

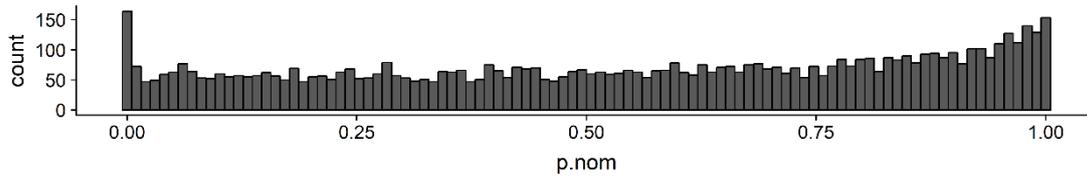
1 **Supplementary material**

2

3

A bimodal p value distribution was observed in GeneiASE output as shown in S1 Fig.1:

4



5

6

Figure S1. Distribution of p values resulted from the beta-binomial test (Static ASE example for spleen in female 1). The uniformity of the p value distribution showed a bimodal (high frequency close to 1 followed by an extensive lower tail close to 0). R machine characteristics for printing float numbers were limited to 6 significant numbers after the decimal point by default which resulted in a high number of absolute zero (0) output-values.

7

8

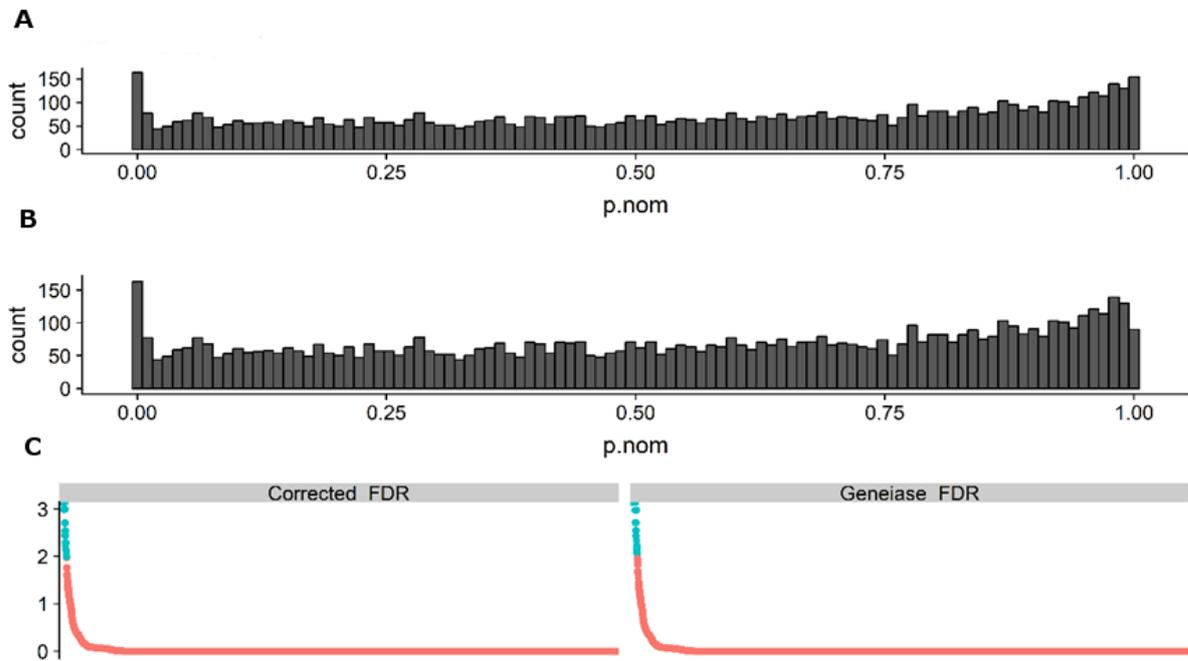
9

10

11

12

13



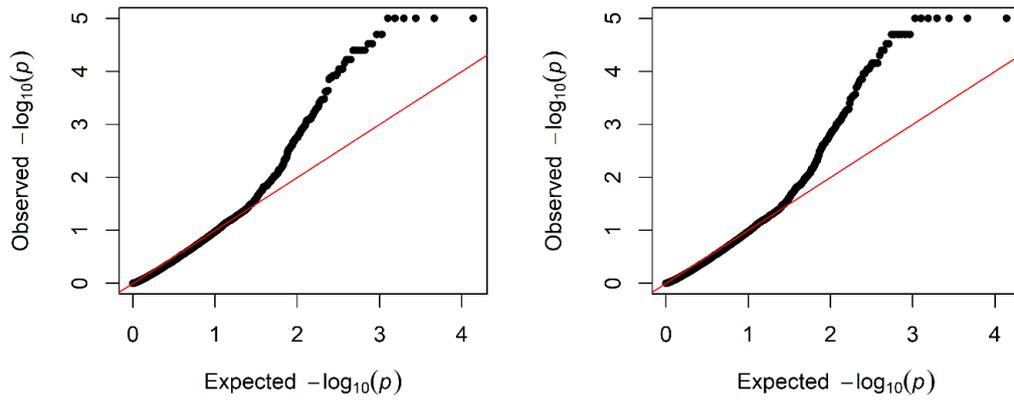
14

15

Figure S2. The figure shows the distribution of p values produced by GeneiASE after cut off introduction (A) and after removal of p value == 1 values (B). The distribution of sorted FDR calculated by GeneiASE and after correction are shown in C (right and left respectively). To correct for multiple testing by recalculation of FDR and eliminating the non-uniformity of p value distribution around 1 the following steps were performed as shown in S1 Fig.2: i) An arbitrary cut off ($1e-7$) was placed over all p values smaller than the said limit including the absolute 0; ii) The genes showing p values equal to 1 were removed from each data frame; iii) The false discovery portion and rate were recalculated in each sample. iv) Gene showing $FDR < 0.1$ selected for further analysis.

