Using NGS to characterize genetics of meat-type chicken lines

INRA, UMR PEGASE, Rennes, France
Animal genetics department
Challenge in genetics: identify QTL and causative mutation underlying complex traits

Purpose: To offer tools for medicine diagnosis in human genetics or for animal selection

Thousands of QTL identified, whatever the species considered, using linkage analysis

But size of QTL are usually very large and contain hundreds genes: it is still hard to identify causal mutation
Biological context

→ **Challenge in genetics**: identify QTL and causative mutation underlying complex traits

→ **Purpose**: To offer tools for medicine diagnosis in human genetics or for animal selection

→ **Thousands of QTL identified**, whatever the species considered, using linkage analysis

→ But size of QTL are usually very large and contain hundreds genes: it is still **hard to identify causal mutation**

Technologic context

→ More and more studies based on **NGS data**

→ Some questions are still raised: impact of sequencing depth, how to filter those data ...

→ Available tools for mapping, filtering, calling, annotating are numerous

→ Interesting to study in a given study, in a given species, how reliable those data are

→ Use NGS to detect SNP and selective sweep in QTL region to highlight candidate mutation
Strategy

**Experimental Design**

Fat Line × Lean Line → F2
Strategy

EXPERIMENTAL DESIGN

Fat Line × Lean Line

F2

LINKAGE ANALYSIS

QTL mapping study on abdominal fatness (AF) and breast muscle weight (BMW) Lagarrigue et al., 2006
**Strategy**

**DNA SEQUENCING**
**ILLUMINA HISEQ 2000**

- 7 F0 LL, 4 F0 FL and 9 F1 hybrid
- Whole genome (20 X)
- Captured 10 Mb region (100 X)

---

**DATA PROCESSING**

- Alignment on WASHUC2.1: BWA
- Filter on mapping quality: Samtools
- Removal of PCR duplicates: Samtools
- Realignment / Recalibration: GATK
- SNP calling: GATK
- Annotation: VeP

---

**VALIDATION STEP**

- Comparison 20 X vs 100 X: BedTools

---

**EXPERIMENTAL DESIGN**

- Fat Line × Lean Line
- F2

---

**LINKAGE ANALYSIS**

- QTL mapping study on abdominal fatness (AF) and breast muscle weight (BMW) Lagarrigue et al., 2006
**Strategy**

**DNA Sequencing**
**Illumina HiSeq 2000**
- 7 F0 LL, 4 F0 FL and 9 F1 hybrid
- Whole genome (20 X)
- Captured 10 Mb region (100 X)

**Data Processing**
- Alignment on WASHUC2.1: BWA
- Filter on mapping quality: Samtools
- Removal of PCR duplicates: Samtools
- Realignment / Recalibration: GATK
- SNP calling: GATK
- Annotation: VeP

**Validation Step**
- Comparison 20 X vs 100 X: BedTools

**Experimental Design**
- Fat Line \times Lean Line \rightarrow F2

**Linkage Analysis**
- QTL mapping study on abdominal fatness (AF) and breast muscle weight (BMW) Lagarrigue et al., 2006

**Differential Genomic Analysis**
- Using whole genome DNASeq data: HapFLK
- Selective sweep
Using whole genome DNASeq data

Selective sweep

HapFLK

Fat Line Lean Line

F2

LINKAGE ANALYSIS

QTL mapping study on abdominal fatness (AF) and breast muscle weight (BMW) Lagarrigue et al., 2006

INTEGRATION PHASE

→ Reduce QTL size with sweeps
→ Identify potential candidate genes and causal mutations using DNASeq

DATA PROCESSING

→ Alignment on WASHUC2.1 BWA
→ Filter on mapping quality Samtools
→ Removal of PCR duplicates Samtools
→ Realignment / Recalibration GATK
→ SNP calling GATK
→ Annotation VeP

