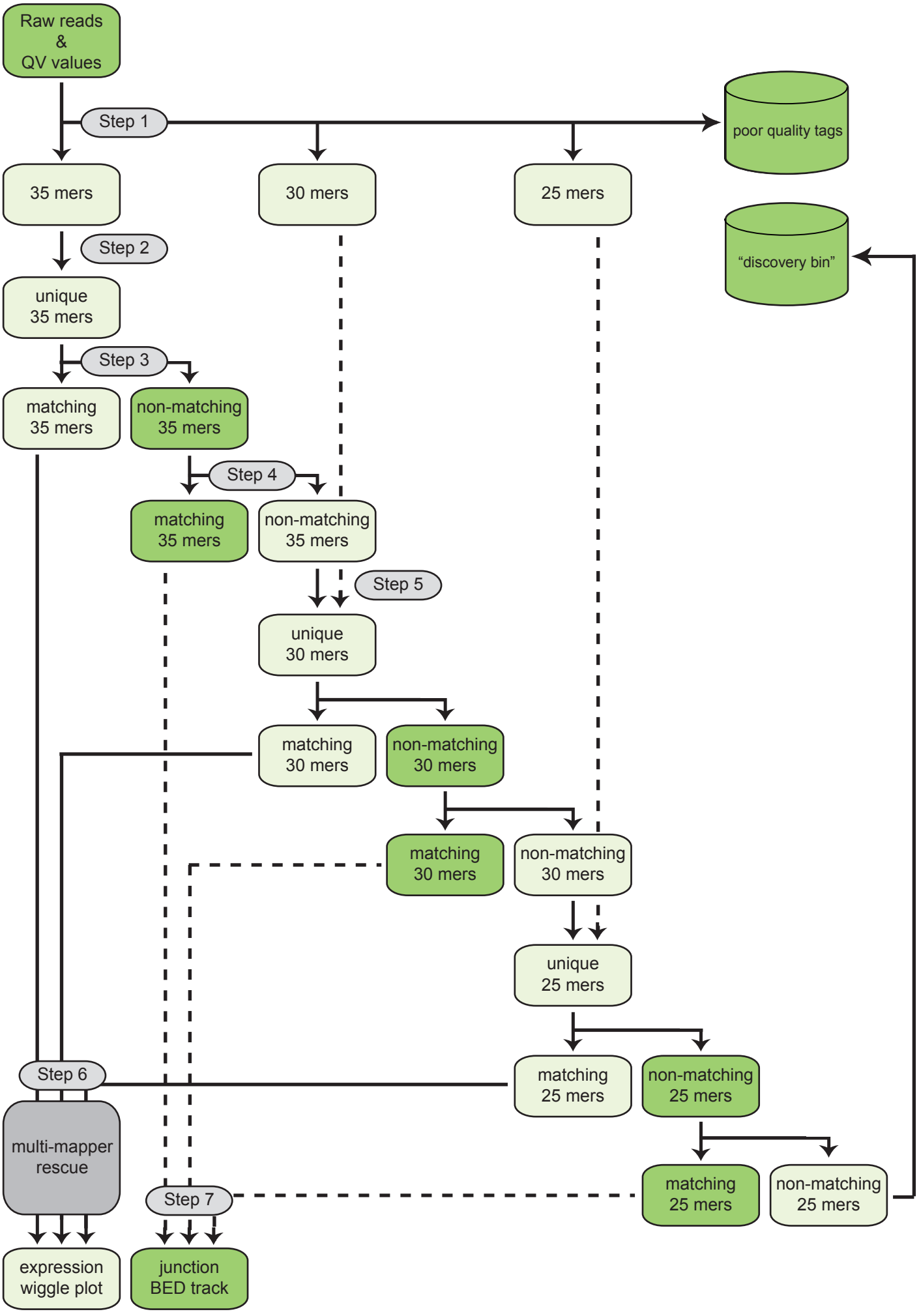
Stem cell transcriptome profiling via massive-scale mRNA sequencing

Nicole Cloonan, Alistair R R Forrest, Gabriel Kolle, Brooke B A Gardiner, Geoffrey J Faulkner, Mellissa K Brown, Darrin F Taylor, Anita L Steptoe, Shivangi Wani, Graeme Bethel, Alan J Robertson, Andrew C Perkins, Stephen J Bruce, Clarence C Lee, Swati S Ranade, Heather E Peckham, Jonathan M Manning, Kevin J McKernan & Sean M Grimmond