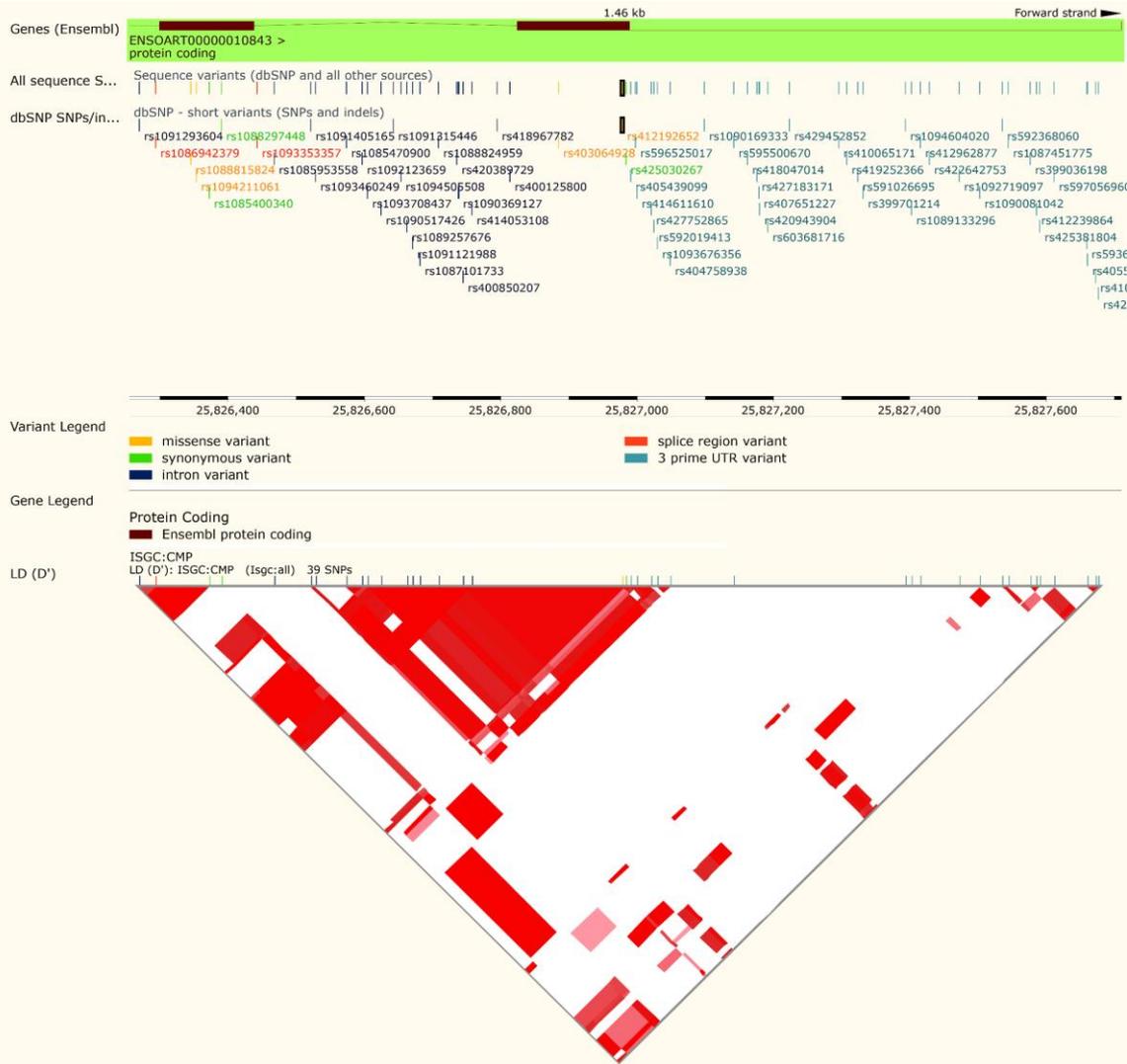
24

25

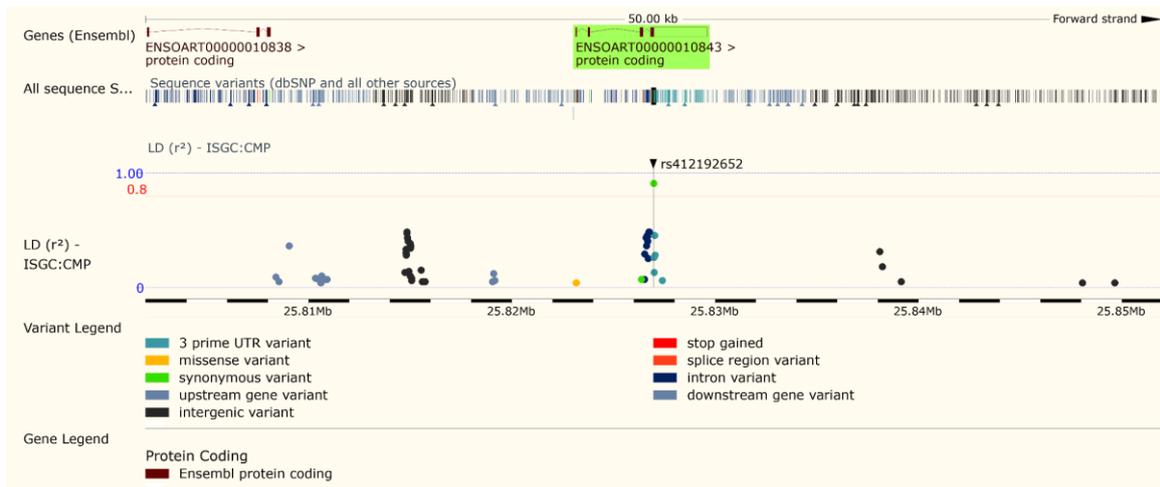


26
27 **Figure S3.** The QQ plot of observed vs expected p values in corrected (left) and GeneiASE's
28 raw (right) calculated outputs.

