



Royal Veterinary College  
University of London

Do selective sweeps in sheep breeds indicate the genomic sites of breed characteristics?

Geoff Pollott

# Background

- Work on wool shedding indicated a dominant gene (Pollott, 2011. JAS, 89: 2316; confirmed by Matika *et al.*, 2012).
- Analysed sheep Hapmap data to try to locate it.
- Noticed long runs of homozygosity (ROH) in some breeds.
- Runs of homozygosity are an indication of the lack of variation in neighbouring DNA.
- A selective sweep is the reduction or elimination of variation in neighbouring DNA as the result of recent and strong positive selection.
- This work looks at within-breed ROH in the Sheep Hapmap 50K SNP dataset and their link with breed characteristics.

# What does a long ROH indicate?

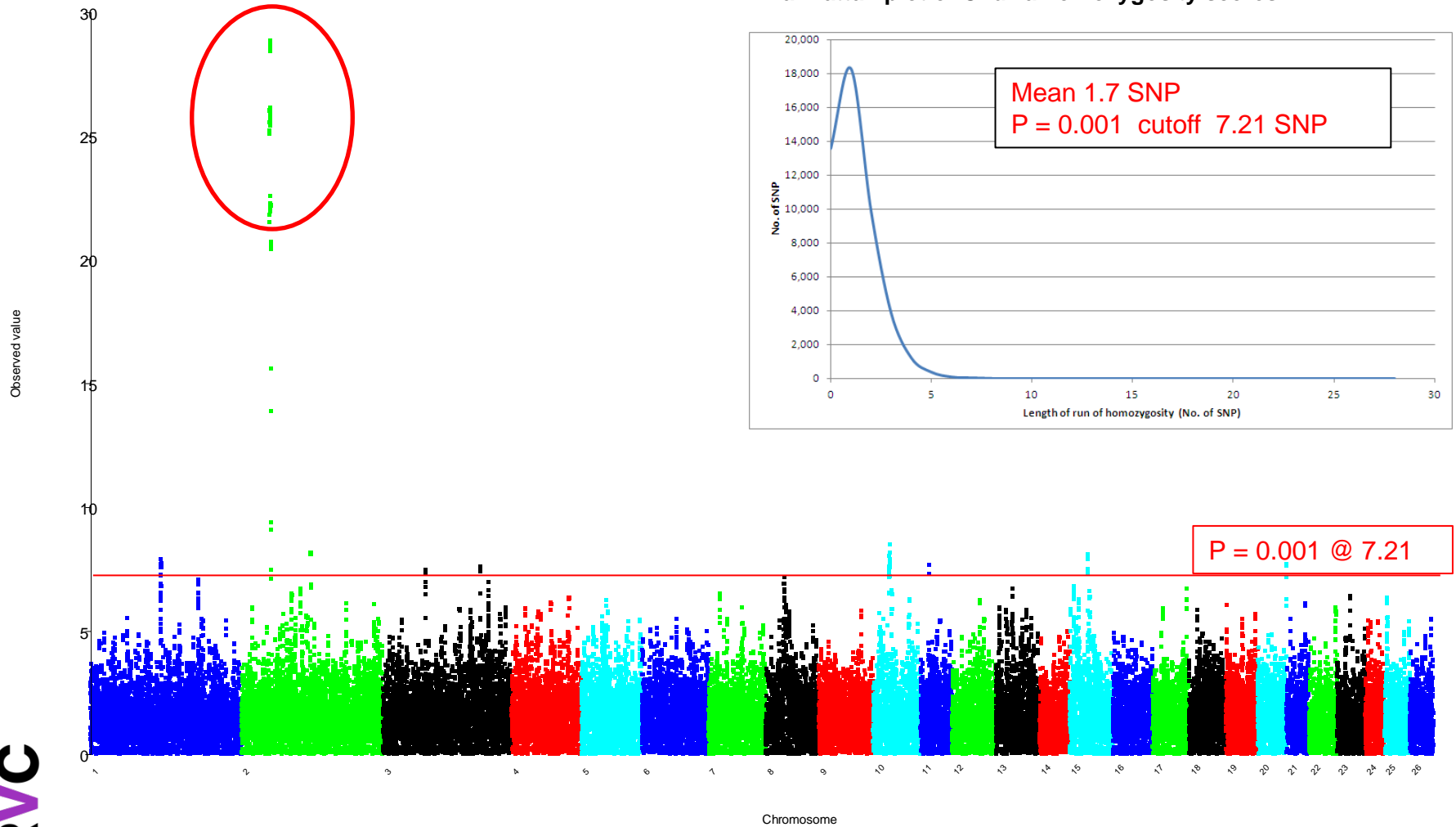
- Possible monomorphic SNP in a breed
- Recessive character of importance
  - commonly used to find recent harmful recessives.
- Inbreeding
- Commonly inherited stretches of DNA

# Calculating ROH from Sheep Hapmap data

For each breed dataset - 24 or more animals:

- Find majority homozygote at each SNP.
- Calculate length of ROH for the major homozygotes within each animal.
- Calculate mean ROH length at each SNP.
- Calculate Poisson distribution of mean ROH values for the breed.
- Find cut-off point for  $P < 0.001$  in terms of mean ROH.
- Identify ROH above this length.
- Look for genes nearby.