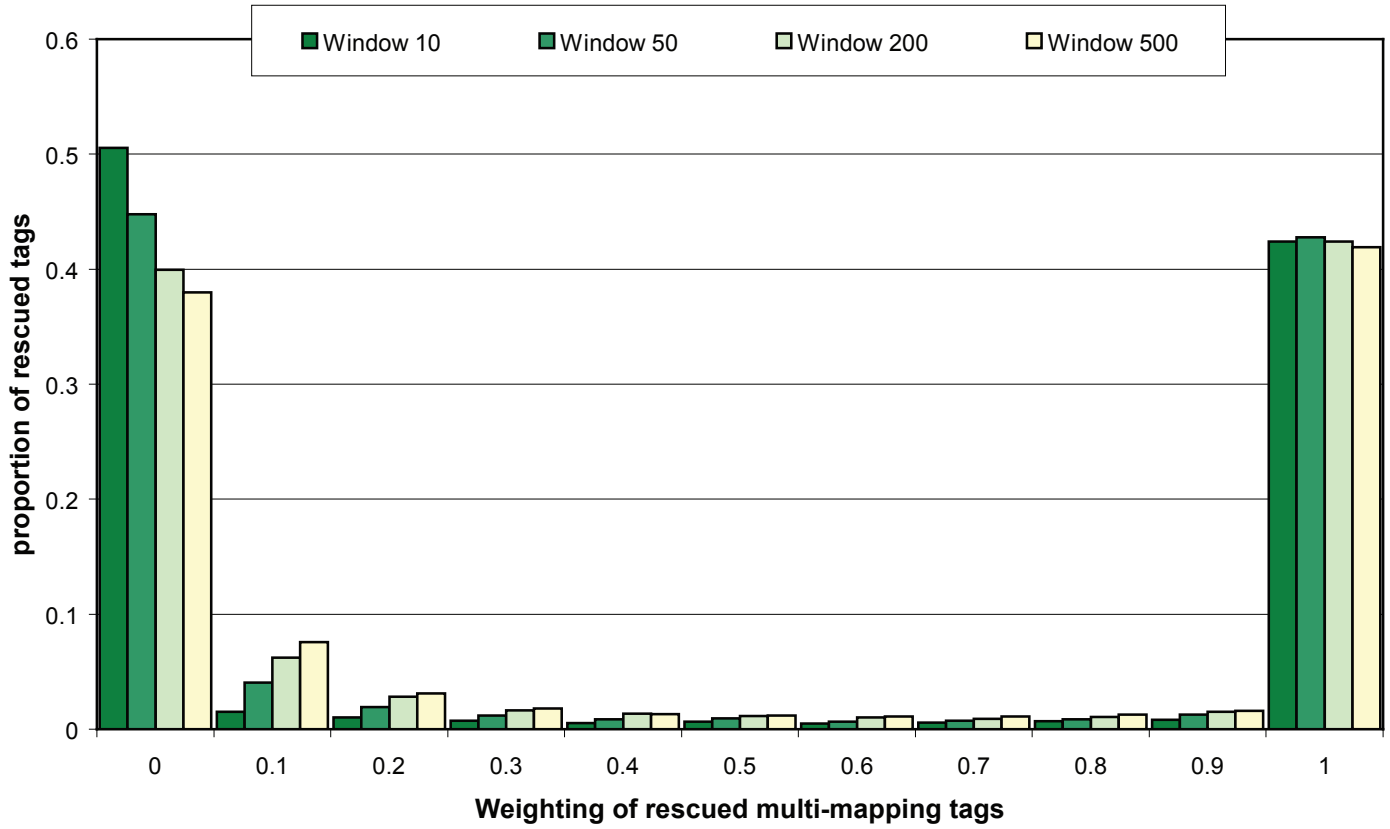
Supplementary figures and text:

Supplementary Figure 1.