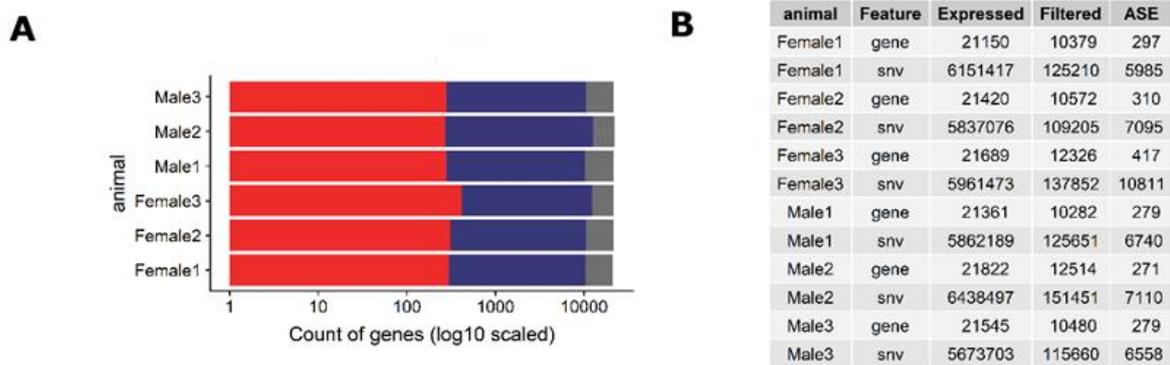
29
30 **A**



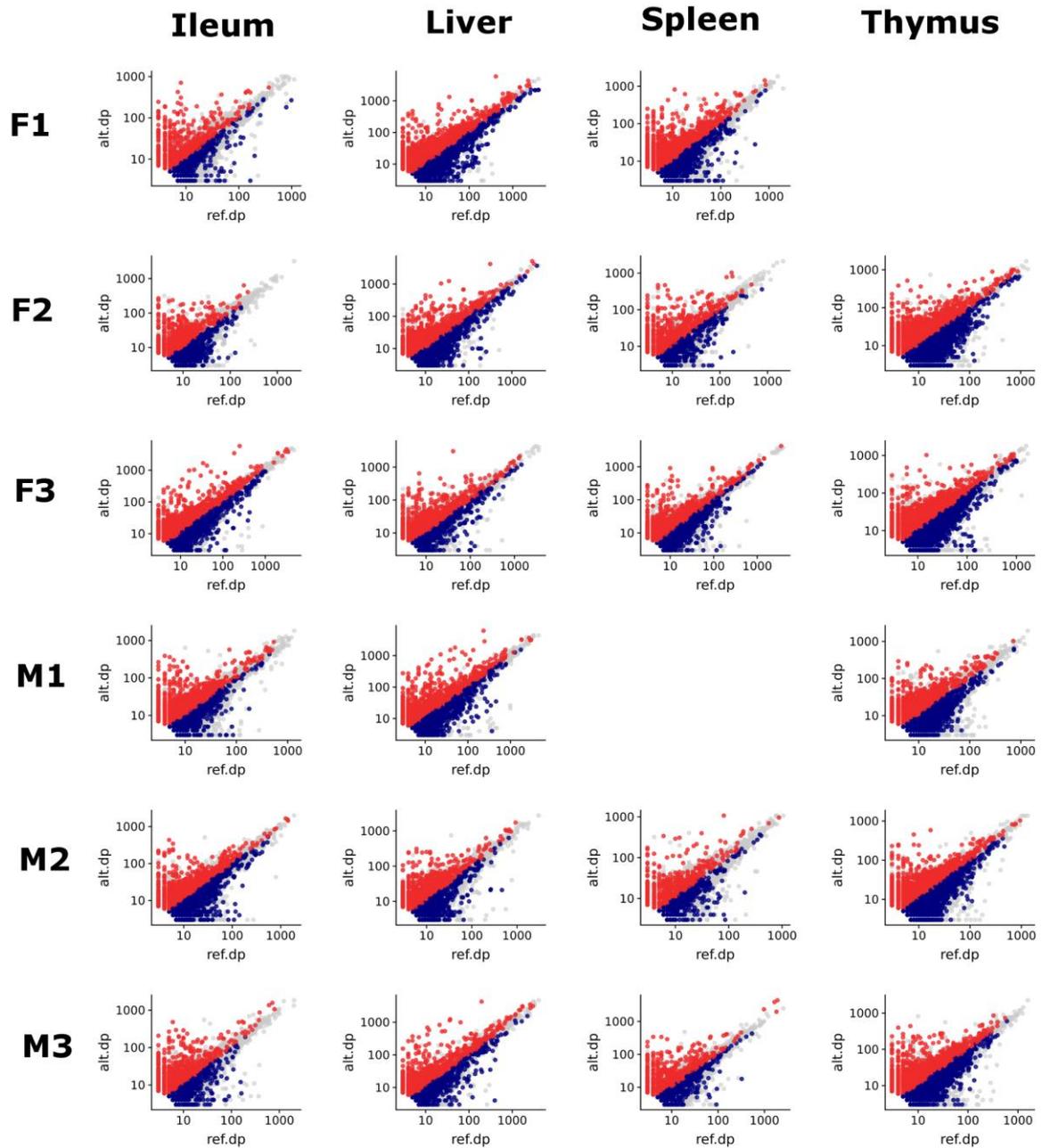
31
32 **B**



33
 34 **Figure S4.** The LD block (A) and LD Manhattan plot (B) of *SAA3* transcript
 35 (*ENSOART0000010843*). The LD block produced as part of International Sheep Genome
 36 Consortium is publicly available on Ensembl using the COMPOSITE population (n = 98).
 37

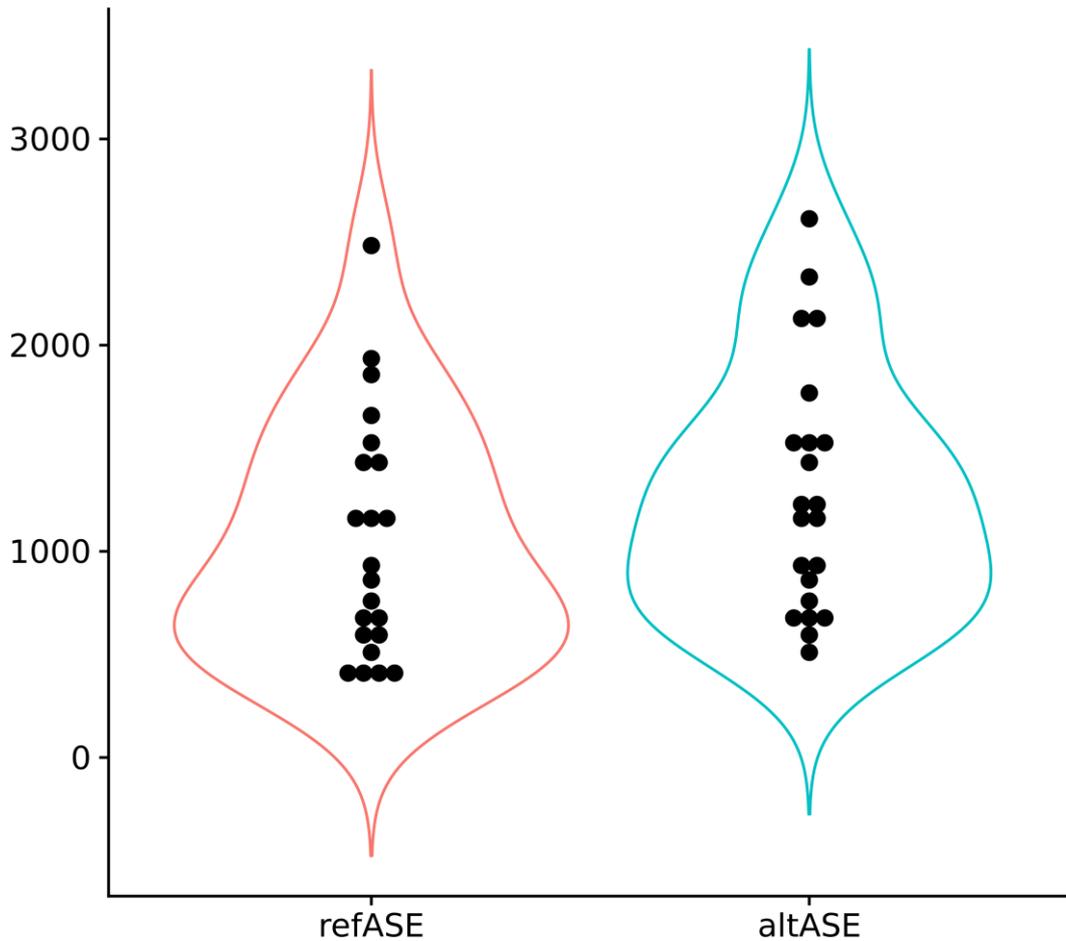


38
 39 **Figure S5.** The proportional distribution of ASE genes and SNVs being expressed across
 40 tissues averaged for each sheep. A) Count of genes in 6 sheep expressed in grey,
 41 informative/filtered in blue and ASE positive in red. B) Break down of genes and SNVs
 42 that are expressed, pass minimum filtration criteria and are ASE positive.
 43