# Churra – an example (N = 120)



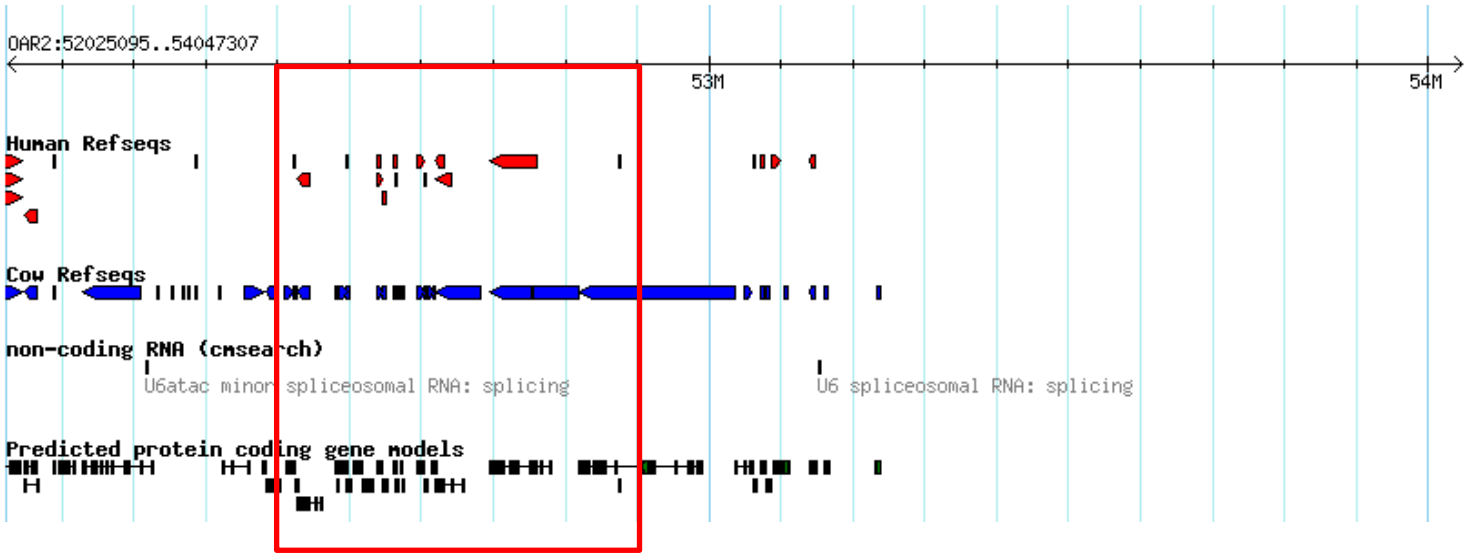
# Churra OAR2 data

chr	snp	bp	mean	sd	cv	min	max	G11	G12	G22
2	OAR2_55617004.1	55,617,004	21.85	17.48	0.80	0	47	0	1	119
2	s60207.1	55,642,214	21.53	17.86	0.83	0	47	5	32	83
2	s28881.1	55,704,709	25.21	16.15	0.64	0	47	0	11	109
2	OAR2_55752685.1	55,752,685	25.11	16.31	0.65	0	47	2	21	97
2	s31618.1	55,777,451	25.57	15.72	0.61	0	47	1	15	104
2	OAR2_55818455.1	55,818,455	26.07	15.40	0.59	0	47	1	11	108
2	OAR2_55853730.1	55,853,730	26.03	15.45	0.59	0	47	1	13	106
2	OAR2_55861669.1	55,861,669	26.03	15.45	0.59	0	47	1	13	106
2	s63030.1	55,922,894	28.46	13.53	0.48	0	47	1	2	117
2	s20468.1	56,248,983	28.77	13.07	0.45	0	47	0	2	118
2	s58048.1	56,288,982	28.85	12.89	0.45	5	47	0	0	120

chr	snp	bp	mean	sd	cv	min	max	G11	G12	G22
2	s20468.1	56,248,983	28.77	13.07	0.45	0	47	0	2	118
2	s58048.1	56,288,982	28.85	12.89	0.45	5	47	0	0	120
2	OAR2_56300406.1	56,300,406	28.61	13.34	0.47	0	47	0	3	117
2	OAR2_56623917.1	56,623,917	25.96	16.84	0.65	0	47	2	20	98
2	s09641.1	56,669,725	26.14	16.55	0.63	1	47	0	0	120
2	OAR2_56707900.1	56,707,900	25.45	17.37	0.68	0	47	2	27	91

