a study has discovered a novel function of PPARD on controlling pinna size. However, limited knowledge of the PPARD causal mutation indicated that the functional analysis of the mutation was essential in identifying how it controls pinna development. In chapter 3, we investigated the contribution of the G32E substitution to the transcription activity of porcine PPARD and deciphered the underlying molecular mechanism involved. We found that the substitution was a loss-of-function mutation that reduced transcription activity of PPARD by promoting nuclear export and changing non-proteolytic ubiquitination [3]. The overexpression of the G32E mutant reduced SOX9 expression, which is a key regulator for all chondrocyte development. In chapter 4, we conducted a function
analysis for an 18-bp insertion/deletion in the 5'-untranslated region (UTR) of \textit{SOX9} when we identified the indel in three pig breeds, including Laiwu, Bamei and Large White. The 18-bp was indispensable to the translation of \textit{SOX9}; the deletion significantly reduced translation efficiency and caused mRNA destability \cite{4}. However, it was not associated with pinna size in the White Duroc x Erhualian population. These findings established an essential basis for the future research on how PPARD controls pinna development.

1 To dissect the genetic basis of ear size in pigs

1.1 Current research strategy

If we look back on the published data on external ear studies in pigs, it is evident that they all follow a very similar research strategy (see Figure 1): first, a linkage map was constructed using over one hundred microsatellite markers covering the whole genome and pedigree information. Second, all traits of interest were measured in F2 offspring. Finally, some specific loci were identified, combining the following information including a linkage map, microsatellite genotypes and phenotypic values. The CI of QTL typically spans more than 20 cM (approx. \(2.0 \times 10^7\) bp). The subsequent fine-mapping step is necessary to identify the quantitative trait nucleotides (QTN). The mapping strategy has to consider marker density, crossover density, QTL detectability and molecular architecture as described by Georges \cite{5}. Therefore, QTLs explaining a large fraction of phenotypic variation in large populations are very good candidates for fine-mapping. To determine the exact location of a QTL, the first step is to enhance the map resolution by adding more markers in the CI. Based on sequences in the Ensembl project, NCBI and UCSC databases, novel markers can be identified in the specific regions \cite{6}. A high-resolution identical-by-descent (IBD) mapping is followed by linkage and linkage-disequilibrium analysis. The QTL genotype of F1 boars and sows can be deduced by genetic markers and phenotype.
segregation in their F2 offspring using linkage analysis. The shared IBD interval can be determined among F1 animals carrying the known QTL genotype [7]. To obtain a high enough resolution map for identifying a causative gene, denser genetic markers and more chromosomes carrying the relevant alleles are required, as described; for instance, for the identification of *IGF2* [8]. A promising positional candidate gene may be chosen according to accumulating functional research of the genes in humans and mice [9]. Then, causative mutations in a candidate gene might be obtained.

### Figure 1. General strategy for QTL identification in pigs

1. **Step 1 linkage based QTL mapping**
   - Pedigree
   - Genotype
   - Phenotype
   - Linkage map
   - Least squares interval analysis
   - Low-resolution linkage mapping

2. **Step 2 IBD mapping**
   - Add markers in CI
   - Linkage disequilibrium analysis
   - Identification of a minimum shared haplotype containing the QTL

3. **Step 3 positional candidate gene**
   - Physiological function
   - Association analysis in several populations
   - Identification of candidate genes and mutations
   - Functional evaluation of candidate mutations

### 1.2 Progress on genetic research strategy for pigs

The pig genome has been sequenced using a clone-by-clone sequencing strategy by the Wellcome Trust Sanger Institute, which began in 2005 [9]. The increasing number of published sequences makes the development of new markers no longer a tedious
and time-consuming effort. High-density markers in the genome are easily achieved; however, traditional linkage analysis only uses intergenerational recombination events in resource populations. As a result, a large quantity of markers has a limited effect on improving QTL mapping. More precise analysis might be linkage disequilibrium and association analysis, which are based on a phenomenon that LD decays quickly when the marker is far away from the QTL [10]. LD is caused by historical events that chromosomal segments inherited from a common ancestor that became shorter as generation recombination occurred. As a result, the history of domestication and the breed determine its LD pattern and mapping resolution. In European breeds, the extent of LD is large—up to 400 kb, whereas in Chinese breeds LD extends generally less than 10 kb. The LD of European wild boar is intermediate between Chinese and European breeds [11]. Consistent with this, SNP spacing should be about 0.1 cM for European breeds and 0.006 cM for Chinese breeds. Therefore, to develop 30k and 500k SNPs per individual is required in European and Chinese pigs [11].

In January 2009, Illumina launched the PorcineSNP60 BeadChip that included 65,000 evenly spaced probes, which were sufficient for whole-genome association studies, determination of genetic merit, identification of QTL, and comparative genetic studies. The BeadChip presented genetic variation in four main commercial porcine breeds, including Duroc, Landrace, Pietrain and Large White. Since the BeadChip was available, the mapping strategy of quantitative traits changed quickly [12-15].