44
45

46 **Figure S6.** Scatter plot of allelic depth (read count) in all 4 tissues of the 6 T×BF sheep.
47 The SNVs passing minimum filtration criteria are coloured in grey and significant ASE
48 (FDR10%) are marked by red and blue. If reference read count > alternate read count
49 the point is shown in blue and vice versa in red. Expression level varies across tissues but
50 does not affect the ASE profile.
51



Count of significant ASE events in every tissue

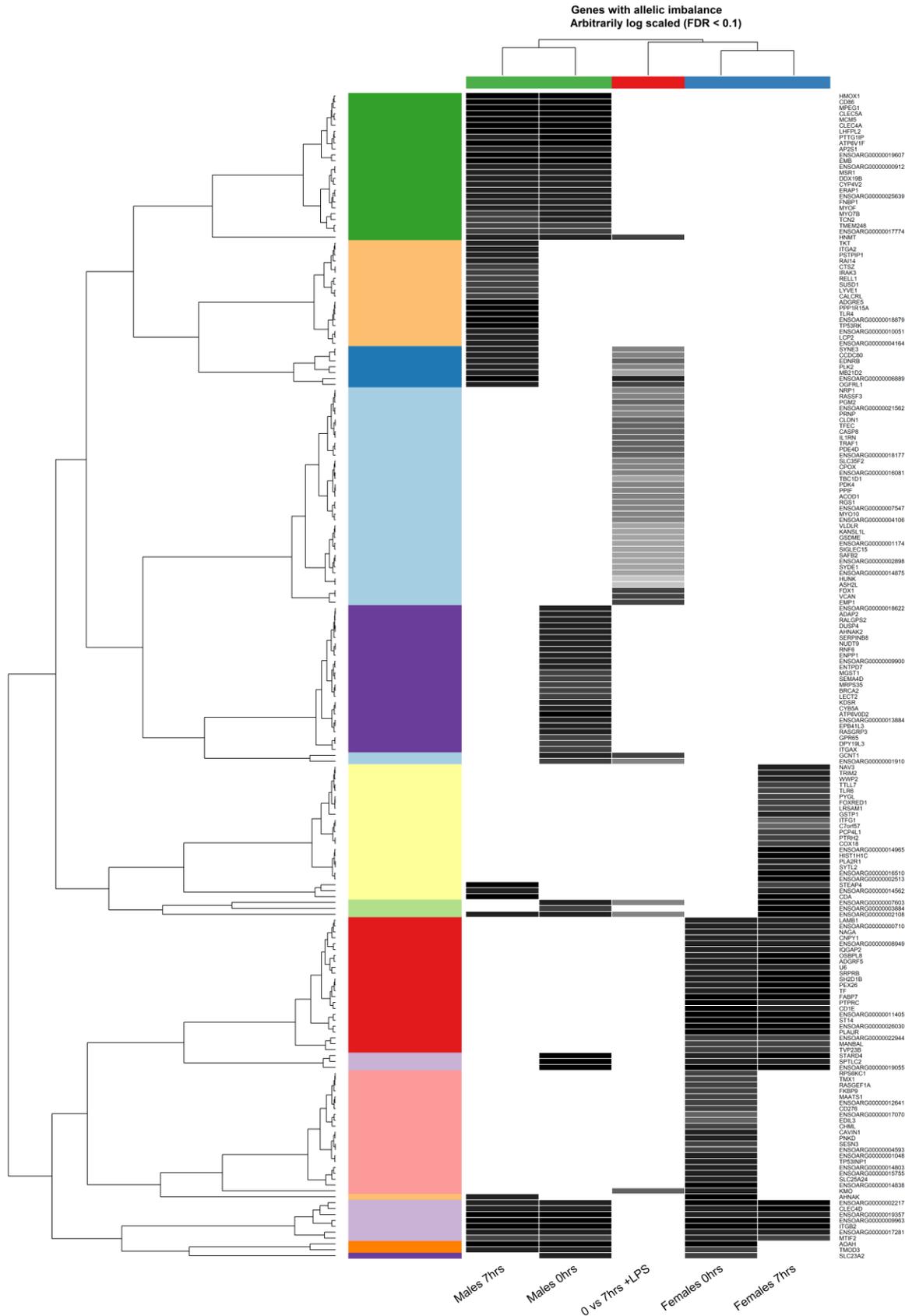
refASE = ref.dp > alt.dp

altASE = alt.dp > ref.dp

$t = -1.9423$, $df = 86$, $p\text{-value} = 0.05537$

52
53
54
55
56
57

Figure S7. Distribution of ASE positive SNVs over the direction of imbalance in the 34 tissues of 6 sheep. Ref.dp > Alt.dp as ASE towards Ref allele and vice versa. There was no significant difference between the count of ASE events given the direction of the allelic imbalance ($p\text{-value} > 0.05$).



