SMARTER low density genotyping panel

The objective was to develop a low density genotype panel for sheep that would be as informative as possible to a range of different breeds and populations represented in SMARTER. The density of the genotype platforms developed was 384, 1,000, 2,000, 3,000, 6,000, 9,000 and 12,000 single nucleotide polymorphisms (SNPs). To develop this, a summary file of the frequency of each allele per SNP was available for five meat sheep breeds from Ireland (i.e., Belclare, Charollais, Suffolk, Texel and Vendeen), two meat sheep breeds from the UK (i.e., Scottish Blackface and Texel), two meat sheep breeds from France (i.e., Charollais and Vendeen), and five French dairy sheep breeds (i.e., Basco-Béarnaise, Black-faced Manech, Corse, Lacaune, and Red-faced Manech). Allele frequency data were available on 44,040, 577,400 and 48,059 SNPs from the sheep in Ireland, the UK, and France, respectively. A total of 38,883 SNPs were common to all sheep populations. All genotyping had been generated using Illumina platforms and the allele frequencies of all SNPs were subsequently aligned to the Illumina "Allele A/B" format; the frequency of the "A" allele was subsequently calculated in each of the 12 populations. The minimum minor allele frequency across all 12 populations was determined per SNP as was the average of the minor allele frequency per population. Each low density SNP panel was generated separately with the genome firstly being divided into N blocks of equal size where N was the number of SNPs to be selected for the panel being developed. Within each block, SNPs were sorted in descending order of the minimum minor allele frequency of that SNP across all 12 populations and then by the average minor allele frequency across all populations. The first SNP (i.e., the most informative) was selected per block; where no SNP existed within a block then an additional SNP was chosen from an adjacent block. Using the 1,000 SNP panel as an example, all selected SNPs were segregating in all populations with the minimum minor allele frequency across the 12,000 combinations (i.e., 1,000 SNPs by 12 populations) being 0.06; the average minor allele frequency of the selected SNPs was 0.36. For the 12,000 SNP panel a total of 527 were monomorphic in at least 1 breed with an average minor allele frequency across all SNPs and populations being 0.30.

An informative SNP was identified where the frequency of the "A" allele was between 0.2 and ≤0.8 in each of two pair-wise breeds compared. A non-informative SNP was defined as when the two breeds were compared, the frequency of the "A" allele in either breed was either <0.2 or >0.8. On average, 49.33% of all 38,883 common SNPs were informative between each of the 66 pairwise comparisons. The pair of breeds with the maximum number of informative SNPs was, as expected, the Irish Texel and Scottish Texel (59.89% of SNPs were informative) breeds; the minimum number of informative SNPs between the Scottish Texel and Irish Suffolk (40.49% of SNPs). The number of SNPs that were informative in one breed but non-informative in another ranged from 1,081 (informative in the Scottish Texel but not in the Irish Texel breed) to 9,350 (informative in the Corse but not in the Suffolk breed). The present study identified both informative and non-informative SNPs across both meat and dairy sheep. The results from the present study will contribute to the development of future genotyping panels for sheep to maximise the applicability of the genotyping panel across multiple sheep populations.