Recent Advances in Sheep Genomics

Noelle E. Cockett
Utah State University
International Sheep Genome Consortium

- Alan Archibald, Roslin Institute
- Steve Bishop, Roslin Institute
- Noelle Cockett, Utah State University
- Brian Dalrymple, CSIRO
- Clare Gill, Texas A&M University
- Olivier Hanotte, ILRI
- Ben Hayes, Vic DPI
- James Kijas, CSIRO
- Terry Longhurst, Meat and Livestock Australia
- Jill Maddox, University of Melbourne
- John McEwan, AgResearch
- Sean McWilliam, CSIRO
- Frank Nicholas, University of Sydney
- Hutton Oddy, University of New England
- Herman Raadsma, University of Sydney
- Laurant Schibler, INRA
- Ross Tellam, CSIRO
- Chris Warkup, Genesis Faraday Partnership
Growth of the Sheep Linkage Map

Number of markers on map

Year

Sheep Linkage Map Version 5 (SM5)

2,528 markers, ~3,800 cM

- 2,383 on autosomals, 145 on X chromosome
- 1,111 SNPs from 1.5K array, rest mainly microsatellites
- 1,420 unique positions, 2.5 cM average spacing between positions

Autosomal map expanded by 220 cM (6%)

- 168 cM (4.8%) at ends of chromosomes and 54 cM (1.2%) map inflation
- Some previously identified double recombinants disappeared with extra markers

Comparison of Male and Female Chromosome Lengths (SM5)

F/M ratios
Sheep 0.79
Other placental mammals 1.3-1.6
Marsupial 0.54
Ovine Radiation Hybrid Map

- 5,000 rad panel (USUoRH5000)
- Distributed to five international labs
- 3567 loci typed on panel
- High density maps for all but OARY
- 2754 loci on maps
CHORI-243
Ovine BAC Library

- 202,752 clones with 184 Kb average insert
- 376,493 sequences from 193,073 clones
- 258,650,691 bp of sequence (6% of the genome)

Arrange ovine BACs into overlapping contigs using reference genome.
Virtual Sheep Genome

www.livestock.genomics.csiro.au/SheepGenomics/
Capturing the Human Genome Annotation

By creating human to virtual sheep coordinate conversion files, features of the human genome can be displayed on the virtual sheep genome.
VSG1.2 vs Linkage Map (SM5) vs VSG2
Ovine SNP Markers

Sequence from a panel of animals

Align to identify SNPs

Rank SNPs and transfer to chip

Diversity panel:

- Texel
  - AgResearch
- Poll Dorset
  - SheepGenomics
- Lacaune
  - INRA
- Red Masai
  - ILRI
- Awassi
  - Uni. Sydney
- Katahdin
  - USU
- Romney
  - AgResearch
- Merino
  - SheepGenomics
- Gulf Coast Native
  - USU
Illumina Ovine 1536 SNP Array

• Integration of SNPs into linkage map  
  (Jill Maddox, University of Melbourne)

• Integration of SNPs into RH map  
  (Noelle Cockett and Chunhua Wu, Utah State University; Brian Dalrymple, CSIRO)

• Genetic diversity, variation and phylogenetics  
  (James Kijas, CSIRO, Australia)

• SNP panels for parentage assignment  
  (Mike Heaton, USDA/ARS MARC)
Illumina Ovine SNP50 BeadChip

**Source 1:** 6 sheep (Romney, Texel, Merino, Dorset, Rambouillet and Suffolk) using Roche 454 FLX at BCM-HGSC to 0.5X coverage each (3X coverage total)

9.7 Gbp of sequence generated

**Source 2:** Pool of 60 sheep using reduced representation and Solexa GA at Illumina

3 Gbp of sequence generated
A subset of 59,454 SNPs selected from 354,448 total SNPs

BeadChip commercially released January, 2009

Using a test panel of 92 animals from 10 breeds:

- 54,200 loci displayed assay signal
- 49,000 SNP passing QC
- 99% displayed polymorphism
- Average MAF = .25
### Sheep Linkage Map (SM6) Ovine SNP50 Chip

<table>
<thead>
<tr>
<th>Description</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total number of SNPs on chip</td>
<td>59,454</td>
</tr>
<tr>
<td>Number of SNPs reported on</td>
<td>54,241#</td>
</tr>
<tr>
<td>Number of SNPs passing QC</td>
<td>~49,000</td>
</tr>
<tr>
<td>Number of SNPs localized to a chromosome using IMF</td>
<td>44,146</td>
</tr>
<tr>
<td>Number of SNPs localized to a chromosome using Falkiner*</td>
<td>49,220</td>
</tr>
<tr>
<td>Number of ChrUn SNPs localized to a chromosome</td>
<td>312</td>
</tr>
<tr>
<td>Number of chromosome discrepancies between SM6 linkage map and VSG/Btau4</td>
<td>137</td>
</tr>
</tbody>
</table>

# Illumina, GeneSeek reported 54,977 loci but 650 dropped

* 2,325 Falkinder male progeny only, still need to add in 1,733 female progeny
More families wanted for 50K SNP linkage map!

• Half-sibling families with at least 20 offspring per sire
  – Can establish phase by genotyping sire’s parents

• Full-sibling families with both parents and grandparents available

Contact Jill Maddox:
  (jillian.maddox@alumni.unimelb.edu.au)
The Sheep HapMap and Breed Diversity Experiment

Total of 3400 animals from 74 breeds genotyped using the SNP50 BeadChip

Wild sheep (argali, urial, mouflon, bighorn, thinhorn)
Out-group species (barbary, goat, thar, cattle..)
Distribution of Gene Diversity within 74 Breeds

Breed Average for Gene Diversity (He)

Proportion of Breeds (%)
• Asian, African, Middle Eastern, European breeds can be differentiated

• European breeds cluster according to geographical location
Proportion of Total Variance (blue squares)
Cumulative Proportion of Total Variance (red triangles)
Range of LD profiles among 73 breeds - selected sample representing extremes