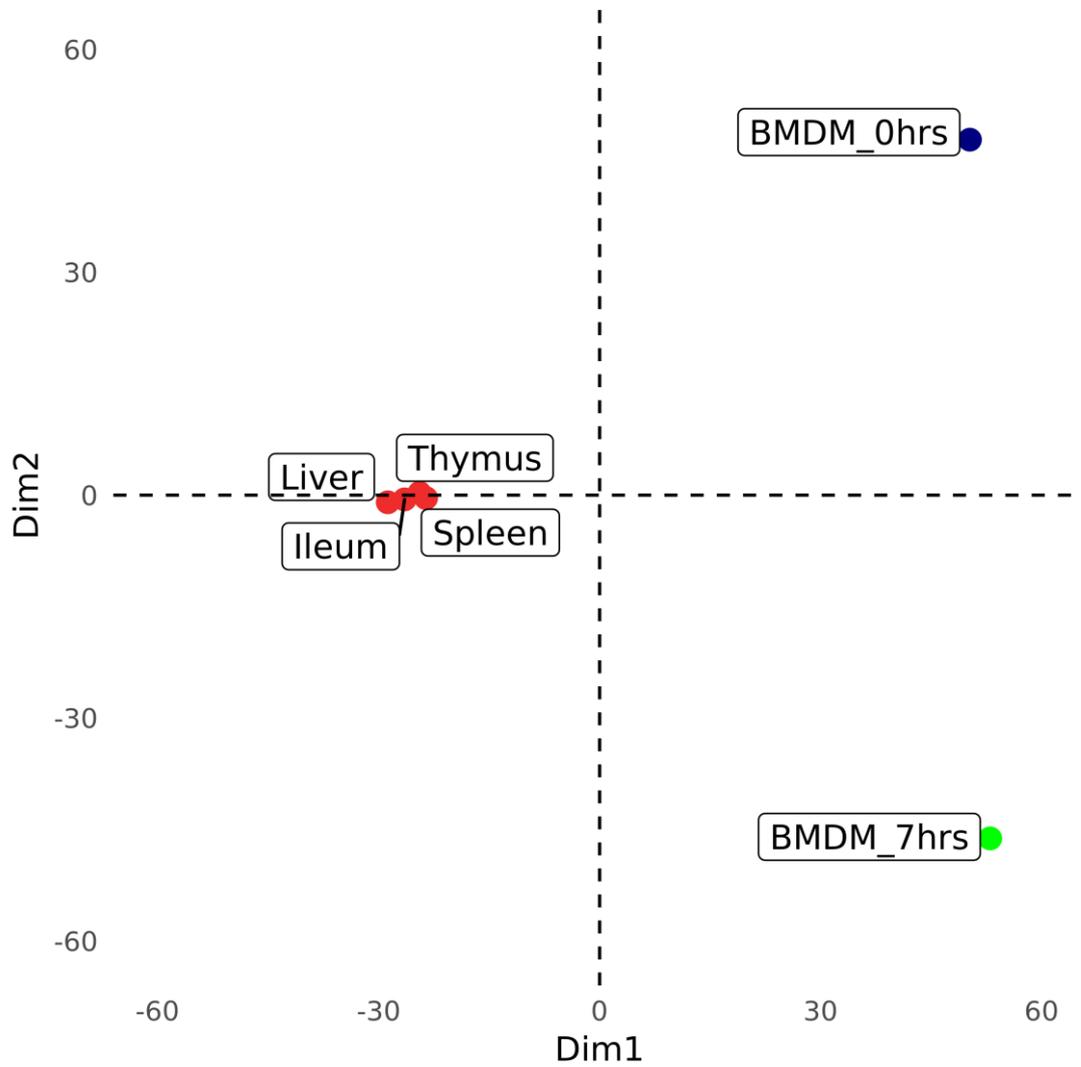
58
59
60
61

Figure S8. The Kmeans cluster analysis of the ASE genes in between static and ICD profiles of BMDMs at 0 and 7hrs samples. Heat map dataframe has been log and centre

62 scaled prior to cluster analysis. The heatmap shows the ASE genes in 5 groups: 0vs 7 hrs
63 ICD set, Males 0hrs , Males 7hrs . Females 0hrs and Females 7hrs.