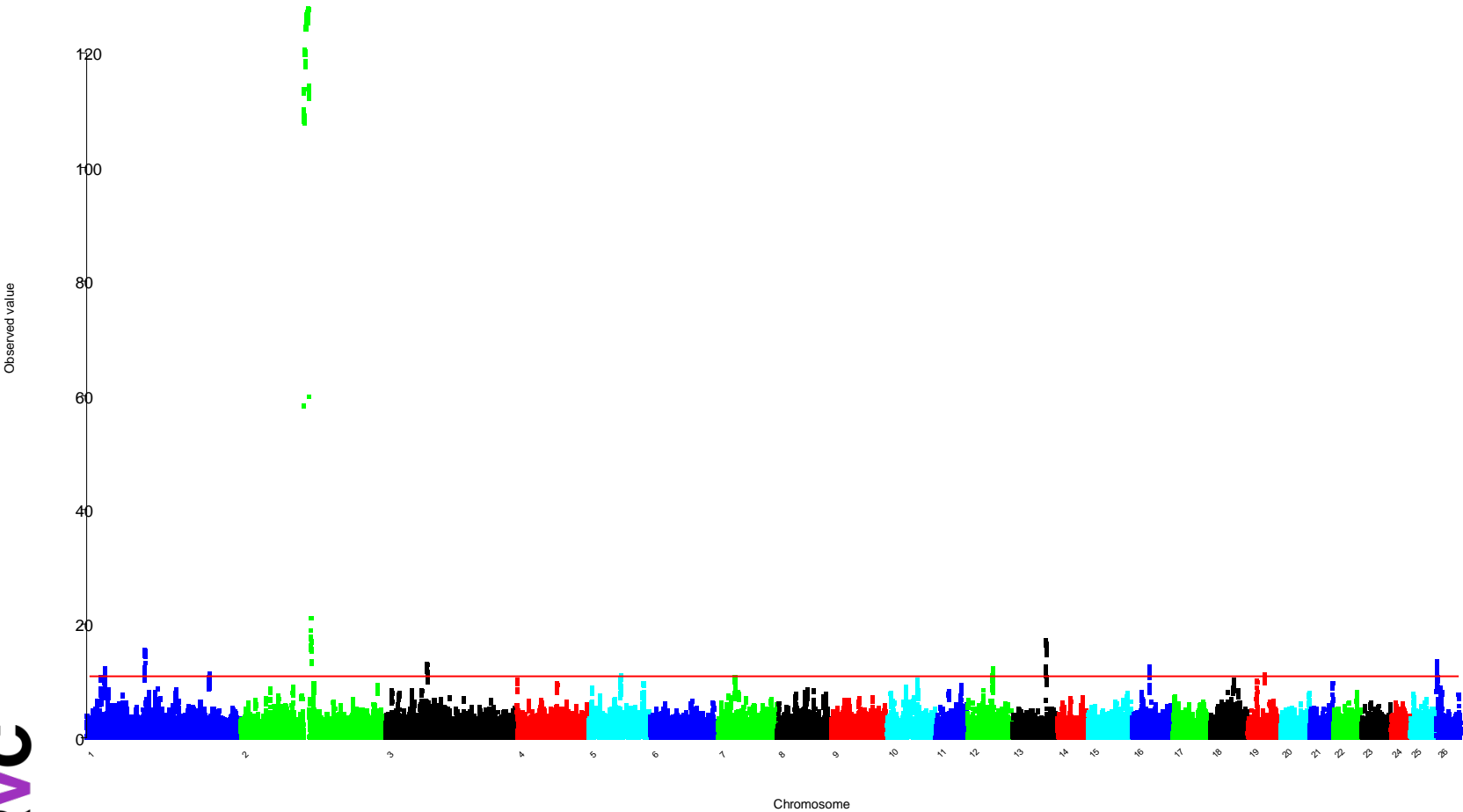
2	OAR2_56869189.1	56,869,189	22.00	19.85	0.90	0	47	3	47	70
2	s23499.1	56,935,114	22.00	19.85	0.90	0	47	3	47	70
2	OAR2_56982765.1	56,982,765	21.98	19.73	0.90	0	47	4	37	79
2	OAR2_57136229.1	57,136,229	22.04	19.66	0.89	0	47	5	32	83
2	OAR2_57167138.1	57,167,138	21.98	19.72	0.90	0	47	2	36	82
2	OAR2_57196512_X.1	57,196,512	21.96	19.76	0.90	0	47	5	42	73
2	s18609.1	57,251,937	22.29	19.40	0.87	0	47	1	21	98
2	s12424.1	57,305,021	22.58	19.16	0.85	0	47	1	27	92
2	OAR2_57526461.1	57,526,461	22.19	18.93	0.85	0	47	1	21	98
2	OAR2_57561444_X.1	57,561,444	22.23	18.66	0.84	0	47	1	11	108
2	OAR2_57596413.1	57,596,413	20.61	19.33	0.94	0	47	3	34	83
2	OAR2_57602414.1	57,602,414	20.73	19.21	0.93	0	47	4	31	85
2	s53985.1	57,679,956	20.44	18.87	0.92	0	47	0	14	106
2	OAR2_57729090.1	57,729,090	15.62	18.74	1.20	0	47	0	36	84
2	s34608.1	57,796,196	13.89	18.85	1.36	0	47	7	38	75
2	OAR2_57832237.1	57,832,237	9.09	16.40	1.80	0	47	12	40	68
2	OAR2_57868140.1	57,868,140	9.41	16.25	1.73	0	47	5	29	86
2	s02336.1	58,046,931	7.48	15.33	2.05	0	47	14	61	45
2	OAR2_58056173_X.1	58,056,173	7.11	15.46	2.18	0	47	30	61	29
2	OAR2_58107393.1	58,107,393	3.71	10.25	2.76	0	47	5	48	67