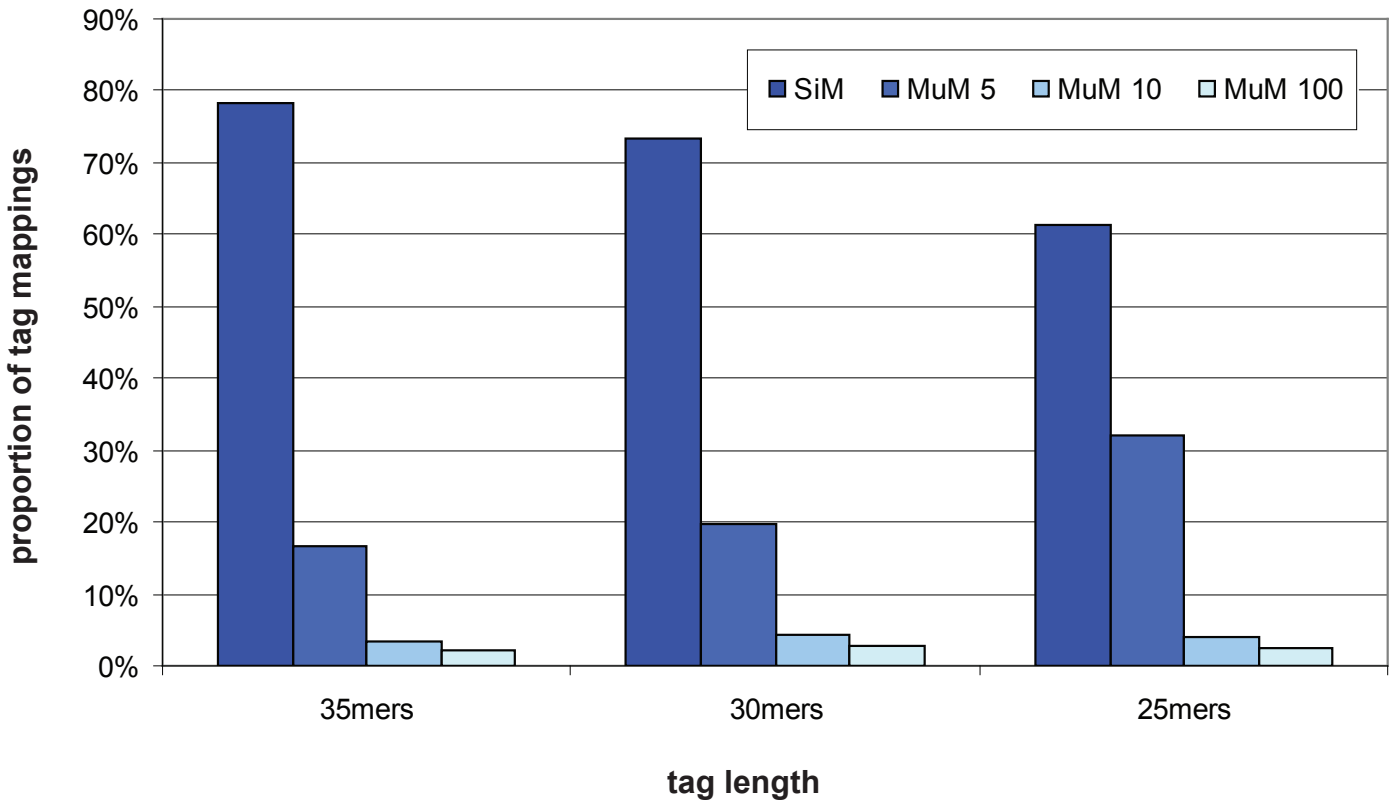
Supplementary Figure 1. Mapping strategy for SQRL gene expression data. Step 1: Sort tags by their raw quality scores (Phred scores). Tags were considered high quality if less than 5 bases were called with a Phred score of less than 10. If a tag was not considered high quality, the last 5nt were removed, and the tag was re-assessed. Tags that were considered poor quality at 25nt were discarded. Step 2: 35mers were clustered and condensed to unique tags, retaining tag frequency information. Step 3: Unique 35mers were mapped to the mm9 genome, with at most 3 colour-space mismatches. The SOLiD sequencing system measures the relationships between nucleotides (referred to as colour-space), rather than the nucleotides directly (base-space). Because any individual nucleotide has two relationships, one to the preceding nucleotide, and one to the subsequent nucleotide, every base-space nucleotide position is measured twice. All mapping is carried out in colour-space, allowing up to 3 colour-space mismatches. This equates to at most 2 adjacent base-space mismatches between the tag and the reference. Tags mapping to more than 100 genomic locations were discarded. Step 4: 35mer tags not mapping to the genome were mapped to the informative junction library – a set of sequences where the 30nt donor and 30nt acceptor sequence both match the genome individually (but not together), and also don't match with another junction within the library. Step 5: The last 5nt were removed from tags not matching the junction library, and combined with sequences that were considered good quality at 30nt. Steps 2-5 are repeated for 30mers and 25mers. Tags that do not map to the genome at 25mers are potential candidates for novel junctions. Step 6: Genome mappings for 35, 30, and 25mers are sorted into single-mappers (SiMs) or multi-mappers (MuMs), based on the number of times a tag maps at its highest stringency. (also see **Supplementary Fig. 2**). Wiggle plots for the positive and negative strand of each sample are created from this MuM-rescued data. Step 7: Junction mappings for 35, 30, and 25mers are sorted into single-mappers (SiMs) or multi-mappers (MuMs), based on the number of times a tag maps at its highest stringency. SiMs are used to create junction BED tracks for each sample. The proportions of single-mapping and multi-mapping tags at different lengths can be found in **Supplementary Fig. 3**.

Supplementary Figure 2.