Average SNP spacing of 50K
Effective population size ~ 50 generations ago

<table>
<thead>
<tr>
<th>Breed</th>
<th>Size</th>
<th>Breed</th>
<th>Size</th>
<th>Breed</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>AfricanDorper</td>
<td>264</td>
<td>Comisana</td>
<td>589</td>
<td>NewZealandRomney</td>
<td>405</td>
</tr>
<tr>
<td>Afshari</td>
<td>465</td>
<td>CyprusFatTail</td>
<td>203</td>
<td>NewZealandTexel</td>
<td>282</td>
</tr>
<tr>
<td>Altamurana</td>
<td>469</td>
<td>Deccani</td>
<td>837</td>
<td>Norduz</td>
<td>285</td>
</tr>
<tr>
<td>AustralianIndustryMerino</td>
<td>853</td>
<td>DorsetHorn</td>
<td>134</td>
<td>Ojala da</td>
<td>588</td>
</tr>
<tr>
<td>AustralianMerino</td>
<td>833</td>
<td>EastFriesianBrown</td>
<td>186</td>
<td>Qezel</td>
<td>1317</td>
</tr>
<tr>
<td>AustralianPollDorset</td>
<td>318</td>
<td>EngadineRedSheep</td>
<td>482</td>
<td>Rambouillet</td>
<td>709</td>
</tr>
<tr>
<td>AustralianPollMerino</td>
<td>918</td>
<td>EthiopianMenz</td>
<td>1260</td>
<td>Rasaaragonesa</td>
<td>780</td>
</tr>
<tr>
<td>AustralianSuffolk</td>
<td>569</td>
<td>Finnsheep</td>
<td>795</td>
<td>RedMaasai</td>
<td>1070</td>
</tr>
<tr>
<td>BangladeshiBGE</td>
<td>582</td>
<td>Galway</td>
<td>322</td>
<td>Sakiz</td>
<td>165</td>
</tr>
<tr>
<td>BangladeshiGarole</td>
<td>437</td>
<td>Garut</td>
<td>337</td>
<td>Santalnes</td>
<td>520</td>
</tr>
<tr>
<td>BarbadosBlackBelly</td>
<td>374</td>
<td>GermanTexel</td>
<td>448</td>
<td>SardinianAncestralBl</td>
<td>306</td>
</tr>
<tr>
<td>BlackHeadedMountain</td>
<td>247</td>
<td>GulfCoastNative</td>
<td>611</td>
<td>ScottishBlackface</td>
<td>528</td>
</tr>
<tr>
<td>BorderLeicester</td>
<td>242</td>
<td>IndianGarole</td>
<td>364</td>
<td>ScottishTexel</td>
<td>305</td>
</tr>
<tr>
<td>BrazilianCreole</td>
<td>371</td>
<td>IrishSuffolk</td>
<td>300</td>
<td>Soay</td>
<td>194</td>
</tr>
<tr>
<td>BundnerOberlanderShe</td>
<td>253</td>
<td>Leccese</td>
<td>556</td>
<td>Spael-white</td>
<td>339</td>
</tr>
<tr>
<td>Castellana</td>
<td>546</td>
<td>MeatLacaune</td>
<td>835</td>
<td>Sumatra</td>
<td>336</td>
</tr>
<tr>
<td>Changthangi</td>
<td>684</td>
<td>Merinolandschaf</td>
<td>407</td>
<td>SwissBlack-BrownMoun</td>
<td>333</td>
</tr>
<tr>
<td>ChineseMerino</td>
<td>397</td>
<td>MilkLacaune</td>
<td>714</td>
<td>SwissMirrorSheep</td>
<td>262</td>
</tr>
<tr>
<td>Chios</td>
<td>334</td>
<td>Moghani</td>
<td>825</td>
<td>SwissWhiteAlpineSheep</td>
<td>374</td>
</tr>
<tr>
<td>Churra</td>
<td>600</td>
<td>MoradaNova</td>
<td>327</td>
<td>Tibetan</td>
<td>903</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ValaisBlacknoseSheep</td>
<td>288</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ValaisRedSheep</td>
<td>173</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Wiltshire</td>
<td>100</td>
</tr>
</tbody>
</table>
Agreement marker positioning LODE vs VSG

Soay

Churro

O. aries genome (Soay)

O. aries genome (Churra)
Yellow fat in Perendales
Single mutation
Partially recessive
30 affected, 30 controls
Gene on OAR21
350 kB region on OAR10

1 gene within the interval
2.2 Mb region on OAR2

15 human Refseqs including *myostatin*
Whole Genome Selection
Australia’s Falkiner Flock

- Largest fine mapping flock in the world
- 15 industry and 5 research sires
- 5,300+ lambs by AI on 4,000+ ewes
- Characterised for 491 traits
- 4,058 progeny genotyped with 50k SNPs
- Funding from MLA and AWI
Ovine Whole Genome Reference Sequence

- Sequence generated from Texel ram used for the CHORI-243 BAC library
  - ~90 Gbp (30X coverage) on Illumina GAII platform (Roslin Institute)
  - ~6 Gbp (2X coverage) on Roche 454 FLX and ABI SOLiD platforms (BCM-HGSC)
  - Gaps will be covered using BAC tiling path on ABI SOLiD platform
- End result will be a whole genome reference sequence equivalent to 7X Sanger sequence and covering 95% of the unique ovine genome
International Sheep Genome Consortium

Funding Sources:
- USDA – ARS (USA)
- USDA – NRICGP (USA)
- USDA – NRSP-8 (USA)
- SheepGenomics (Australia)
- Australian Government ISL
- AgResearch (New Zealand)
- INRA (France)
- 7th Framework Program (EC)