# Genes in this region



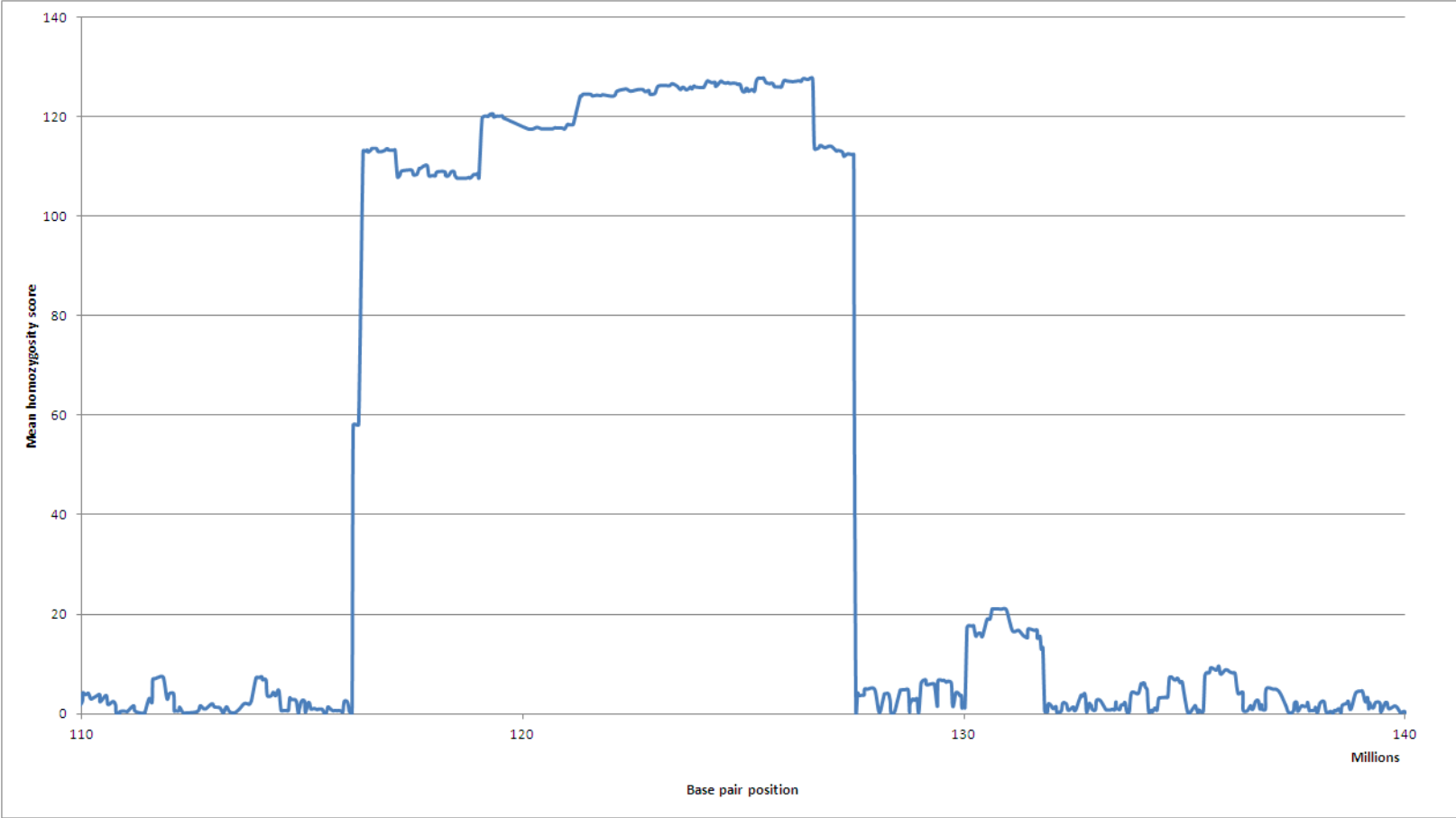
# Large ROH – Scottish Texels (N = 80)