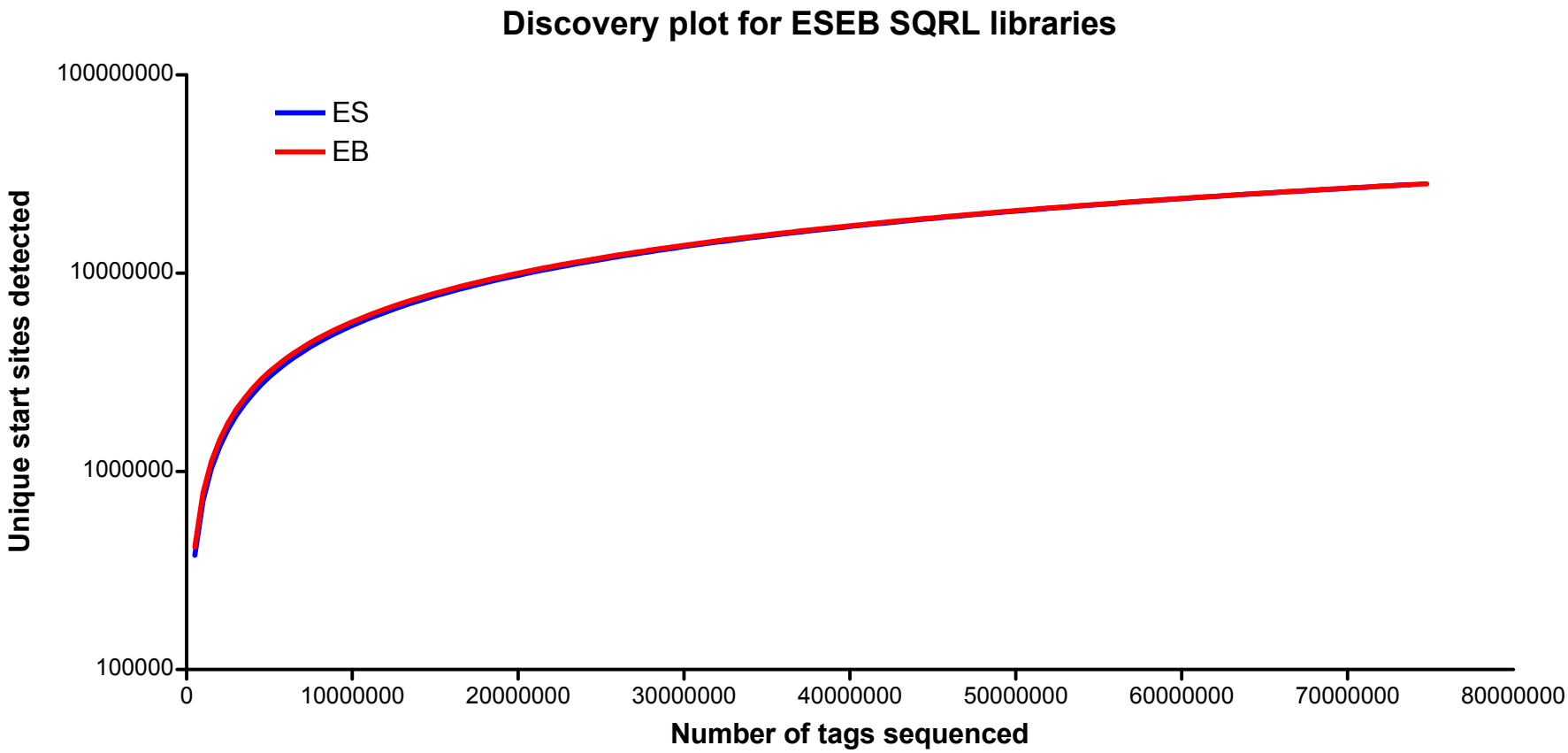
Supplementary Figure 2. Evaluation of MuM rescue of SQRL gene expression tags. MuMs are assigned to their most likely source probabilistically, based on a modified version of our previously described algorithm (Faulkner *et al* 2008). The use of different windows of expression surrounding each potential mapping site was investigated, finding that for SQRL gene expression data, small windows outperformed larger ones. Using a window of 10nt, more than 90% of MuM locations could be unambiguously assigned as either on (a weighting of 1) or off (a weighting of 0). Weightings for each MuM location are multiplied by the frequency of the MuM tag in the SQRL library to obtain a final tag count for mapping at that position.

Supplementary Figure 3.



Supplementary Figure 3. Tag matching at multiple stringencies. Distribution of tags which are map uniquely (SiM), <5 times (MuM 5), <10 times (MuM 10) or <100 times (Mum) at mapped as 35mers, 30mers and 25mers (with up to 3 mismatch in colour space). There is only a slight increase in the mismatch rate when mapping at 25mers.

