Welcome to the first newsletter from the ISGC. Our aim for these short information sheets is to provide a summary of our activities and to encourage participation from researchers who have a common interest in the genomics of sheep.

Best Regards, James Kijas

ISGC Secretary (James.Kijas@csiro.au)

SNP50 BeadChip Starts to Leave its Mark

Release of the ovine SNP50 BeadChip in January 2009 represented a landmark for the consortium. The chip, developed by researchers from AgResearch, Baylor UCSC, CSIRO, and USDA in partnership with Illumina Inc, contains approximately 50,000 SNPs that are evenly spaced across the sheep genome. Research groups across the community started using the tool as soon as it became available, and the Plant and Animal Genome conference in January was the first time many of the results were presented. High call rates across populations and good quality data was reported by each of the groups. The chip has been used to successfully map a number of single gene traits including chondrodysplasia in Texels, yellow fat in Perendales (picture provided by John McEwan) and microphthalmia in Texels. Other data presented included GWAS for infectious disease (White et al), detailed analysis of genetic diversity and the impact of selection across breeds (Kijas et al) as well as a first look at genome wide levels of linkage disequilibrium (Raadsma, Khatkar, Hayes et al). Copies of the abstracts are available (http://www.intl-pag.org/) and the presentations have been uploaded at (http://www.sheephapmap.org/pag.php). Looking forward, the first half of 2010 is likely to see many more results emerge, particularly from researchers within the ISGC HapMap project which will be the focus of the next newsletter.
Background on the Design of the SNP50 BeadChip

We are regularly asked what strategy was used to design the 50k chip and whether any known functional mutations were included on it. Illumina required that the SNP selection process exclude any DNA variants with known intellectual property claims. As a result the consortium was unable to include any SNP, which at the time, were known to directly underpin disease or production phenotypes. The consortium developed an algorithm which used genomic position, predicted MAF, SNP type and assay suitability to select 49,545 SNP. Some were also chosen based on how well they performed in the pilot 1.5k SNP project (Kijas et al., 2009). In addition a SNP on chromosome Y was included, along with several mitochondrial SNPs and some SNPs representing unpositioned sequences (ChrUn). We also endeavoured to include as many SNPs as possible from the parentage SNP set that is being developed by Mike Heaton (http://cgemm.louisville.edu/USDA/sheep/isgc_snps.html). The algorithm used for SNP selection was developed by Wes Barris and Brian Dalrymple (CSIRO) and further details will be published during 2010.

Ongoing Access of the SNP50 Chip and a New Low Density Array

Interest is growing for the consortium to again coordinate placement of a large order (>10,000 samples) for SNP50 genotyping later in 2010. Feedback gathered at PAG indicates likely demand from Utah State, USDA/ARS, EU 3SR, U Wyoming, AgResearch, U Sydney and Virginia Tech is projected to exceed 8,000 samples and representatives from additional groups (eg sheepCRC and INRA) were not present. The team from Illumina reaffirmed their ongoing willingness to work with the consortium on pricing. The consortium encourages anyone intending to purchase chips to register their interest with John McEwan (john.mcewan@agresearch.co.nz), James Kijas (james.kijas@csiro.au) and Diane Lince (dlince@illumina.com). Looking beyond the existing product, a group led by John McEwan is working on design of a 3 – 5K Golden Gate chip. The opportunity exists to manage the content and associated cost through coordination across the community, contact John for details.

Work Underway on the Reference Genome

The consortium’s major planned activity for 2010 is construction of the sheep reference genome. The work is being coordinated by Alan Archibald and Richard Talbot (Roslin Institute), Donna Muzny and Richard Gibbs (Baylor College of Medicine), John McEwan and Rudi Brauning
(AgResearch), Brian Dalrymple and James Kijas (CSIRO Livestock Industries), Hutton Oddy (University of New England) and Noelle Cockett (Utah State University). The operational approach consists of two stages. In the first phase, a mix of Next Gen sequence will be generated at the Roslin Institute (Illumina GAII) and Baylor College (454) using the single inbred Texel ram used previously for construction of a CHORI BAC library and the virtual sheep genome. The resulting sequence will be assembled and evaluated before a second phase of sequencing is undertaken to fill gaps and improve the interim de novo assembly.

The first phase commenced in December 2009 at the Roslin Institute and preliminary evaluation of the first 40 Gb of sequence returned promising results. The team at the Roslin is working to generate a 30 x genome coverage using 100 bp paired end reads from libraries constructed across a range of insert sizes (220 bp up to 5 kB). To assist in the assembly process, researchers at Baylor College will focus on 454 sequencing of larger insert libraries (8 – 20 kB). All of the data will be deposited in the public domain and the assembly made available for the benefit of the research community.

Additional information on the activities of the ISGC can be found at www.sheephapmap.org

The ISGC represents a community of those interested in sheep genomics. It is free to join and the only requirement is that you have an interest in sheep genomics. The ISGC has regular phone meetings, generally each fortnight, on Tuesdays at 9:00 a.m. Australian Eastern Standard time and minutes of the meetings are placed on the ISGC web site (http://www.sheephapmap.org/).

ISGC funding comes from participants pooling their resources to drive projects that are of interest to the genomics community. Where researchers recognise potential value in leveraging their funding applications with other researchers in the field, they are encouraged to interact within the consortium. Our aim is to achieve maximal benefit for the sheep genomics community from the sometimes limited funding for available ovine genomics.