

## SNP Based Parentage Assignment in Sheep: Application in Australian Flocks



James Kijas <sup>1</sup>, Julius van der Werf <sup>2,3</sup>, Mohammad Ferdosi <sup>2,3</sup>, Amy Bell <sup>1,3</sup>, Sam Gill <sup>3,4</sup>, Klint Gore <sup>3,5</sup>, Felice Driver <sup>4</sup>, Jill Maddox <sup>6</sup> and John Henshall <sup>1,3</sup>

<sup>1</sup> CSIRO Livestock Industries, Australia. <sup>2</sup> University of New England, Australia. <sup>3</sup> Sheep Cooperative Research Center, Australia. <sup>4</sup> Meat and Livestock Australia. <sup>5</sup> Animal Genetics and Breeding Unit, University of New England . <sup>6</sup> Port Melbourne, Australia.

# The use of accurate pedigrees is important for livestock production systems and research projects. We present the development and attributes of a SNP panel for the assignment of parentage in sheep.



The genomic position and average minor allele frequency (MAF) of each SNP is given in Table 1. SNP were selected using population allele frequencies obtained from over 70 breeds sampled from 5 continents. The panel is biased towards high MAF markers (Figure 2) to ensure it will be of use across a wide range of breeds.

#### Figure 2: MAF



MAF was calculated by genotyping 2384 sheep from 74 breeds using the *ovine* SNP50 BeadChip [2]. MAF distributions are shown for all 49034 SNP on the BeadChip (red bars) and 89 SNP in the parentage panel (green bars, Table 1).

The core panel of 89 markers was supplemented with additional SNP for testing within Australian flocks that included markers for Horn Poll and disease genes. A total of 6 Sequenom (SQ) multiplexs (383 SNP) were developed and tested for their genotyping accuracy and performance. This sought to identify the minimum number of SNP panels required for achieve high rates of assignment.

#### **Key Points**

- \* We have identified a technically robust set of SNP suitable for parentage analysis in a wide variety of sheep.
- **\*** Disclosure of the SNP and their attributes is intended to promote uptake by commercial partners.
- \* The core panel is being promoted to the International Society of Animal Genetics (ISAG) as the standard for testing in sheep.

\_

- **\*** Using industry Merinos, we find application of 3 SQ plexes (190 SNP) returned high rates of assignment.
- \* Wide applicability of the SNP panels opens the way for international level trace-back and product of origin testing.

[1] Kijas et al. (2009) PLoS ONE 4:e4668	
[2] Kijas et al. (2012) PLoS Biology 10:e1001258	<ul><li>[4] http://www.livestockgenomics.csiro.au/sheep/oar2.0.php</li></ul>
[3] Crawford et al. (1995) Genetics 140:703-724.	[5] Marshall et al. (1998) Molecular Ecology 7: 639-655.

Acknowledgments. Aspects of the work were performed within the International Sheep Genomics Consortium (ISGC). We thank Marc Murphy (Karbullah Merino Stud, Queensland, Australia) for allowing access to his animals for blood collection and DNA genotyping.