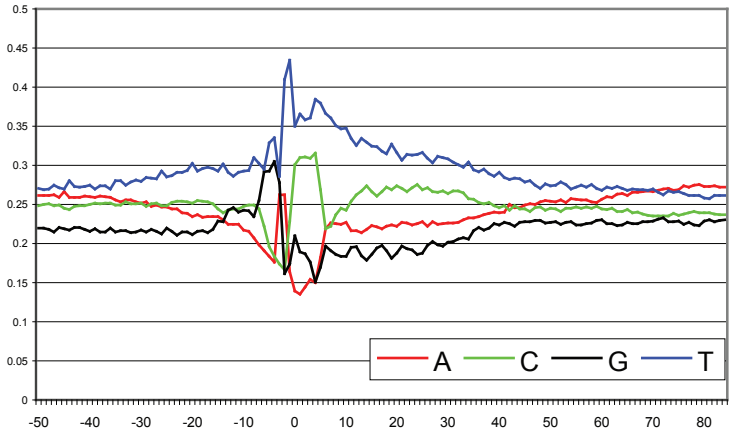
Supplementary Figure 4.



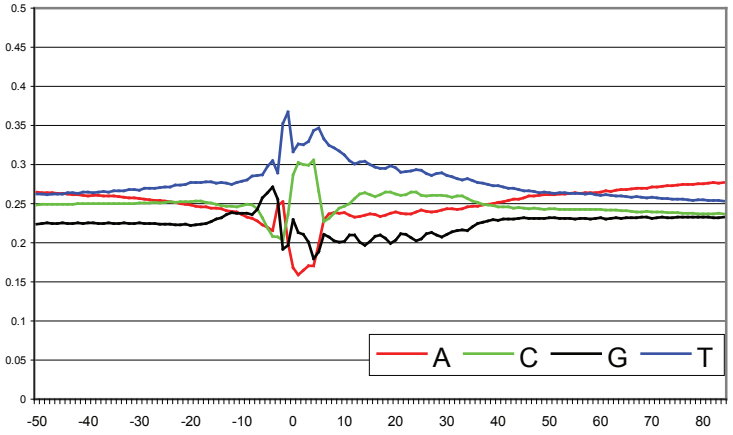
Supplementary Figure 4. Cumulative discovery plots of ESC and EB sequencing libraries. All single-mapping tags (either because they map once at their highest stringency, or because the multi-mapping rescue algorithm is able to unambiguously map them) of any length were randomly selected and the cumulative number of unique start sites was determined. This graph clearly shows that at this depth of sequencing, the number of unique start sites discovered has reached a plateau.

Supplementary Figure 5.

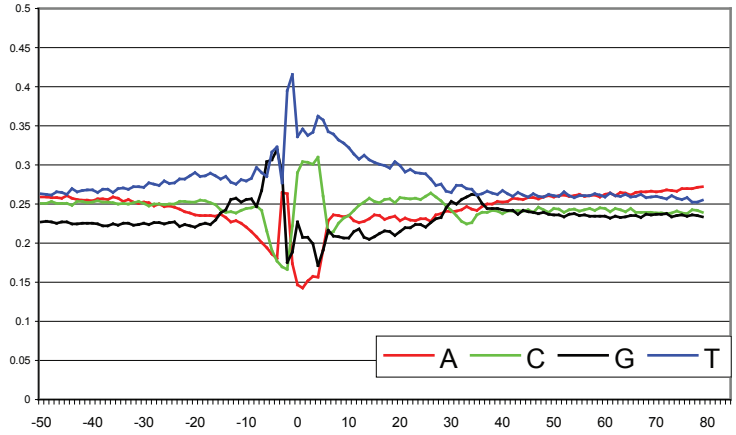
35nt tags (all tags)



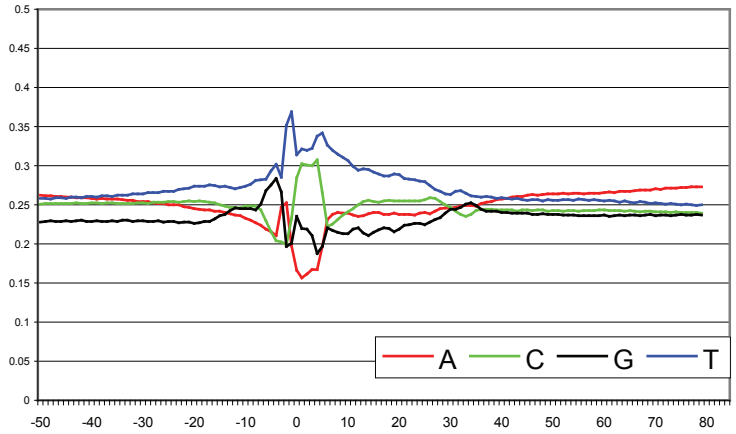
35nt tags (unique tags only)