Manhattan plot of Scottish Texel homozygosity scores

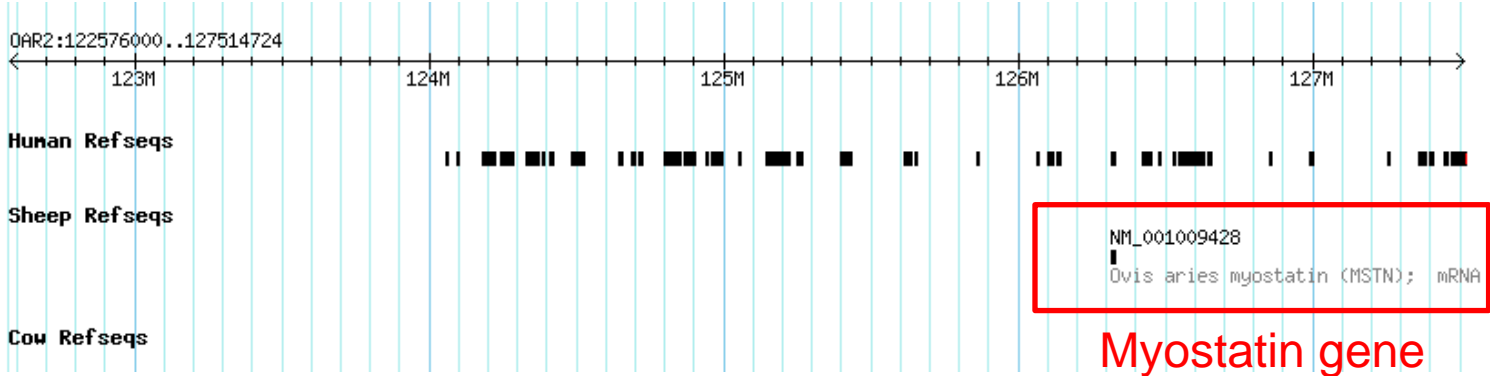




# Detail of OAR2 – Scottish Texels

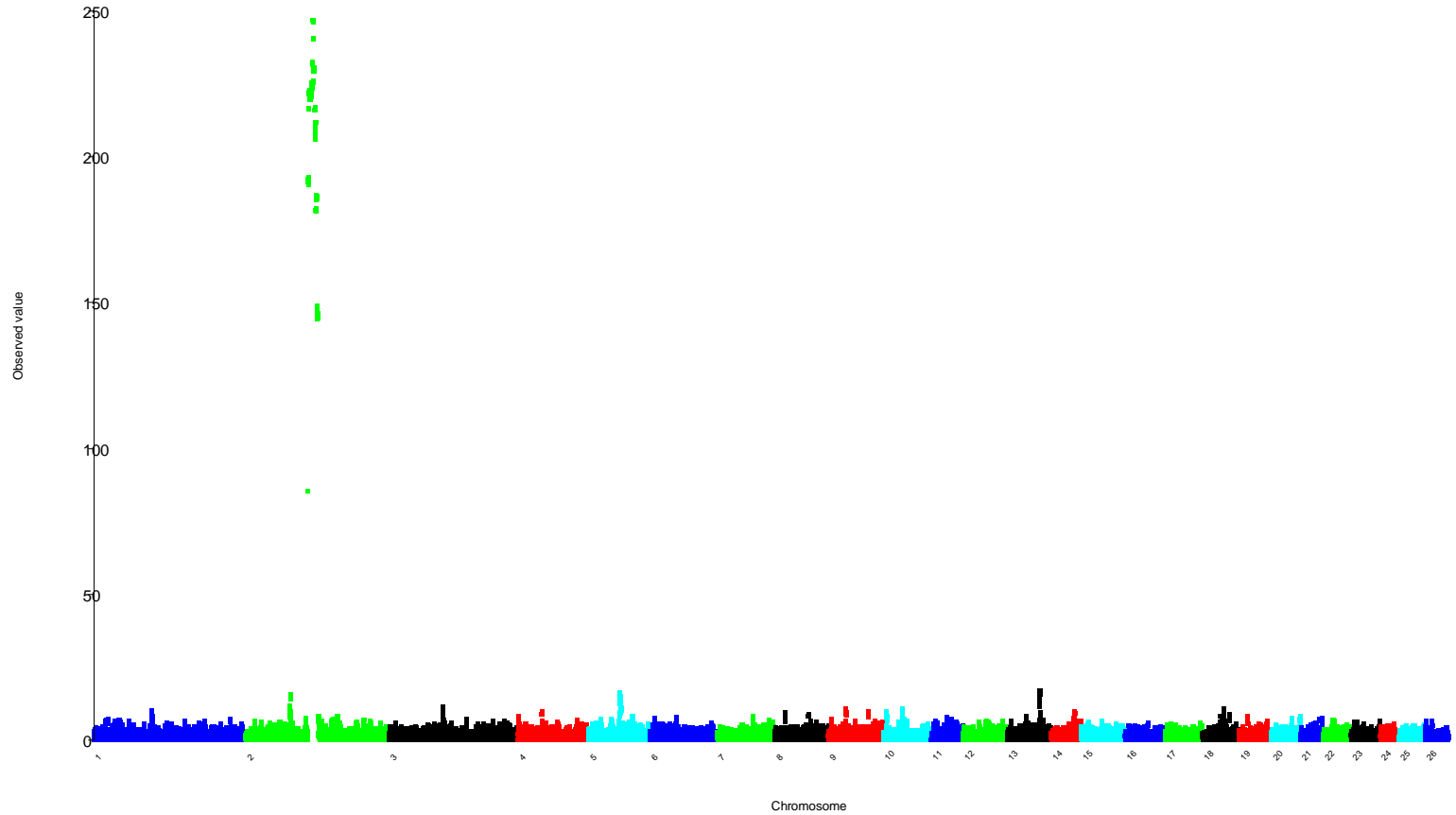


# Genes in Scottish Texel Region

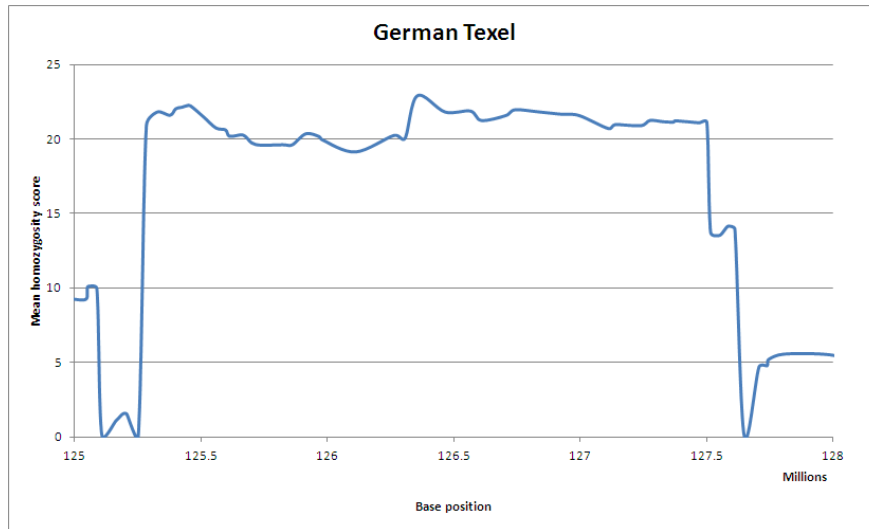


# New Zealand Texels (N = 24)

Manhattan plot of New Zealand Texel homozygosity scores

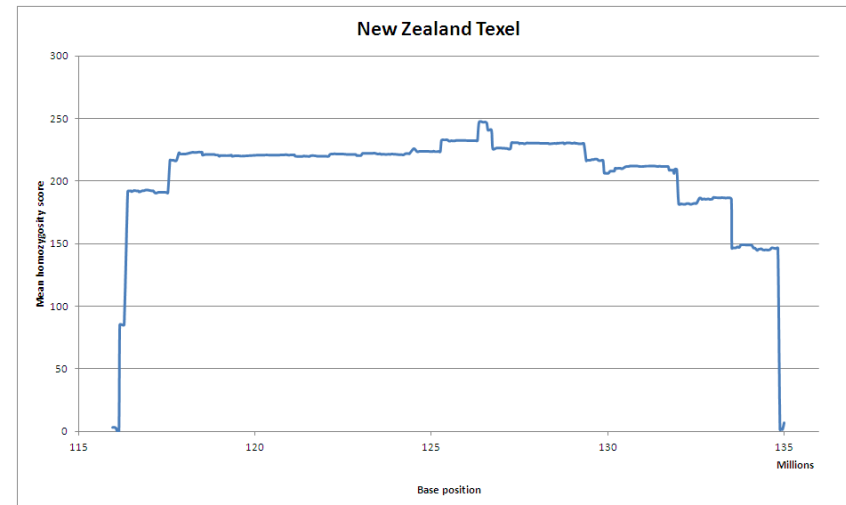


# German and New Zealand Texels



Next SNP upstream from Myostatin gene

- was completely homozygous for the same genotype in all 3 samples