VALIDATION STEP

Comparison 20 X vs 100 X BedTools

DIFFERENTIAL GENOMIC ANALYSIS

Using whole genome DNASeq data

→ Selective sweep HapFLK

EXPERIMENTAL DESIGN

7 F0 LL, 4 F0 FL and 9 F1 hybrid
Whole genome (20 X)
Captured 10 Mb region (100 X)

DNA SEQUENCING
ILLUMINA HiSeq 2000

7 F0 LL, 4 F0 FL and 9 F1 hybrid
Whole genome (20 X)
Captured 10 Mb region (100 X)
Results

**COMPARISON BETWEEN 20 X AND 100 X SNP CALLING RESULTS**
Results

**COMPARISON BETWEEN 20 X AND 100 X SNP CALLING RESULTS**

↑ Context

**Context**

- **SNP 20 X**
  - 95,118 / 10 Mb

- **SNP 100 X**
  - 108,649 / 10 Mb

→ Same individuals sequenced with both 20 X and 100 X on a 10 Mb window
**Results**

**COMPARISON BETWEEN 20 X AND 100 X SNP CALLING RESULTS**

- **Context**
- **Intersection** between the two sets of SNP

**Intersection**

- **SNP 20 X**
  - 95,118 / 10 Mb
- **SNP 100 X**
  - 108,649 / 10 Mb

- Same individuals sequenced with both 20 X and 100 X on a 10 Mb window

Intersection size: 93,330
COMPARISON BETWEEN 20 X AND 100 X SNP CALLING RESULTS

Context
Intersection between the two sets of SNP
Characteristics of 20 X specific SNP

20 X specific SNP

Same individuals sequenced with both 20 X and 100 X on a 10 Mb window

SNP 20 X: 98% in intersection
2% specific

But specific SNP are, in average, of poor quality (average call rate < 90%) and must be considered as false positives

SNP 20 X:
95,118 / 10 Mb

SNP 100 X:
108,649 / 10 Mb

Average call rate (%)

1788
93,330

Specific Intersection
Results

**COMPARISON BETWEEN 20 X AND 100 X SNP CALLING RESULTS**

- **Context**
- **Intersection** between the two sets of SNP
- **Characteristics** of 20 X specific SNP
- **Characteristics** of 100 X specific SNP

**100 X specific SNP**

- **SNP 20 X**: 95,118 / 10 Mb
- **SNP 100 X**: 108,649 / 10 Mb

- **Same individuals sequenced with both 20 X and 100 X on a 10 Mb window**

- **SNP 20 X**: 98 % in intersection
  - 2 % specific

  But specific SNP are, in average, of poor quality (average call rate < 90 %) and must be considered as false positives

- **SNP 100 X**: 86 % in intersection
  - 14 % specific

  Specific SNP have an average call rate of 92 % i.e. among them there are false positives and reliable SNP
**Results**

**COMPARISON BETWEEN 20 X AND 100 X SNP CALLING RESULTS**

- **Context**
- **Intersection** between the two sets of SNP
- **Characteristics** of 20 X specific SNP
- **Characteristics** of 100 X specific SNP

**Conclusion**

- Same individuals sequenced with both 20 X and 100 X on a 10 Mb window

**SNP 20 X**
- 98% in intersection
- 2% specific
- But specific SNP are, in average, of poor quality (average call rate < 90%) and must be considered as false positives

**SNP 100 X**
- 86% in intersection
- 14% specific
- Specific SNP have an average call rate of 92% i.e. among them there are false positives and reliable SNP

20 X re-sequencing is sufficient to have information for almost all SNP
CHARACTERIZATION OF WHOLE GENOME SNP IN OUR LINES
Results

CHARACTERIZATION OF WHOLE GENOME SNP IN OUR LINES

.general characteristics

- **Chicken genome size**: 1.05 Gb
- **SNP per individual**: 2.7 M (± 0.5)
- **SNP density**: 2.6 / kb
- **SNP in the global scheme**: 9.4 M
Results

CHARACTERIZATION OF WHOLE GENOME SNP IN OUR LINES

General characteristics

Chicken genome size : 1.05 Gb
SNP per individual : 2.7 M (± 0.5)
SNP density : 2.6 / kb
SNP in the global scheme : 9.4 M
Results

**Characterization of whole genome SNP in our lines**

- General characteristics
- Genome wide functional consequences: global overview

### Functional consequences

<table>
<thead>
<tr>
<th></th>
<th>SNP</th>
<th>Genes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>All types of SNP</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>%</td>
</tr>
<tr>
<td>Intergenic</td>
<td>4513004</td>
<td>48,68</td>
</tr>
<tr>
<td>Regulatory regions</td>
<td>776080</td>
<td>8,37</td>
</tr>
<tr>
<td>Intronic</td>
<td>3873604</td>
<td>41,78</td>
</tr>
<tr>
<td>Coding - Synonymous</td>
<td>65276</td>
<td>0,70</td>
</tr>
<tr>
<td>Coding - Non Synonymous</td>
<td>43410</td>
<td>0,47</td>
</tr>
</tbody>
</table>