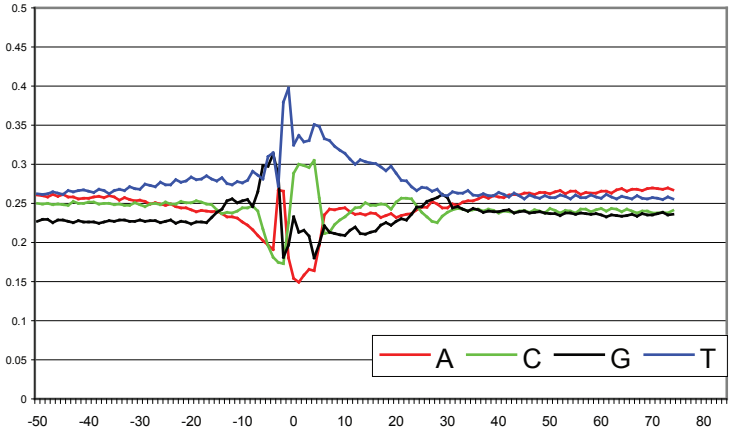
30nt tags (all tags)



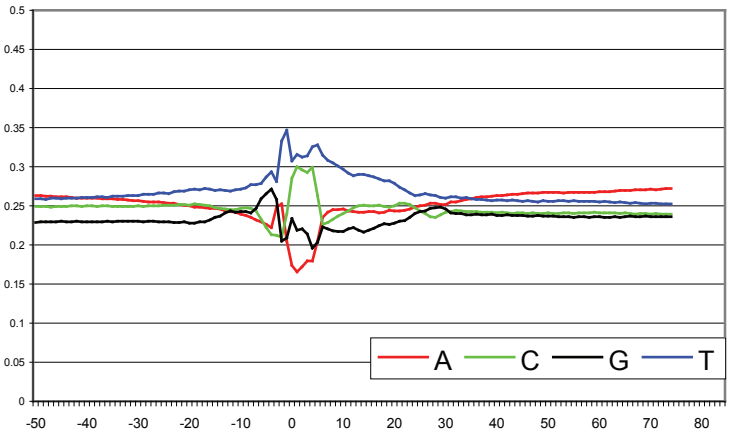
30nt tags (unique tags only)



25nt tags (all tags)