GWAS is a strategy to detect the statistical association between a trait and a genome-wide panel of markers. Both LD and linkage influence observed associations. So GWAS is effective in identifying loci for complex traits in a large number of individuals in a commercial breed if the existing 60k Chip is used. However, a higher resolution marker panel is required if GWAS is used for the comparative analysis of several breeds due to a small LD between breeds. Furthermore, the statistical formulas have to take population stratification effects and kin structures of the
population into consideration. As a result, a large number of samples might be required to get an accurate QTL. For example, if we identify QTLs for a trait with a heredity capacity of 0.25 and \( r^2 = 0.5 \) using GWAS, over 1,800 animals are necessary to detect the QTL, explaining < 4% genetic variance [10]. It might be better to confirm the results in a large sample of the same breed [10]. These loci can be applied in the future breeding for the same breed. Although the PorcineSNP60 BeadChip has no sufficient markers to detect some Chinese breeds, it is effective to detect a QTL by combining several F2 resource families created by the same founder breeds.

The second and next-generation sequencing methods were introduced in 2007 and 2011, respectively. The new methods provide cheap, efficient, and reliable large-scale DNA sequencing. Groenen et al. found that genes involved in RNA splicing and decay were selected in domestic history using a selective sweep analysis [16]. Rubin et al. reported that three loci associated with \( NR6A1, PLAG1 \) and \( LCORL \) explained most of the phenotypic variation of elongation of the backs and an increased number of vertebrae in domestic European pigs [17]. These pioneering studies indicated that (1) selective sweep analysis was a promising method to identify QTL, (2) whole-genome sequencing resulted in numerous new markers for future higher resolution PorcineSNP BeadChips, and (3) a large number of non-synonymous substitutions in domestic pigs were very useful resources for the future identification of QTN.

The above studies only focused on the association between DNA sequence variants and phenotype; however, they all omitted intermediate steps of genetic information flow. Using next-generation sequencing technology, some studies began to analyze quantitative trait transcripts and used system genetics approaches to identify QTG or QTN Identification [18-20]. System genetics usually include: (1) identification of the eQTL by calculating the association between transcript abundance and DNA variants; (2) identification of the QTT by analyzing the association between transcript abundance and phenotypic values; (3) construction of a
co-expression network of relevant transcripts that are associated with both molecular variant and phenotype; (4) identification of a QTG or QTN by combining the results of eQTL, QTT, and co-expression network [21]. To date, no QTN has been detected using system genetics in domestic animals. Excitingly, two studies of body weight and obesity in mice and humans showed a strong level of system genetics to detect the biological basis of variation for quantitative traits [22, 23].

These emerging methods in identifying the molecular mechanisms of complex traits- including ear traits - in pigs is underway.

2 To dissect biological function of causative mutations

When a potential causative mutation is identified, a functional confirmatory experiment has to follow to reveal its biological function. Furthermore, functional experiments are also effective to exclude those nonfunctional variants in the refined region. Types of genetic variation include SNPs, insertion/deletion, duplication and inversion. Here I give some examples to describe the usual methods to dissect biological functions for different types of variation.

2.1 SNP

There are numerous SNPs in the whole genome. Here we divide SNPs into 2 groups including protein-coding and regulatory mutation. To a mutation in the protein-coding region, the first thing is to prove the mutation is associated with the interested traits. A nonconservative substitution (R200Q) in 5'-AMP-activated protein kinase subunit gamma-3 gene (PRKAG3) caused high glycogen content in skeletal muscle of Hampshire pigs [24]. Milan et al. found that loss-of-function in the homologous gene in yeast resulted in defects in glucose metabolism and less glycogen storage [24]. However, in their original paper, no evidence proved the function of alkaline arginine to neutral glutamine substitution at position 200 (R200Q) in PRKAG3. Then they
made 2 transgenic mice overexpressing wild type \textit{PRKAG3} or R225Q mutant subunit and knockout mice [25]. They found that the mutation was gain-of-function whose transgenic mice showed a marked increase in glycogen content[25]. To SNPs in introns, the most important thing is to find whether the mutation affect gene expression. In 2003, the first QTN was identified as a G3072A substitution in intron 3 of \textit{IGF2} [8]. They found that the mutation located in a functional element by EMSA and could increase \textit{IGF2} expression by a luciferase reporter assay. Using oligonucleotide capture and mass spectrometry, researchers found that the function element could be bound by a repressor ZBED6 [26]. To mutations in UTR, the efficient experiment is to confirm whether and how the mutation affects RNA stability and translation efficiency. A G to A transition in the 3’-UTR of \textit{GDF8} gene creates a target site for \textit{mir1} and \textit{mir206} [27]. By luciferase reporter assay, they found \textit{mir1} and \textit{mir206} inhibited \textit{GDF8} translation and hence caused muscular hypertrophy of Texel sheep [27]. No study describes the causative mutation in the 5’-UTR.