SNPs density in accordance with literature
Results

Characterization of whole genome SNP in our lines
- General characteristics
- Genome wide functional consequences: global overview

Functional consequences

- SNPs density in accordance with literature
- Approximately 50% of SNP are in intergenic regions

<table>
<thead>
<tr>
<th>All types of SNP</th>
<th>SNP</th>
<th>Genes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
</tr>
<tr>
<td>Intergenic</td>
<td>4513004</td>
<td>48.68%</td>
</tr>
<tr>
<td>Regulatory regions</td>
<td>776080</td>
<td>8.37%</td>
</tr>
<tr>
<td>Intronic</td>
<td>3873604</td>
<td>41.78%</td>
</tr>
<tr>
<td>Coding - Synonymous</td>
<td>65276</td>
<td>0.70%</td>
</tr>
<tr>
<td>Coding - Non Synonymous</td>
<td>43410</td>
<td>0.47%</td>
</tr>
</tbody>
</table>
Results

Characterization of whole genome SNP in our lines

- General characteristics
- Genome wide functional consequences: global overview

Functional consequences

- SNPs density in accordance with literature
- Approximately 50% of SNP are in intergenic regions
- Only 1.2% are located on genes

<table>
<thead>
<tr>
<th>All types of SNP</th>
<th>SNP</th>
<th>Genes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
</tr>
<tr>
<td>Intergenic</td>
<td>4513004</td>
<td>48.68</td>
</tr>
<tr>
<td>Regulatory regions</td>
<td>776080</td>
<td>8.37</td>
</tr>
<tr>
<td>Intron</td>
<td>3873604</td>
<td>41.78</td>
</tr>
<tr>
<td>Coding - Synonymous</td>
<td>65276</td>
<td>0.70</td>
</tr>
<tr>
<td>Coding - Non Synonymous</td>
<td>43410</td>
<td>0.47</td>
</tr>
</tbody>
</table>
CHARACTERIZATION OF WHOLE GENOME SNP IN OUR LINES

- General characteristics
- Genome wide functional consequences: global overview

Functional consequences

<table>
<thead>
<tr>
<th>All types of SNP</th>
<th>SNP</th>
<th>Genes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
</tr>
<tr>
<td>Intergenic</td>
<td>4513004</td>
<td>48.68</td>
</tr>
<tr>
<td>Regulatory regions</td>
<td>776080</td>
<td>8.37</td>
</tr>
<tr>
<td>Intrinsic</td>
<td>3873604</td>
<td>41.78</td>
</tr>
<tr>
<td>Coding - Synonymous</td>
<td>65276</td>
<td>0.70</td>
</tr>
<tr>
<td>Coding - Non Synonymous</td>
<td>43410</td>
<td>0.47</td>
</tr>
</tbody>
</table>

- SNPs density in accordance with literature
- Approximately 50% of SNP are in intergenic regions
- Only 1.2% are located on genes
- Among the 17,838 genes listed on the chicken genome, 60% have at least 1 non synonymous SNP
CHARACTERIZATION OF WHOLE GENOME SNP IN OUR LINES

- General characteristics
- Genome wide functional consequences: global overview
- Genome wide functional consequences: focus on coding SNP

**Functional consequences**

<table>
<thead>
<tr>
<th>SNP</th>
<th>SNP</th>
<th>Genes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coding SNP</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Synonymous</td>
<td>65276</td>
<td>59.45%</td>
</tr>
<tr>
<td>Missense</td>
<td>27711</td>
<td>25.24%</td>
</tr>
<tr>
<td>Initator or stop codon</td>
<td>495</td>
<td>0.45%</td>
</tr>
<tr>
<td>Splicing site</td>
<td>16286</td>
<td>14.83%</td>
</tr>
<tr>
<td>Mature mi-RNA</td>
<td>34</td>
<td>0.03%</td>
</tr>
</tbody>
</table>

SNPs density in accordance with literature:
- Approximately 50% of SNP are in intergenic regions
- Only 1.2% are located on genes
- Among the 17,838 genes listed on the chicken genome, 60% have at least 1 non-synonymous SNP
- 60% of the SNP located on genes are synonymous
Characterization of whole genome SNP in our lines