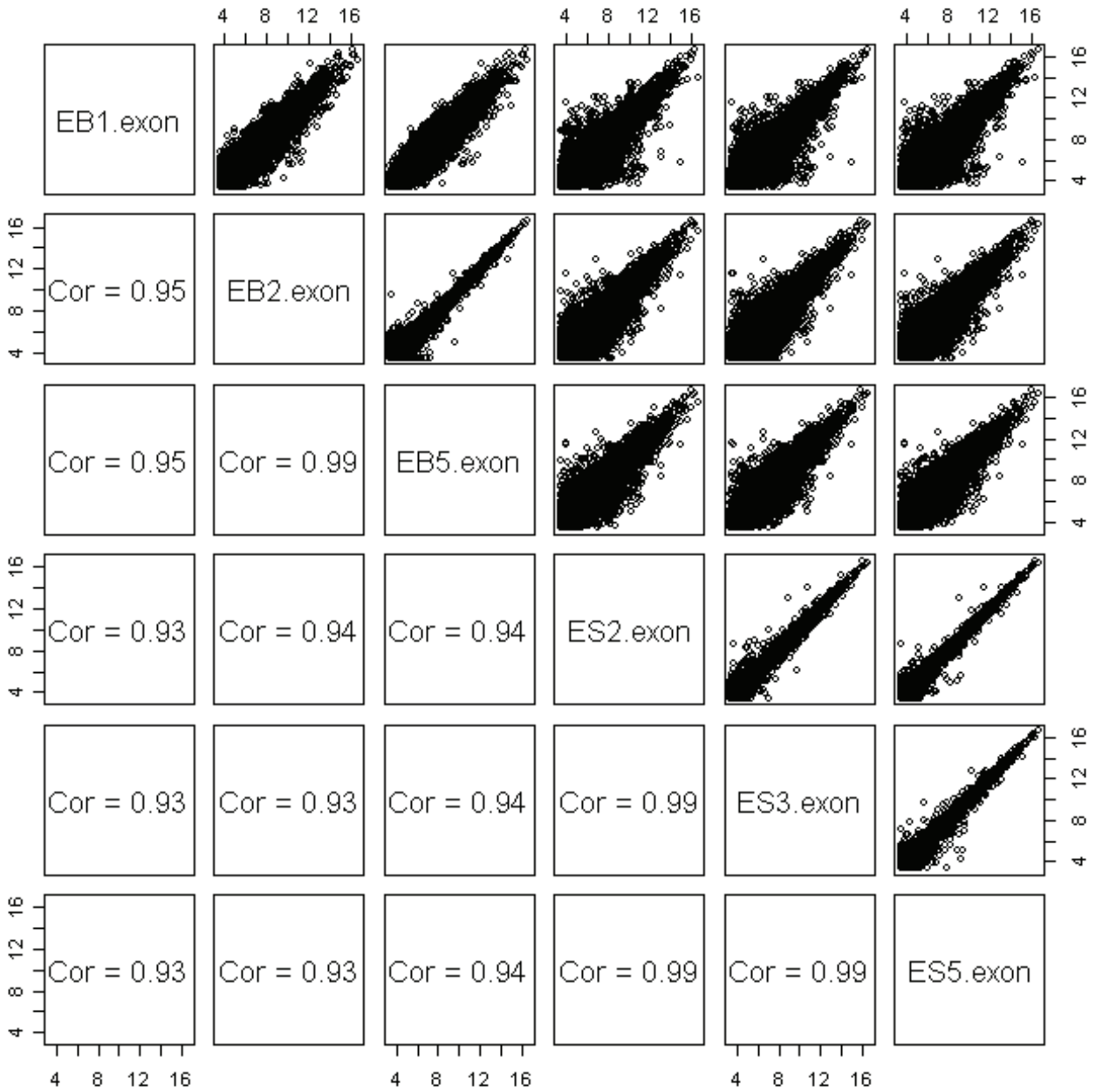


25nt tags (unique tags only)