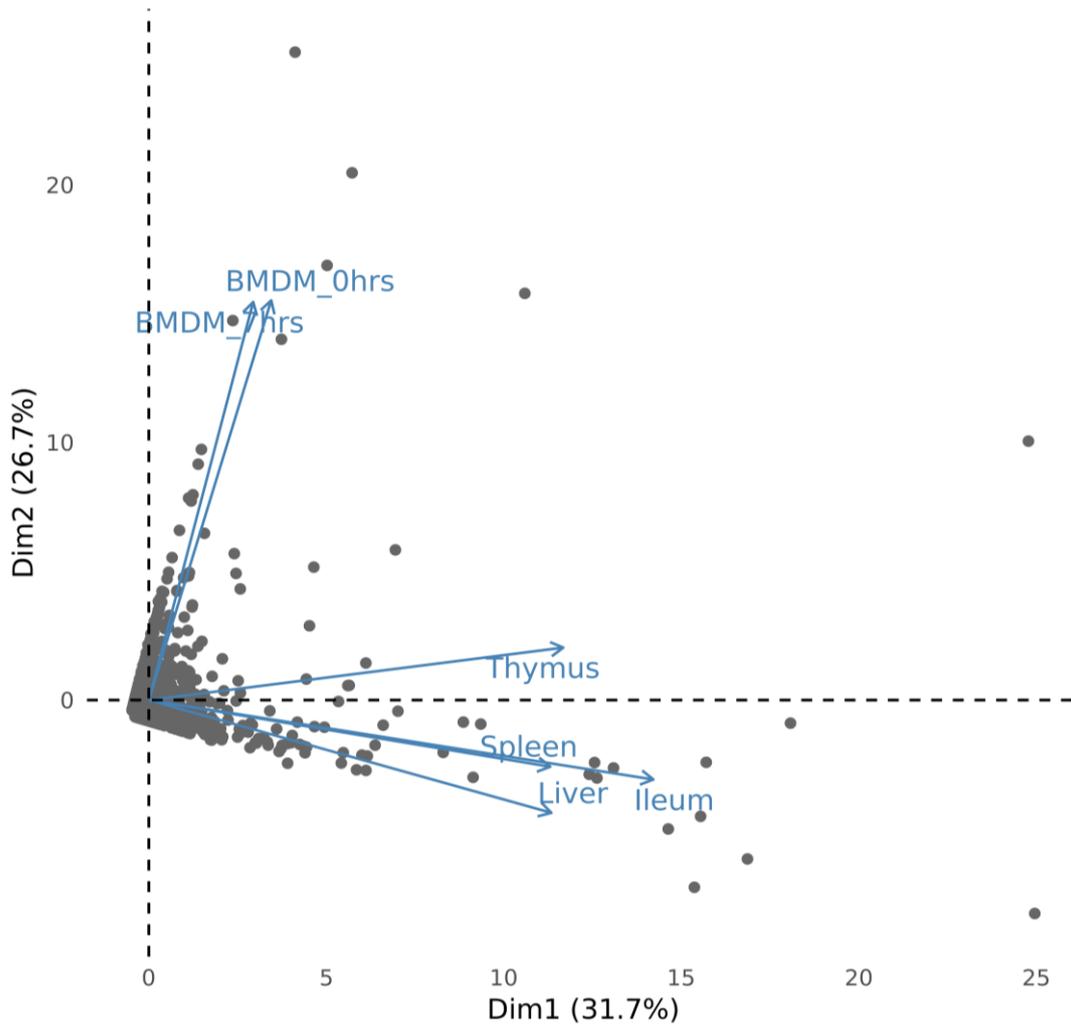
64
65

A



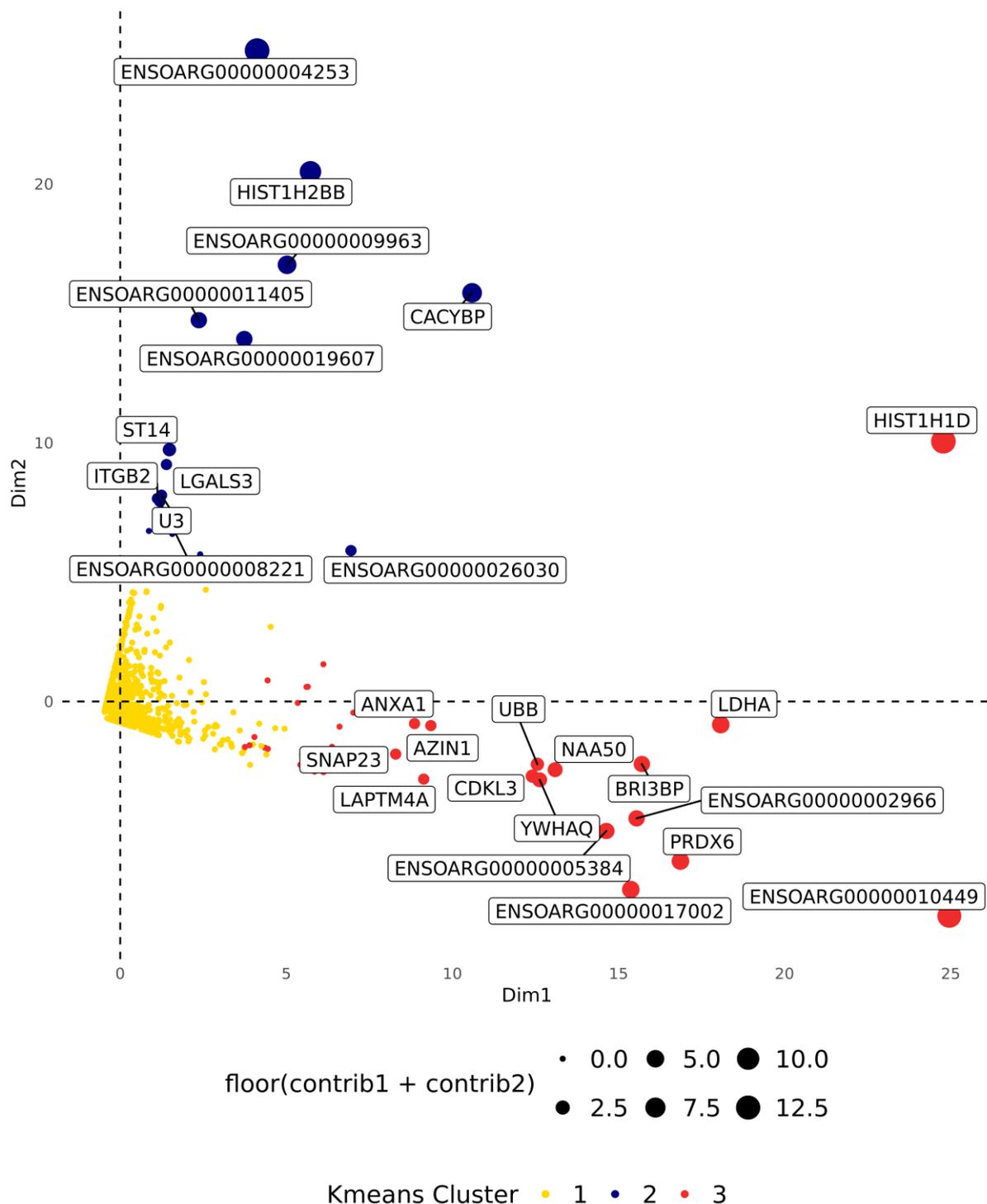
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82