- General characteristics
- Genome wide functional consequences: global overview
- Genome wide functional consequences: focus on coding SNP

### Functional consequences

#### Table: SNP and Genes Distribution

<table>
<thead>
<tr>
<th>Coding SNP</th>
<th>SNP</th>
<th>%</th>
<th>Genes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Synonymous</td>
<td>65276</td>
<td>59.45</td>
<td>11905</td>
</tr>
<tr>
<td>Missense</td>
<td>27711</td>
<td>25.24</td>
<td>8214</td>
</tr>
<tr>
<td>Initiator or stop codon</td>
<td>495</td>
<td>0.45</td>
<td>461</td>
</tr>
<tr>
<td>Splicing site</td>
<td>16286</td>
<td>14.83</td>
<td>7155</td>
</tr>
<tr>
<td>Mature mi-RNA</td>
<td>34</td>
<td>0.03</td>
<td>30</td>
</tr>
</tbody>
</table>

- SNPs density in accordance with literature
- Approximately 50% of SNP are in intergenic regions
- Only 1.2% are located on genes
- Among the 17,838 genes listed on the chicken genome, 60% have at least 1 non-synonymous SNP
- 60% of the SNP located on genes are synonymous
- Missense SNP concern 45% of genes
Results

Genome scan to detect selective sweeps using HapFLK
Results

Genome scan to detect selective sweeps using HapFLK

Selective sweeps

<table>
<thead>
<tr>
<th>Chr</th>
<th>Chr size (Mb)</th>
<th>Selective sweeps (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>200.99</td>
<td>36</td>
</tr>
<tr>
<td>2</td>
<td>154.87</td>
<td>25</td>
</tr>
<tr>
<td>3</td>
<td>113.66</td>
<td>17</td>
</tr>
<tr>
<td>4</td>
<td>94.23</td>
<td>21</td>
</tr>
<tr>
<td>5</td>
<td>62.24</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>37.40</td>
<td>5</td>
</tr>
<tr>
<td>7</td>
<td>38.38</td>
<td>3</td>
</tr>
<tr>
<td>8</td>
<td>30.67</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>22.56</td>
<td>6</td>
</tr>
<tr>
<td>11</td>
<td>21.93</td>
<td>5</td>
</tr>
<tr>
<td>12</td>
<td>20.54</td>
<td>1</td>
</tr>
<tr>
<td>13</td>
<td>18.91</td>
<td>1</td>
</tr>
<tr>
<td>14</td>
<td>15.82</td>
<td>4</td>
</tr>
<tr>
<td>17</td>
<td>11.18</td>
<td>1</td>
</tr>
<tr>
<td>20</td>
<td>13.99</td>
<td>1</td>
</tr>
<tr>
<td>24</td>
<td>6.40</td>
<td>1</td>
</tr>
<tr>
<td>TOTAL</td>
<td>1050.9</td>
<td>129</td>
</tr>
</tbody>
</table>
Results

**Genome scan to detect selective sweeps using HapFLK**

Number of sweeps

### Selective sweeps

<table>
<thead>
<tr>
<th>Chr</th>
<th>Chr size (Mb)</th>
<th>Selective sweeps (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>200.99</td>
<td>36</td>
</tr>
<tr>
<td>2</td>
<td>154.87</td>
<td>25</td>
</tr>
<tr>
<td>3</td>
<td>113.66</td>
<td>17</td>
</tr>
<tr>
<td>4</td>
<td>94.23</td>
<td>21</td>
</tr>
<tr>
<td>5</td>
<td>62.24</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>37.40</td>
<td>5</td>
</tr>
<tr>
<td>7</td>
<td>38.38</td>
<td>3</td>
</tr>
<tr>
<td>8</td>
<td>30.67</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>22.56</td>
<td>6</td>
</tr>
<tr>
<td>11</td>
<td>21.93</td>
<td>5</td>
</tr>
<tr>
<td>12</td>
<td>20.54</td>
<td>1</td>
</tr>
<tr>
<td>13</td>
<td>18.91</td>
<td>1</td>
</tr>
<tr>
<td>14</td>
<td>15.82</td>
<td>4</td>
</tr>
<tr>
<td>17</td>
<td>11.18</td>
<td>1</td>
</tr>
<tr>
<td>20</td>
<td>13.99</td>
<td>1</td>
</tr>
<tr>
<td>24</td>
<td>6.40</td>
<td>1</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td><strong>1050.9</strong></td>
<td><strong>129</strong></td>
</tr>
</tbody>
</table>