\textbf{2.2 Chromosomal rearrangement}

Another type of genetic variant is commonly chromosomal rearrangement including deletion, duplication, inversion and translocation. Usually, this type of mutation in the protein-coding region does not need to do additional functional experiments. However, if the mutation locates in intron, its function has to be elucidated. The horse graying with age is caused by a 4.6-kb duplication in the sixth intron of \textit{STX17} gene [28]. The duplication increases the expression of two neighboring genes \textit{STX17} and \textit{NR4A3} in melanoma by RT-qPCR. [28]. The two copies are strongly associated with increased expression in neural crest-derived cells, which are precursors of melanocytes in vertebrates. Pea-comb of Chicken is caused by a massive amplification of a duplicated sequence in the first intron of \textit{SOX5} [29]. The expression of \textit{SOX5} was investigated in both Pea-comb and wild-type embryos during development by immunohistochemical staining and situ-hybridization. Pea-comb embryos showed a
strikingly high SOX5 expression in mesenchymal cells located beneath the surface ectoderm where the comb and wattles develop [29]. Both above examples are classical functional elements duplication. Sometimes some simple repeat elements might repress or activate gene expression on more complex conditions. The CGG expansion might be abnormally modified by methylation which lead to chromatin condensation and repress neighboring gene expression [30]. Furthermore, the GAA repeats are prone to form nonusual B-DNA structure and heterochromatin that lead to a direct transcriptional silencing [31]. In our case, the 18-bp deletion locates in the 5’-UTR of porcine SOX9 which reduces translational efficiency due to rapid RNA decay.

3 The pig is a potential biomedical model for ear size

The availability of a whole genome sequence of the pig makes it a promising large animal model for human diseases because the two species share similar size, anatomy, and physiology. The pig has undergone intense selection pressures, causing differentiated subpopulations and phenotypes. The genetic dissection of extreme phenotypes is not so difficult and time-consuming today with the availability of new sequencing technologies and marker genotyping tools in the last two decades. Today, the pig model has been developed for some human diseases, including Alzheimer’s, cancer, and muscular dystrophy. Recently, it has been reported that 112 new gene variants might be models for human diseases by sequencing several strains of wild boars and domestic pigs [32]. Hence, it is reasonable to expect that more new causative genes for ear size will be identified in the future.

4 Future prospect

Although we have already known that PPARD G32E is a loss-of-function mutation, how the variant leads to rapidly increased ear size is still unclear. The studies on
Hoxa2 might be the pioneer to dissect molecular basis of microtia/ anotia genes in mice. As we known, Hoxa2 is a key transcription factor, which regulates the expression of target genes for formation of ear in the second branchial arch. Deficient Hoxa2 causes anotia and abnormality of the middle ear. The well-annotated whole genome sequence and next-generation sequencing technology provide unique opportunity to dissect how this transcription factor controls outer ear development in embryonic stage. Donaldson et al. used chromatin immunoprecipitation and sequencing (ChIP-seq) to identify binding sites of Hoxa2 in whole genome at Embryonic Day 11.5 of mouse [33]. This study confirmed previous findings that Hoxa2 activated the following target genes Robo2, Six2, Runx2, Msx1, Meis1, and Meis2, of which Six2 had been evidenced to involve in outer ear development in mice [34-39]. Furthermore, they first reported that Hoxa2 regulated gene expression within Wnt-β-catenin signaling pathway including a well-known small ear gene Wnt5a [33]. Interestingly, Hoxa2 inactivation causes anotia, while some of its target genes null mice show microtia. This study indicates ChIP-seq is a simple but very effective method to dissect the mechanism of transcription factors.