83 **B**



84

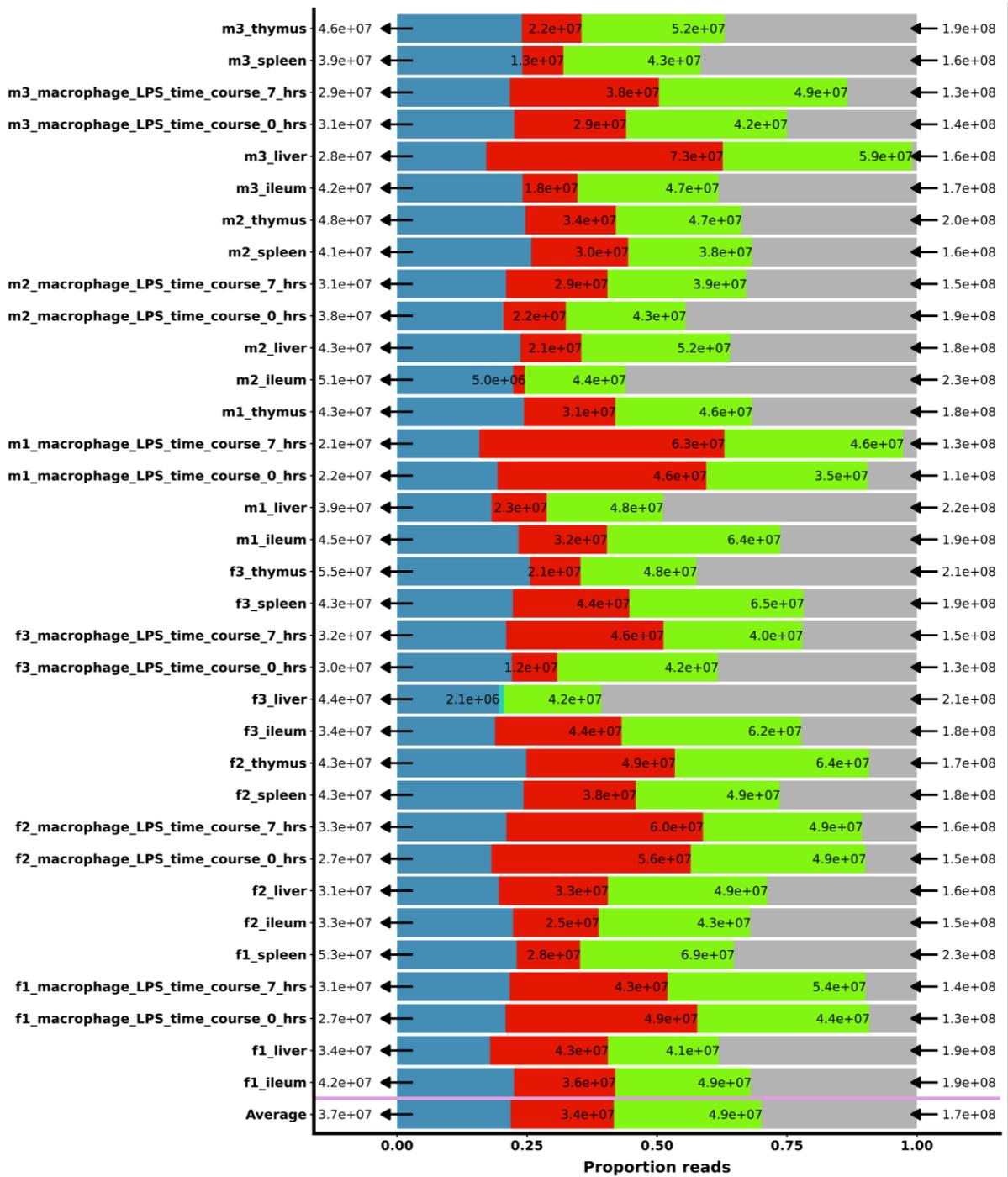
85 C



86

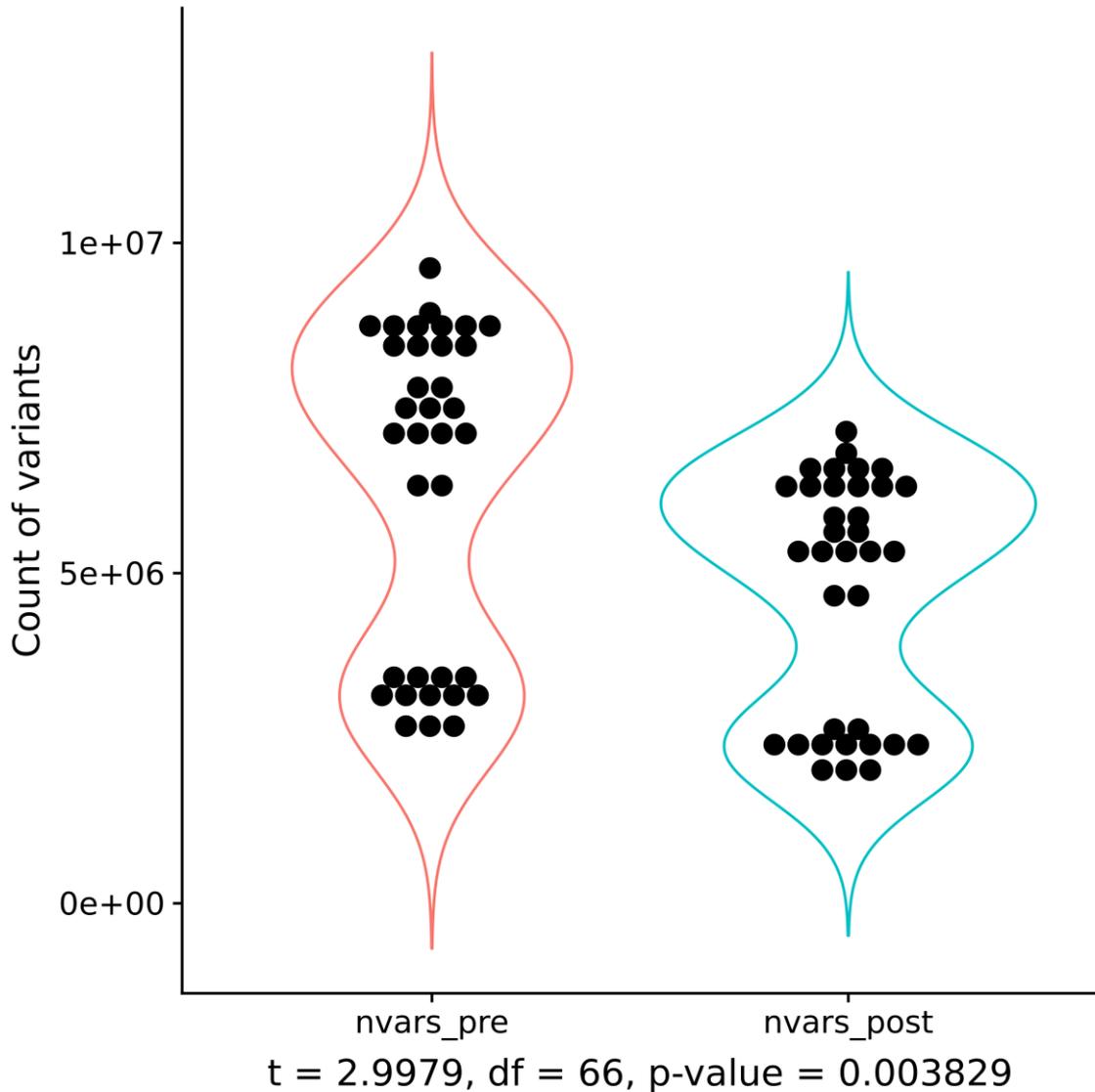
87 **Figure S9.** Principle component analysis and Kmeans clustering of ASE gene profiles of 4
 88 tissues and 1 cell type (2 treatment) in 6 sheep. The gene list comprised of all ASE genes
 89 in each tissue or cell type was used to perform PCA which established the presence of a
 90 hierarchical stratification resulting in 3 groups in the dataset. This grouping was used to
 91 cluster the dataset using K means algorithm with 3 centres. **A)** Samples PCA plot of the
 92 first 2 PCs (Matrix 6 sample rows \times 3000 gene columns) **B)** Genes PCA plot of the first 2
 93 PC (Matrix 3000 gene rows \times 6 sample columns) **C)** Kmeans Clusters overlay on the PC1

94 vs PC2 scatter plot from section B (Genes PCA plot). Gene name labels were included only
 95 for the genes with overall contribution to the first 2 components > 1.
 96
 97
 98



99
 100 **Figure S10.** The read loss proportion during the WASP reference mapping bias removal steps.
 101 The graph shows the input (read counts on the right) and output (read counts on the 2nd left)
 102 RNA-Seq reads from the study before and after WASP processing respectively. Each bar
 103 represents 4 overlaid read proportions labelled by colour and raw read count. The total portion
 104 of reads coloured in grey shows 100% of the reads processed by WASP (input 170 million
 105 reads on average). The green portion shows the non-intersecting reads (and their duplicates)
 106 mapped within the range of Ensemble gene coordinate boundaries (exonic, intronic and UTRs