**129 sweeps distributed on 16 chromosomes**
Results

**GENOME SCAN TO DETECT SELECTIVE SWEEPS USING HapFLK**

- Number of sweeps
- Size of sweeps

### Selective sweeps

<table>
<thead>
<tr>
<th>Chr</th>
<th>Chr size (Mb)</th>
<th>Selective sweeps (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>200.99</td>
<td>36</td>
</tr>
<tr>
<td>2</td>
<td>154.87</td>
<td>25</td>
</tr>
<tr>
<td>3</td>
<td>113.66</td>
<td>17</td>
</tr>
<tr>
<td>4</td>
<td>94.23</td>
<td>21</td>
</tr>
<tr>
<td>5</td>
<td>62.24</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>37.40</td>
<td>5</td>
</tr>
<tr>
<td>7</td>
<td>38.38</td>
<td>3</td>
</tr>
<tr>
<td>8</td>
<td>30.67</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>22.56</td>
<td>6</td>
</tr>
<tr>
<td>11</td>
<td>21.93</td>
<td>5</td>
</tr>
<tr>
<td>12</td>
<td>20.54</td>
<td>1</td>
</tr>
<tr>
<td>13</td>
<td>18.91</td>
<td>1</td>
</tr>
<tr>
<td>14</td>
<td>15.82</td>
<td>4</td>
</tr>
<tr>
<td>17</td>
<td>11.18</td>
<td>1</td>
</tr>
<tr>
<td>20</td>
<td>13.99</td>
<td>1</td>
</tr>
<tr>
<td>24</td>
<td>6.40</td>
<td>1</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td><strong>1050.9</strong></td>
<td><strong>129</strong></td>
</tr>
</tbody>
</table>

- **129 sweeps** distributed on **16 chromosomes**
- **Average size**: 98 kb (± 90)
Results

Genome scan to detect selective sweeps using HapFLK

- Number of sweeps
- Size of sweeps
- Number of SNP per sweep

Selective sweeps

- **129 sweeps** distributed on **16 chromosomes**

<table>
<thead>
<tr>
<th>Chr</th>
<th>Chr size (Mb)</th>
<th>Selective sweeps (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>200.99</td>
<td>36</td>
</tr>
<tr>
<td>2</td>
<td>154.87</td>
<td>25</td>
</tr>
<tr>
<td>3</td>
<td>113.66</td>
<td>17</td>
</tr>
<tr>
<td>4</td>
<td>94.23</td>
<td>21</td>
</tr>
<tr>
<td>5</td>
<td>62.24</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>37.40</td>
<td>5</td>
</tr>
<tr>
<td>7</td>
<td>38.38</td>
<td>3</td>
</tr>
<tr>
<td>8</td>
<td>30.67</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>22.56</td>
<td>6</td>
</tr>
<tr>
<td>11</td>
<td>21.93</td>
<td>5</td>
</tr>
<tr>
<td>12</td>
<td>20.54</td>
<td>1</td>
</tr>
<tr>
<td>13</td>
<td>18.91</td>
<td>1</td>
</tr>
<tr>
<td>14</td>
<td>15.82</td>
<td>4</td>
</tr>
<tr>
<td>17</td>
<td>11.18</td>
<td>1</td>
</tr>
<tr>
<td>20</td>
<td>13.99</td>
<td>1</td>
</tr>
<tr>
<td>24</td>
<td>6.40</td>
<td>1</td>
</tr>
<tr>
<td>TOTAL</td>
<td>1050.9</td>
<td>129</td>
</tr>
</tbody>
</table>

- **Average size**: 98 kb (± 90)

- **Average number of SNP in sweeps**: 850 (± 700)
Results

**GENOME SCAN TO DETECT SELECTIVE SWEEPS USING HapFLK**

- Number of sweeps
- Size of sweeps
- Number of SNP per sweep
- Number of genes per sweep

**Selective sweeps**

- **129 sweeps** distributed on **16 chromosomes**

- **Average size**: 98 kb (± 90)