- had the highest homozygosity score in all 3 samples



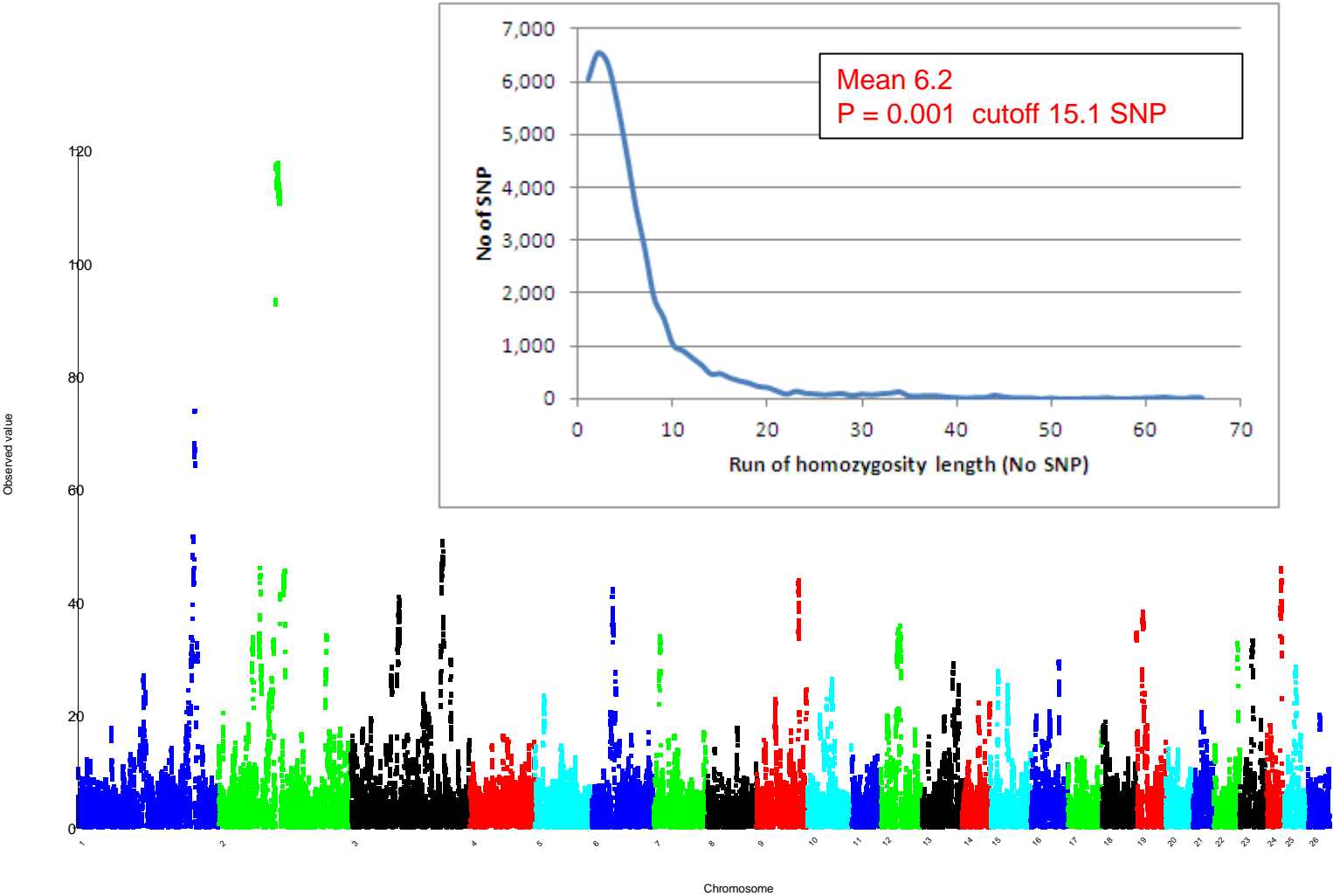
# Two Lacaune types

Breed	Longest mean HZ score (no. SNP)	Chr	Position	Start	End	Length (Mb)
Milk Lacaune	14.2	6	41,583,796	41,003,295	43,064,935	2.06
	10.8	1	137,856,107	137,186,574	138,571,443	1.38
	10	10	29,469,450	29,381,795	29,867,192	0.49
Meat Lacaune	17.6	6	40,780,034	39,081,346	42,094,768	3.01
	9.2	11	18,701,428	18,568,846	19,092,204	0.52
	9.1	10	29,469,450	29,381,795	29,907,137	0.53