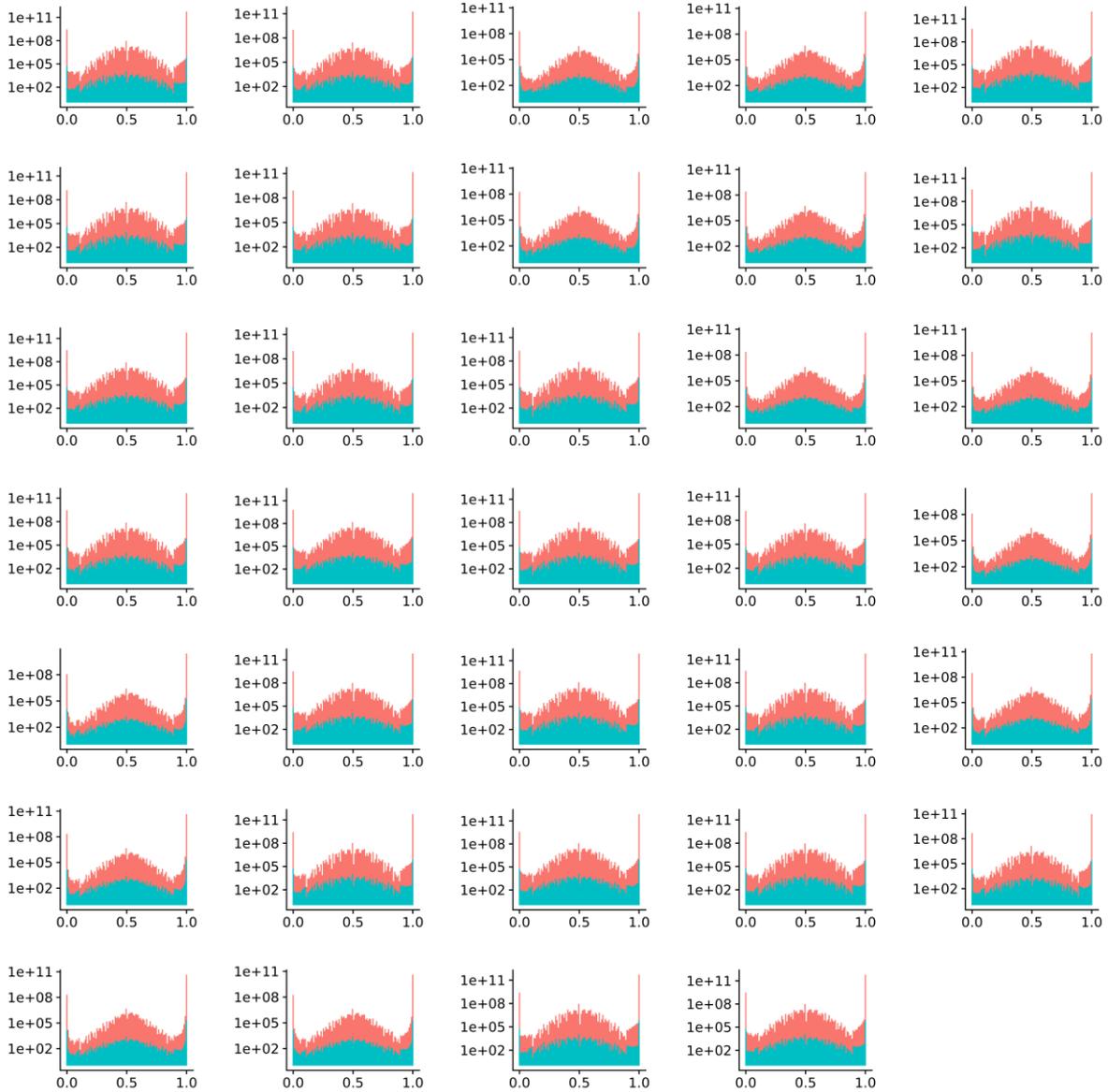
107 reads 49 million reads on average). The red portion of the reads represents reads intersecting
 108 bi-allelic heterozygote sites and selected for synthetic read production step (34 million reads
 109 on average). The blue portion of the graph represents the final output of WASP's pipeline
 110 (reference biased and duplicated reads removed from red and green portions respectively). The
 111 output portion (37 million reads on average) was used in GATK ASEReadCounter for allelic
 112 read counts. The last row on the bottom shows the average of these proportions across both
 113 tissue and cell type samples. Reference biased reads count for average 20% of input RNA-Seq
 114 data.



115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126

Figure S11. Number of SNVs before and after WASP reference bias correction. There is a significant reduction in number of variants captured (25% on average 1.5×10^7) between two steps. Tissues samples dots mainly present within the upper quantile and BMDMs in the lower. Average number of variants beforehand = 6,325,570 and afterwards = 4,729,639.

127
128
129



130
131
132
133
134
135
136
137
138
139
140
141

Figure S12. The comparison of reference allele count ratio before and after WASP mapping bias correction. The distribution of refRatio (refCount/refCount+altcounts) in all the tissue and cell samples before (in red) and after WASP (in green). The overall counts, left skew (viz. range 0.75-1) and refRatio=1 inflation is significantly reduced after WASP bias correction in all samples.

Table S1. Mean allelic imbalance ($\log_2(\text{ASE}_{+LPS}/\text{ASE}_{-LPS})$) in all 6 sheep. Only 1 gene (*VCAN* including 7 SNVs) exhibited inducible ASE which was shared between 3 sheep (in bold), while two (*GCNT1* and *SYNE3*) showed inducible ASE in two sheep. Diverse and highly individual ASE response to LPS was observed in each sheep.

Gene	Female1	Female2	Female3	Male1	Male2	Male3
<i>ACOD1</i>	-	-	-	6.8	-	-
<i>ASH2L</i>	-	-	-	-	3.9	-

Unbiased ovine allele-specific expression

<i>CASP8</i>	-	-	-	-	-	9.8
<i>CCDC80</i>	-	-	9.1	-	-	-
<i>CLDN1</i>	-	-	-	-	-	10.3
<i>CPOX</i>	-	-	-	-	-	6.5
<i>EDNRB</i>	-	-	-	-	-	11.4
<i>EMP1</i>	-	-	-	-	-	23.9
<i>ENSOARG00000001174</i>	-	-	-	-	5.0	-
<i>ENSOARG00000001910</i>	-	-	-	-	-	7.8
<i>ENSOARG00000002108</i>	-	-	-	-	-	8.0
<i>ENSOARG00000002898</i>	-	-	-	-	-	5.3
<i>ENSOARG00000004106</i>	-	-	-	-	-	7.5
<i>ENSOARG00000006889</i>	-	-	31.9	-	-	-
<i>ENSOARG00000007547</i>	-	-	-	-	8.2	-
<i>ENSOARG00000007603</i>	-	-	-	6.9	-	-
<i>ENSOARG00000014875</i>	-	-	-	-	-	5.6
<i>ENSOARG00000016081</i>	6.4	-	-	-	-	-
<i>ENSOARG00000018177</i>	-	-	11.0	-	-	-
<i>ENSOARG00000021562</i>	-	-	-	-	-	9.0
<i>FDX1</i>	-	-	-	-	17.3	-
<i>GCNT1</i>	-	-	-	16.9	-	30.4
<i>GSDME</i>	-	-	-	4.6	-	-
<i>HNMT</i>	-	-	-	-	-	15.1
<i>HUNK</i>	-	-	-	-	-	4.0
<i>IL1RN</i>	-	-	-	-	-	9.7
<i>KANSL1L</i>	-	-	-	-	-	4.7
<i>KMO</i>	-	-	-	-	-	9.7
<i>MB21D2</i>	-	-	-	5.3	-	-
<i>MYO10</i>	-	-	-	-	-	7.8
<i>NRP1</i>	-	-	9.3	-	-	-
<i>OGFRL1</i>	-	-	-	-	-	16.5
<i>PDE4D</i>	-	-	-	-	-	11.7
<i>PDK4</i>	-	-	-	6.9	-	-
<i>PGM2</i>	-	-	-	-	-	9.5
<i>PLK2</i>	-	-	-	-	-	6.7
<i>PPIF</i>	6.9	-	-	-	-	-
<i>PRNP</i>	-	-	-	-	-	8.8
<i>RASSF3</i>	-	-	-	-	9.2	-
<i>RGS1</i>	-	-	-	-	-	8.4
<i>SAFB2</i>	-	-	-	-	-	5.3
<i>SIGLEC15</i>	-	-	-	4.8	-	-
<i>SLC35F2</i>	-	-	-	-	6.5	-
<i>SYDE1</i>	-	-	-	-	-	5.1
<i>SYNE3</i>	-	-	7.2	-	-	10.7
<i>TBC1D1</i>	-	-	-	6.3	-	-
<i>TFEC</i>	-	-	10.3	-	-	-
<i>TRAF1</i>	12.1	-	-	-	-	-

Unbiased ovine allele-specific expression

142	<i>VCAN</i>	23.5	-	-	-	14.0	13.9
	<i>VLDLR</i>	-	-	-	-	4.7	-