- **Average number of SNP in sweeps**: 850 (± 700)

- **Average number of genes in sweeps**: 2.2 (± 1.5)
Results

Overlying QTL Mapping Results with Selective Sweep Analysis Results
Results

OVERLAYING QTL MAPPING RESULTS WITH SELECTIVE SWEEP ANALYSIS RESULTS

QTL mapping results

7 QTL considered in the analysis (6 for AF, 1 for BMW)
Results

Overlaying QTL mapping results with selective sweep analysis results:

- QTL mapping results
- Overlaying with HapFLK results

- 7 QTL considered in the analysis (6 for AF, 1 for BMW)
- All QTL have at least one selective sweep
Results

Overlapping QTL mapping results with selective sweep analysis results

- QTL mapping results
- Overlaying with HapFLK results
- Examples

Examples of co-location between QTL and sweeps

7 QTL considered in the analysis (6 for AF, 1 for BMW)

All QTL have at least one selective sweep

4 QTL have only one selective sweep
Results

Overlaying QTL Mapping Results with Selective Sweep Analysis Results

- QTL mapping results
- Overlaying with HapFLK results
- Examples

7 QTL considered in the analysis (6 for AF, 1 for BMW)

All QTL have at least one selective sweep
**Conclusion**

- **20 X** sequencing depth is sufficient to study **most of SNP** and to give a highly reliable information after filtering steps.

- **HapFLK**, adapted for **analyses on re-sequencing data**, and to take into account genotypes of F1 individuals, allowed us to identify numerous selective sweeps along chicken genome.

- The overlaying step between **QTL** and **selective sweep** results allowed a **reduction of the size of the region to focus on**.

- On these reduced QTL regions we finally identified some interesting **functional candidate genes**.

- NGS data also allowed identification of **candidate mutations**.
Using whole genome DNASeq data

**Selective sweep**

**HapFLK**

**Perspective**

**DNA Sequencing**
**Illumina HiSeq 2000**

7 F0 LL, 4 F0 FL and 9 F1 hybrid
Whole genome (20 X)
10 Mb region (100 X)

**Data processing**
- Alignment on WASHUC2.1: BWA
- Filter on mapping quality: Samtools
- Removal of PCR duplicates: Samtools
- Realignment / Recalibration: GATK
- SNP calling: GATK
- Annotation: VeP

**Validation step**
Comparison 20 X vs 100 X: BedTools

**Experimental design**

Fat Line Lean Line

F2

**Linkage analysis**

QTL mapping study on abdominal fatness (AF) and breast muscle weight (BMW) Lagarrigue et al., 2006

**Integration phase**

- Reduce QTL size with sweeps
- Identify potential candidate genes and causal mutation using DNASeq
- Identify position of causal mutation using AS

**RNA Sequencing**
**Illumina HiSeq 2000**

8 F1 hybrid individuals
6.67 Gb / ind i.e. 33 millions of paired end reads / ind

**Experimental design**

Using F1 hybrid RNASeq data
Identification of parental alleles preferentially expressed for differentially expressed genes

**Differential genomic analysis**
Using whole genome DNASeq data
- Selective sweep: HapFLK

**Allele specific expression analysis**
Using F1 hybrid RNASeq data
Identification of parental alleles preferentially expressed for differentially expressed genes
64th Annual Meeting of the European Federation of Animal Science

INRA, UMR PEGASE, Rennes, France
Genetics & Genomic Team
→ Sandrine Lagarrigue
→ Olivier Demeure
→ Colette Désert
→ Frederic Lecerf

INRA, LGC, Toulouse, France
→ Simon Boitard
→ Bertrand Servin
→ Frédérique Pitel

INRA, SIGENAE, Toulouse, France
→ Anis Djari

INRA, Genotoul Plate-form, Toulouse, France
→ Diane Esquerre

INRA, UMR GABI, CRB GADIE, Jouy-en-Josas, France
→ Sylvain Marthey
→ Marco Moroldo

INRA, UMR GABI, Jouy-en-Josas, France
→ Tatiana Zerjal
→ Jordi Estelle
→ Bertrand Bed’hom
→ Michelle Tixier - Boichard

INRA, UR 38, Avian research, Tours, France
→ Elisabeth Le Bihan - Duval
28th August 2013

Pierre-François ROUX

Many thanks for you attention