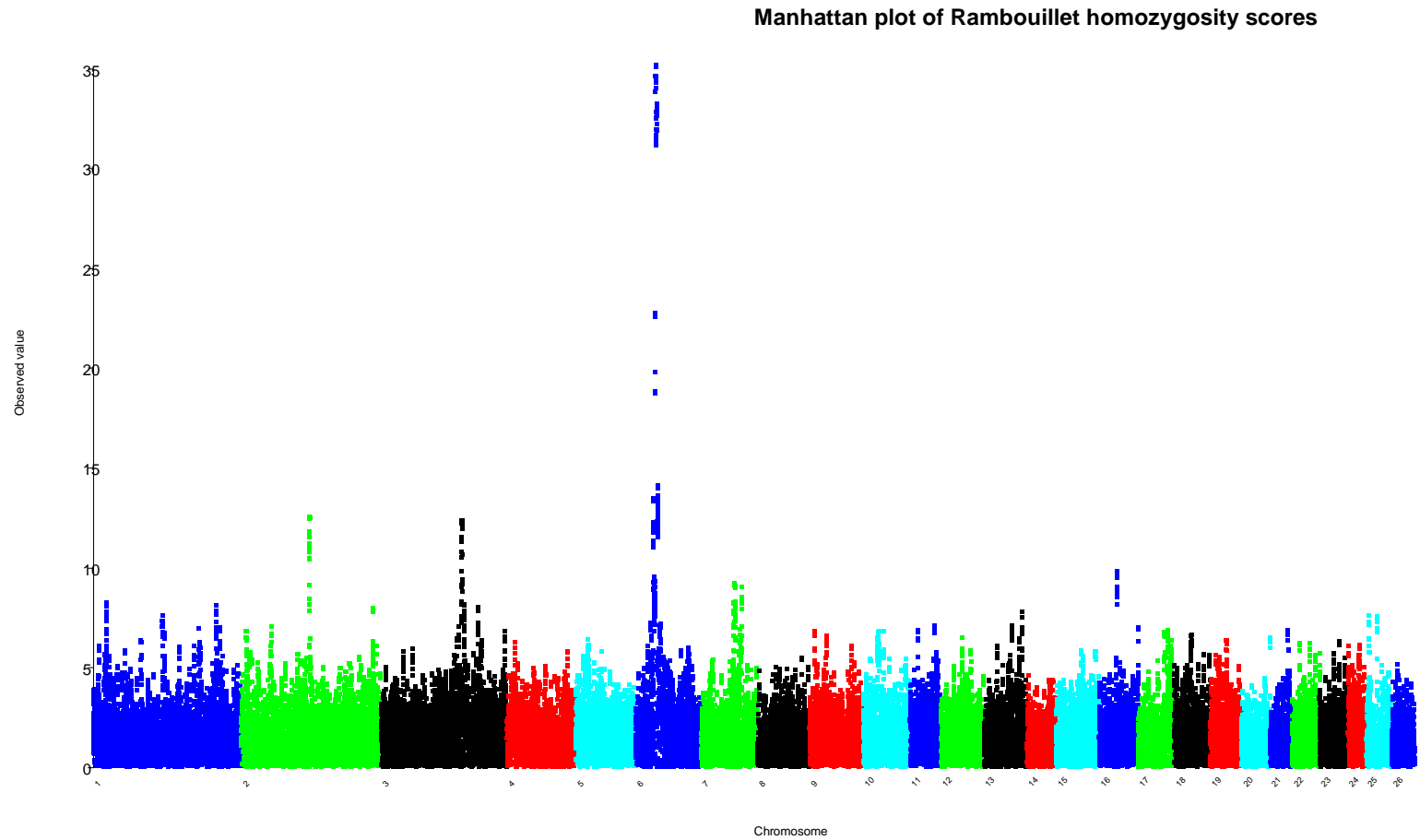
# Overlapping ROH between breeds

Breed	Longest mean HZ score (No. SNP)	Chr	Position	Start	End	Length (Mb)
Scottish Blackface	27.8	2	55,922,894	55,138,384	56,557,812	1.42
Irish Suffolk	11.7	2	56,288,982	55,493,630	56,383,269	0.89
Soay	117.7	2	119,604,666	116,066,653	125,735,242	9.67
Finnsheep	10.8	2	122,611,468	121,800,892	123,533,453	1.73
Australian Suffolk	11.7	2	125,326,201	124,923,303	125,596,219	0.67
Rambouillet	12.6	2	130,761,692	130,195,322	131,507,986	1.31
Finnsheep	11	2	130,953,296	130,195,322	131,656,211	1.46
Meat Lacaune	17.6	6	40,780,034	39,081,346	42,094,768	3.01
Milk Lacaune	14.2	6	41,583,796	41,003,295	43,064,935	2.06
Rambouillet	35.2	6	41,768,532	40,169,159	43,010,674	2.84
Australian Suffolk	10.8	10	29,469,450	29,341,212	29,907,137	0.57
Milk Lacaune	10	10	29,469,450	29,381,795	29,867,192	0.49
Meat Lacaune	9.1	10	29,469,450	29,381,795	29,907,137	0.53
Irish Suffolk	18.3	10	37,886,155	37,775,025	40,399,035	2.62
Finnsheep	16.3	10	38,966,500	38,159,341	40,399,035	2.24
Scottish Blackface	17.3	10	39,605,193	38,772,235	40,475,642	1.70
Australian Suffolk	11.3	11	28,950,668	26,956,590	29,487,277	2.53
Australian Polled Dorset	16.5	11	28,950,668	26,927,248	29,443,457	2.52

# Soay data (N = 110)



# Rambouillet data (N = 102)





# Discussion points

Can ROH be used to investigate 'unknown' regions of breed interest?

