The Chicken Genome Project

Leif Andersson

Uppsala University
&
Swedish University of Agricultural Sciences
Current resources for chicken genomics

1. A draft genome sequence (since 2004)

2. A database comprising 2.8 million Single Nucleotide Polymorphisms (SNPs)

3. A consensus linkage map comprising >1000 markers

4. Resources for expression arrays
>5 SNP/kb

~1 SNP/kb
The draft genome sequence of chicken - shortcomings

- It is not a finished sequence

- Still many gaps and errors

- Poor coverage of the Z and W chromosomes

- No assembly for 9 microchromosomes

- Duplicated regions not well annotated

- ~10% of all known cDNAs do not have a match in the assembly!
Improvements of the chicken genome assembly
Dr. Wes Warren, WashU genome sequencing center

1. A new assembly for Z and W based on BAC contigs

2. A new assembly including 454 reads to ~10X coverage

Will this be sufficient to generate a finished genome sequence?
A New SNP-based Linkage Map for the Chicken
Martien Groenen, Leif Andersson, et al.

• ~10,000 SNPs ~ 1 marker/100 kb

• Total length New map = 3280 cM
  Old map = 4200 cM

• Highly significant heterogeneity in recombination rates between populations

• No significant difference in total map length between sexes

• Hot and cold spots for recombination
  Rec rate on average 3.11 cM/Mb
  Range = 0 - ~20 cM/Mbp
A linkage map for chromosome Z
Emerging opportunities for genome resequencing in chicken

1. Roche 454 (long reads)
2. Illumina Solexa (short reads, high throughput)
3. ABI SOLiD (short reads, high throughput)
4. Helicos (not yet ready for whole genome reseq)
5. Pacific Biosciences (future opportunities)
6. ?
Whole genome resequencing in chicken
A pilot project

Red junglefowl

Broiler

White Leghorn Line A

White Leghorn Line B
Whole genome resequencing in chicken

Aims

1. Comprehensive lists of all high frequency polymorphisms that distinguish domestic lines from their wild ancestor

   - Identification of candidate causal mutations

     • Are there any loss-of-function mutations that have been fixed by selection during chicken domestication?
     • Less is more!
Whole genome resequencing in chicken

Aims

2. Detection of selective sweeps
Most recent example *yellow skin*

Whole genome resequencing in chicken

Aims

3. Detection of structural variation: duplications, deletions, inversions

- Number of reads in sliding windows across the genome
- Paired reads that are matched to the genome

~5 kb

Match in assembly
>5 kb = deletion
<5 kb = insertion/duplication
Whole genome resequencing in chicken

Limitations

• Can only align the short reads to the reference sequence

• Insertions/Deletions will occur as gaps in the alignment
Preliminary Results (July 11-08)

• ABI SOLiD runs, 35 bp fragments

• Two lines of White Leghorn (A and B)

• One full slide per breed

• Total cost for experiment <10,000 Euro
WL – Line A

- 2.7 Gbp aligned to the chicken genome
- 75% physical coverage
- 3.6 fold coverage
WL – Line B

- 3.4 Gbp aligned to the chicken genome
- 77.5% physical coverage
- 4.4 fold coverage
Polymorphism detection

<table>
<thead>
<tr>
<th>Population</th>
<th>Total variants</th>
<th>Confirmed SNPs</th>
</tr>
</thead>
<tbody>
<tr>
<td>WL A</td>
<td>4,352,372</td>
<td>1,109,887</td>
</tr>
<tr>
<td>WL B</td>
<td>4,625,772</td>
<td>1,472,277</td>
</tr>
</tbody>
</table>

>1000 potential deletions larger than 10 kb

>500 potential duplications larger than 10 kb
Detection of a duplicated region on chromosome Z
Implications for chicken breeding

1. Genomic selection feasible – will partially replace phenotypic selection

2. Resequencing of major breeding stocks

3. Detection of causal mutations

4. Better understanding of the genetic basis for traits of major importance

5. Genetic engineering
Acknowledgements

Uppsala Genome Sequencing Centre

Max Ingman

Karolinska Institutet, Stockholm

Ellen Sherwood