#### Table 1: SNP ID and Genomic Location for Core SNP

ore SNP	Chr	Mb	Allele	MAF	Core SNP	Chr	Mb	Allele	MAF	Core SNP	Chr	Mb	Allele	MAF
		Pos					Pos	-				Pos	-	
00290101_408.1	1	7.8	A	0.337	DU194639_560.1	6	56.7	G	0.442	DU426312_454.1	15	44.4	G	0.375
0U518561_359.1	1	14.2	G	0.381	CZ925803_293.1	6	100.8	A	0.443	DU301502_402.1	15	73.7	G	0.441
U351298_316.1	1	69.6	Α	0.445	DU337465_337.1	6	106.0	Α	0.338	DU241306_191.1	15	78.6	G	0.279
U232924_365.1	1	95.8	G	0.250	CL635944_160.1	6	115.0	A	0.490	DU324670_456.1	17	10.2	A	0.400
0U271929_382.1	1	97.5	Α	0.483	DU467751_524.1	7	10.6	Α	0.429	DU206327_107.1	17	14.4	Α	0.499
U502334_443.1	2	19.1	A	0.437	DU499587_509.1	7	74.0	A	0.325	DU378819_632.1	17	22.3	Α	0.475
U469454_586.1	2	26.2	G	0.394	CZ920950_468.1	7	74.8	Α	0.456	DU511222_139.1	17	27.4	Α	0.351
U425907_184.1	2	50.1	G	0.358	DU530067_219.1	7	100.0	G	0.327	DU300156_445.1	17	38.0	G	0.456
U501115_497.1	2	62.8	A	0.239	DU213735_493.1	8	6.6	А	0.437	DU463532_137.1	17	56.0	Α	0.443
0U492516_411.1	2	63.4	т	0.478	DU411204_551.1	8	13.8	А	0.361	DU492379_209.1	18	3.9	Α	0.385
U470875_383.1	2	91.5	G	0.357	DU189970_325.1	9	86.6	С	0.374	DU488903_267.1	18	21.4	G	0.334
50506CS_*1	2	100.9	G	0.345	DU471913_499.1	9	91.1	G	0.490	DU325612_517.1	18	25.4	Α	0.433
U191879_495.1	2	157.6	A	0.335	DU364754_308.1	9	93.9	А	0.397	DU440765_491.1	18	60.5	Α	0.474
U480434_533.1	2	192.2	А	0.480	DU372582_268.1	9	94.4	G	0.247	DU345394_399.1	18	61.1	А	0.450
U260201_585.1	2	226.7	А	0.422	DU468275_284.1	10	33.1	А	0.352	DU264531_279.1	19	0.6	А	0.388
U503161_123.1	2	237.2	А	0.352	DU310747_445.1	10	38.2	G	0.470	DU258053_237.1	19	57.1	Α	0.400
U425259_620.1	3	21.4	A	0.461	DU269694_582.1	11	1.9	А	0.473	DU411432_523.1	19	57.2	С	0.406
U231007_156.1	3	59.0	G	0.463	DU433863_261.1	11	15.5	А	0.419	DU183112_480.1	20	31.1	А	0.453
U225323_218.1	3	91.0	А	0.467	DU417675_79.1	11	19.6	А	0.344	DU442373_141.1	20	48.4	Α	0.342
U260081_579.1	3	108.8	А	0.383	DU508448_227.1	11	25.3	А	0.485	DU380983_440.1	21	28.3	G	0.451
U394537_289.1	3	181.6	G	0.371	DU326572_241.1	11	59.5	А	0.446	DU383863_376.1	21	38.2	G	0.443
L635241_413.1	3	181.9	А	0.455	DU314655_578.1	12	26.7	А	0.365	DU196132_525.1	21	42.7	G	0.388
U408817_431.1	3	205.0	А	0.343	DU310703_497.1	12	75.3	А	0.492	DU413316_575.1	22	13.1	Α	0.419
U202116_405.1	4	58.2	А	0.444	DU275428_276.1	13	10.9	А	0.460	DU302760_528.1	23	11.6	G	0.494
U460511_423.1	4	61.1	G	0.443	DU435573_466.1	13	30.1	А	0.449	DU313102_671.1	23	17.3	G	0.484
U305004_417.1	4	70.1	A	0.270	DU411403_398.1	13	41.3	G	0.427	CZ920359_258.1	24	3.2	G	0.382
U369175_467.1	4	73.0	G	0.375	DU462008_263.1	14	44.6	А	0.330	DU455254_479.1	25	0.1	G	0.453
0U446213_412.1	5	12.5	А	0.394	DU223894_556.1	14	57.5	G	0.449	DU512685_259.1	25	1.2	G	0.495
U444709_372.1	5	56.0	А	0.489	DU381045_479.1	14	60.7	A	0.403	oY1	Y	0.0	G	0.320
U453259_440.1	5	64.8	G	0.346	DU464373_638.1	15	2.3	А	0.467					
_					-									

Table 1. The genomic location of each SNP (Chr / Mb Pos) is taken from the genome assembly version OAR2.0 available at [4]. SNP identifiers can be used to obtain additional information about each SNP [4]. The minor allele is given along with its frequency (MAF) in 2384 animals [2].

A flock of 290 Merinos was collected and genotyped using all 6 SQ plexes. High levels of relatedness were expected within both the sires and dams. Figure 3 shows promising results using a set of 190 SNP (3 SQ plexes).

### **Figure 3: Performance Within an Australian Flock**



To set thresholds for assignment, simulation was performed using the observed genotypes of parents in each of the real populations. 1000 lambs were simulated, each consistent with the genotypes of a random sire and dam. LOD was calculated [5] as the ratio of likelihood correct/incorrect. The LOD of the most likely sire was termed mLOD1, and the LOD of the next most likely sire mLOD2. mLOD1 was plotted against Δ1 (the difference between mLOD1 and mLOD2) and coloured blue if it was correct or red if incorrect. The equivalent plot for mLOD2 is shown in light blue. Results and thresholds are shown using either 127 (A) or 190 (B) SNP. Application of the thresholds within a population of Australian Merinos revealed the majority of progeny sire assignments exceed the threshold (C).