Significant cut-off point to use?

Indicators of breed history?

What will WGS tell us?

# Conclusions

- Runs of homozygosity can be linked to breed characteristics
- Analysis of ROH simple but informative
- Different way of characterising breeds
- Indicator of key breed characteristics at the molecular genetic level

# Acknowledgements



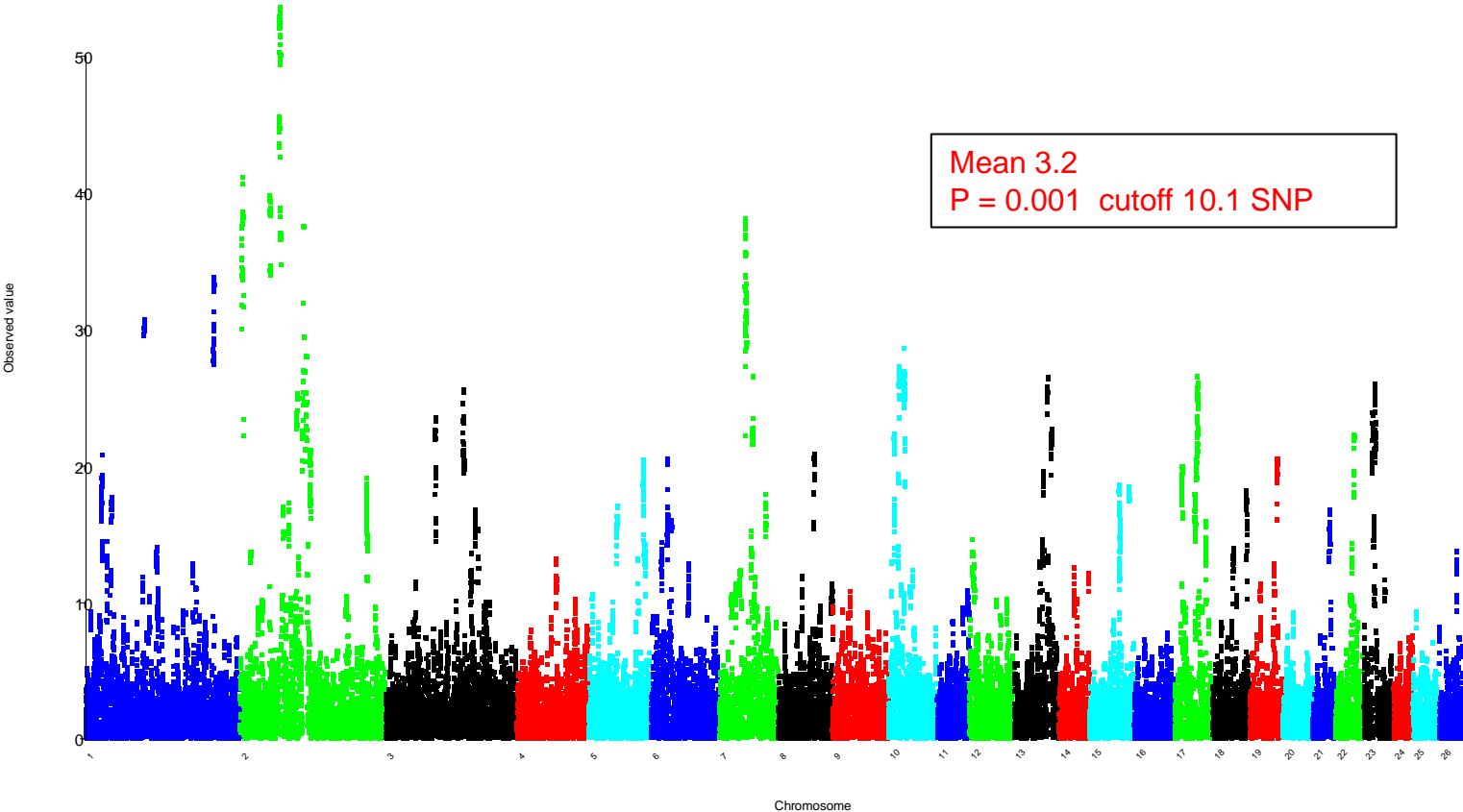
Sheep Hapmap Project for the data.

# Two Suffolk types

Breed	Longest mean HZ score (No. SNP)	Chr	Position	Start	End	Length (Mb)
Australian Suffolk	11.7	2	125,326,201	124,923,303	125,596,219	0.67
	11.3	11	28,950,668	26,956,590	29,487,277	2.53
	10.8	10	29,469,450	29,341,212	29,907,137	0.57
	10.5	13	67,997,490	67,466,397	69,088,841	1.62
	9.8	25	7,203,123	6,682,842	7,637,326	0.95
Irish Suffolk	28.2	14	13,715,221	13,046,967	16,133,656	3.09
	18.3	10	37,886,155	37,775,025	40,399,035	2.62
	18.1	3	165,060,142	164,029,380	165,578,228	1.55
	13.6	13	67,857,725	67,067,820	68,528,151	1.46
	10.9	1	141,808,510	141,348,822	142,494,315	1.15
	12.7	18	39,240,181	38,771,786	39,877,482	1.11
	11.7	2	56,288,982	55,493,630	56,383,269	0.89
	11.6	3	137,891,445	137,556,243	138,674,916	1.12

# Border Leicester data (N = 48)

Manhattan plot of Border Leicester homozygosity scores



# Santa Ines data (N = 47)