As expected and excitingly, next-generation sequencing technology and PorcineSNP60 BeadChip had accelerated genetic dissection in pigs. However, most studies on molecular mechanism of growth and development were carried out in mice or simpler animal models. Here we are planning to identify molecular basis behind PPARD G32E, which cause increased ear size. In chapter 3, we have known that the nonconservative G32E substitution reduced its transcriptional activity by nuclear export and posttranslational modification. Mutant has less effective quantity of nucleus PPARD than wild type. It indicates that some different target genes between mutant and wild type can be identified using Chip-Seq. So we are planning to do a Chip-seq of PPARD in ears of Erhualian (mutant) and western commercial pigs (wild type) to identify binding sites of PPARD and detect different sites between both pigs. Furthermore, it also needs to elucidate how transcription activity of target genes is
controlled by PPARD. Some target genes involved in cell growth pathway or ear development pathway might be candidate for the following RT-PCR. Because the activation of PPARD is ligand-dependent [40], it might be effective to detect different expression of candidate genes identified in the ChIP-Seq between with and without ligand treatment in cultured primary ear chondrocyte isolated from Erhualian and western commercial pigs. By this way, we can know which genes in ChIP-Seq can be activated by PPARD. And the significant difference needs to be confirmed again in ear cartilage of wild type and mutant in a large number. PPARD specific ligand GW0742 inhibits the growth of primary ear chondrocytes of western commercial pigs. So we hope to identify a gene to satisfy the following hypothesis: (1) its quantity of binding site in ChIP-Seq is distinct between wild type and mutant; (2) its transcript can be activated by the specific ligand of PPARD in the primary ear chondrocytes of western commercial pigs but not in chondrocyte of Erhualian; (3) expression distinction should be detected in pigs ear cartilage between Erhualian and western commercial pigs; (4) the activity of some cell growth markers controlled by the identified gene would be different in pig ear cartilage between Erhualian and western commercial pigs.

References


metabolism. BMC Genomics 12.


Acknowledgements

I am grateful to all people who helped me when I studied at the Institute of Veterinary Medicine, Georg-August-University of Göttingen, the Key Laboratory for Animal Biotechnology of Jiangxi Province and the Ministry of Agriculture of China in Jiangxi Agricultural University and the Institution of Developmental Biology and Molecular Medicine, Fudan University.

In particular, I wish to express my deepest respect and most sincere gratitude to my supervisor Prof. Dr. Dr. Bertram Brenig, for his continued encouragement, patient guidance and constructive advice at all stages of my work, which help me live through difficult time in the last three years. I am deeply thankful to my co-supervisor Prof. Dr. Christoph Knorr for his help on my lab work. My sincere thanks are due to my co-director Prof. Dr. Lusheng Huang (Jiangxi Agricultural University, China) for his arrangement for my Ph.D. education. I also deliver my appreciation to Prof. Dr. Xiaohui Wu (Fudan University, China) for his valuable guidance, help and encouragement when I studied in Fudan University.

I am thankful to Dr. Julia Beck, Dr. Jan Soller, Dr. Carolin Fromm-Dornieden, Dr. Jeniffer Mayer, Dr. Rangsun Charoensook, Claudia Floren for their assistance for my research, study and life at the Georg-August-University of Göttingen. In addition to the aforementioned, Sabrina Pach has created a good experimental environment and helped with reagent order.

I also do not wish to miss the gratitude to Beibei Ying, Dr. Lin Sun for their encouragement and numerous fruitful discussions, to Dr. Wufan Tao, Dr. Yun Zheng, Dr. Rener Xu, Yan Jin, Rongbo Li, Zengli Guo, Herui Wang, Fubiao Shi, Chi Zhang, Chao Peng, Jian Ye, Xingrong Du and Yan Ye for their help for my research and study at Fudan University.

I am thankful to my good friends and colleagues at Jiangxi Agricultural University in my home country, Prof. Dr. Jun Ren for his valuable advice on my research work and the time he invested in amendment of our manuscripts, Dr. Junwu
Ma, Dr. Jun Gao, Dr. Nengshui Ding, Dr. Yuyun Xing, Dr. Shijun Xiao, Dr. Xueming Yan, Dr. Yuanmei Guo, Dr. Dongren Ren, Dr. Bin Yang, Dr. Zhiyan Zhang, Wanbo Li, Lihua Zhou, Lin Li and other members for their help and encouragement.

Finally, I must express my gratitude to Fei Yuan, my husband, for his permanent love and encouragement. I am continually amazed by his patience to take care of my little brother to complete senior high school education, and by his persistence to take more than 24 hours to visit me by train every month when I studied at Fudan University. I’d like to express my deep thanks to my family for their support to help me face ups and downs of my researches.
Curriculum Vitae

Family Name: Duan
First Name: Yanyu
Sex: Female
Date of birth: Feb. 19, 1981
Place of Birth: Nei Mongol, China
Marital Status: Married
Parents: Feng Duan and Chunping Guo
Nationality: P. R. China

Education

2011.5—
University of Göttingen, Fudan University, Jiangxi Agricultural University
Cooperated postgraduate student,
Major: Molecular genetics in livestock

2010.4—2011.4
Institute of Veterinary Medicine, University of Göttingen,
Postgraduate student,
Major: Molecular genetics in livestock

2007.9—2010.3
Jiangxi Agricultural University,
Lecturer in Animal Biochemistry

2004.9—2007.7
Key Laboratory for Animal Biotechnology of Jiangxi Province and the Ministry of Agriculture of China,
Jiangxi Agricultural University,
Postgraduate student for master degree

1999.9—2003.7
Undergraduate student in Bioengineering,
Inner Mongolia Agricultural University