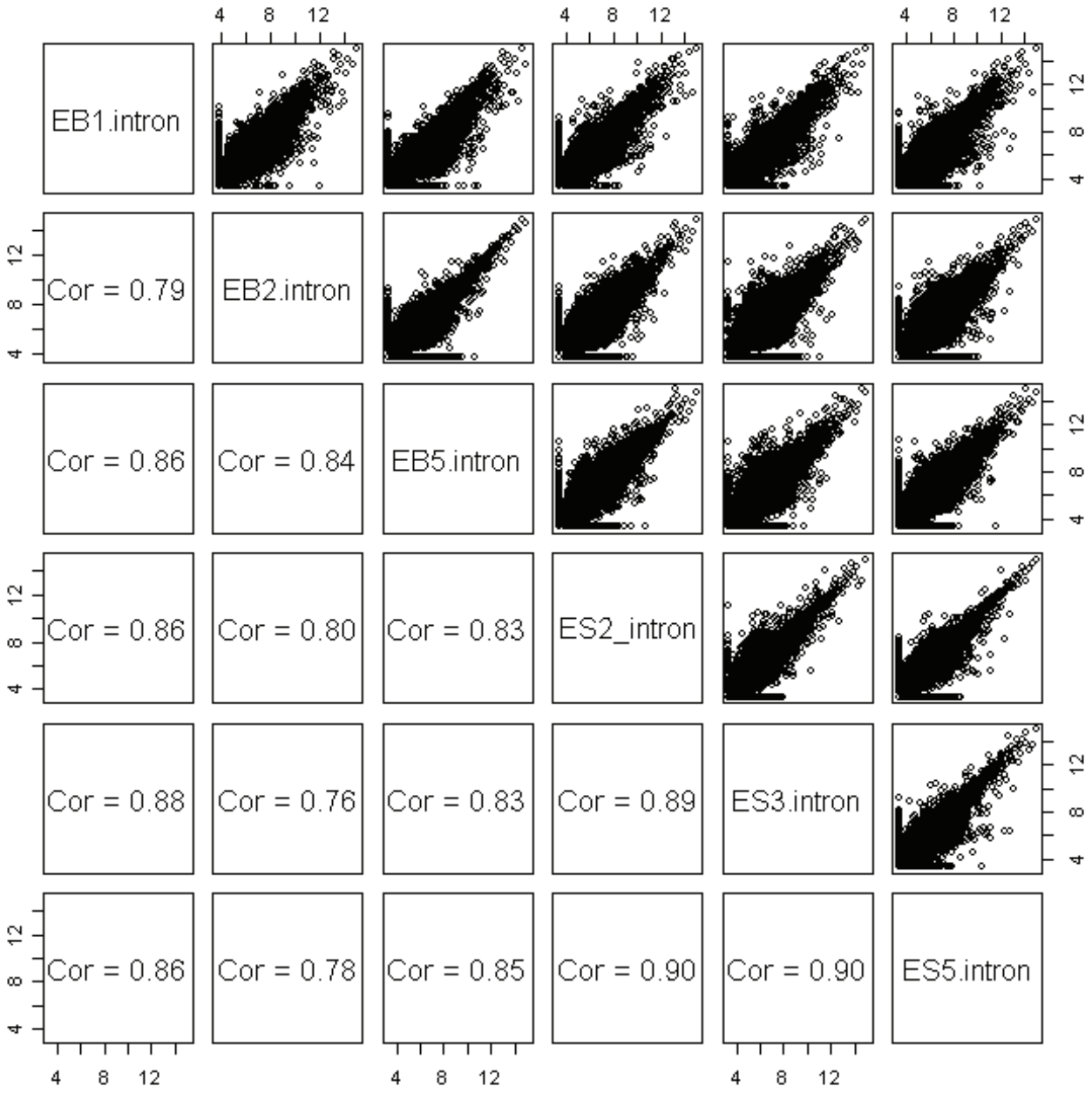


Supplementary Figure 5. Tag bias by sequence content. All tags were matched to the genome and the ratios of A, C, G and T were calculated for each position relative to the tag. 1 indicates tag start, negative numbers are derived from genomic sequence upstream of the matched tag and positive numbers represent the tag and downstream sequence. Plots show that the type of bias is similar at the 3 matching lengths, and only differs in magnitude in a comparison between all tags and those derived from unique start sites. The major source of bias occurs at -3 to 1 and 0-5 which is likely to represent bias in the random-hexamer FDV primer.

Supplementary Figure 6a.

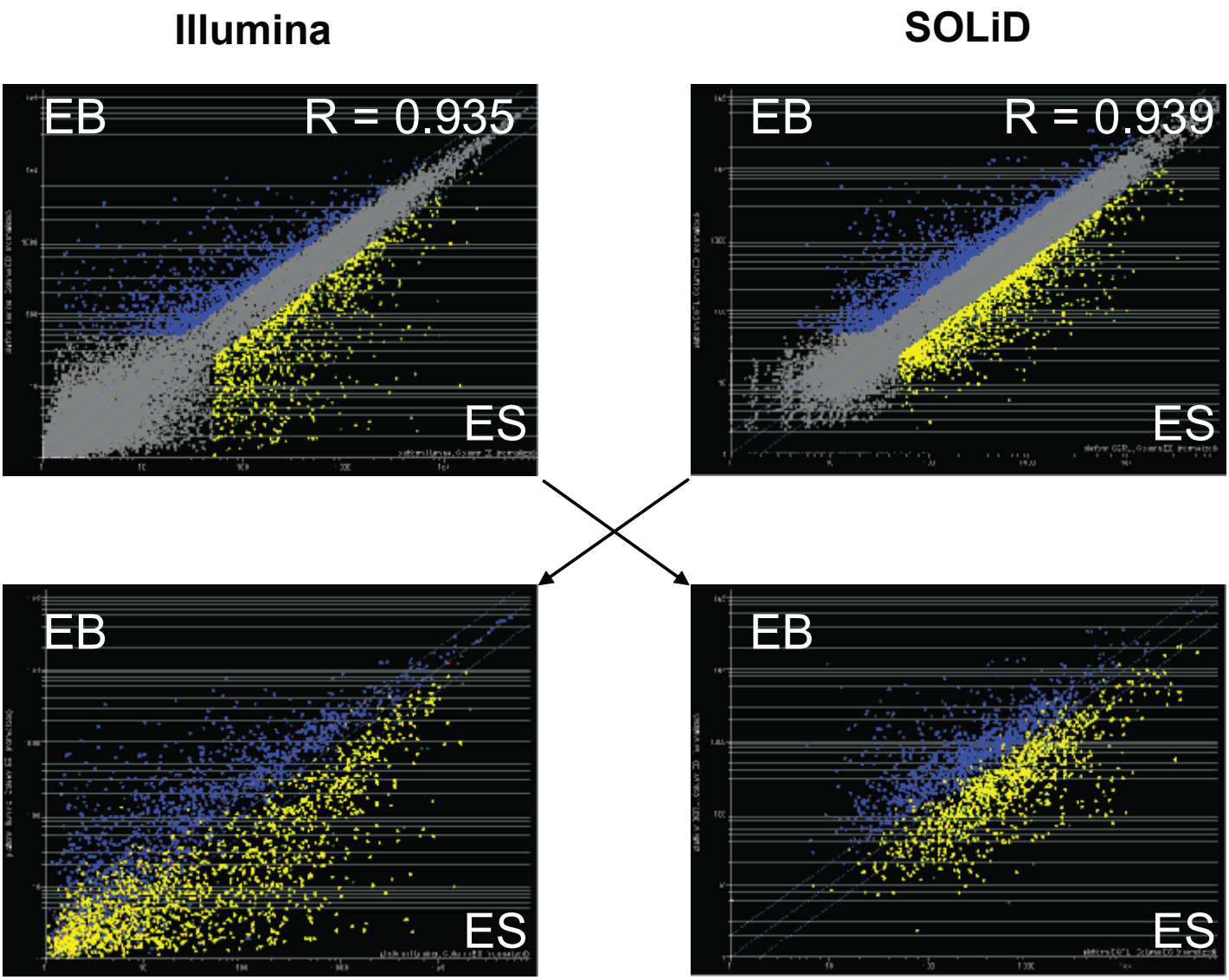


Supplementary Figure 6b.



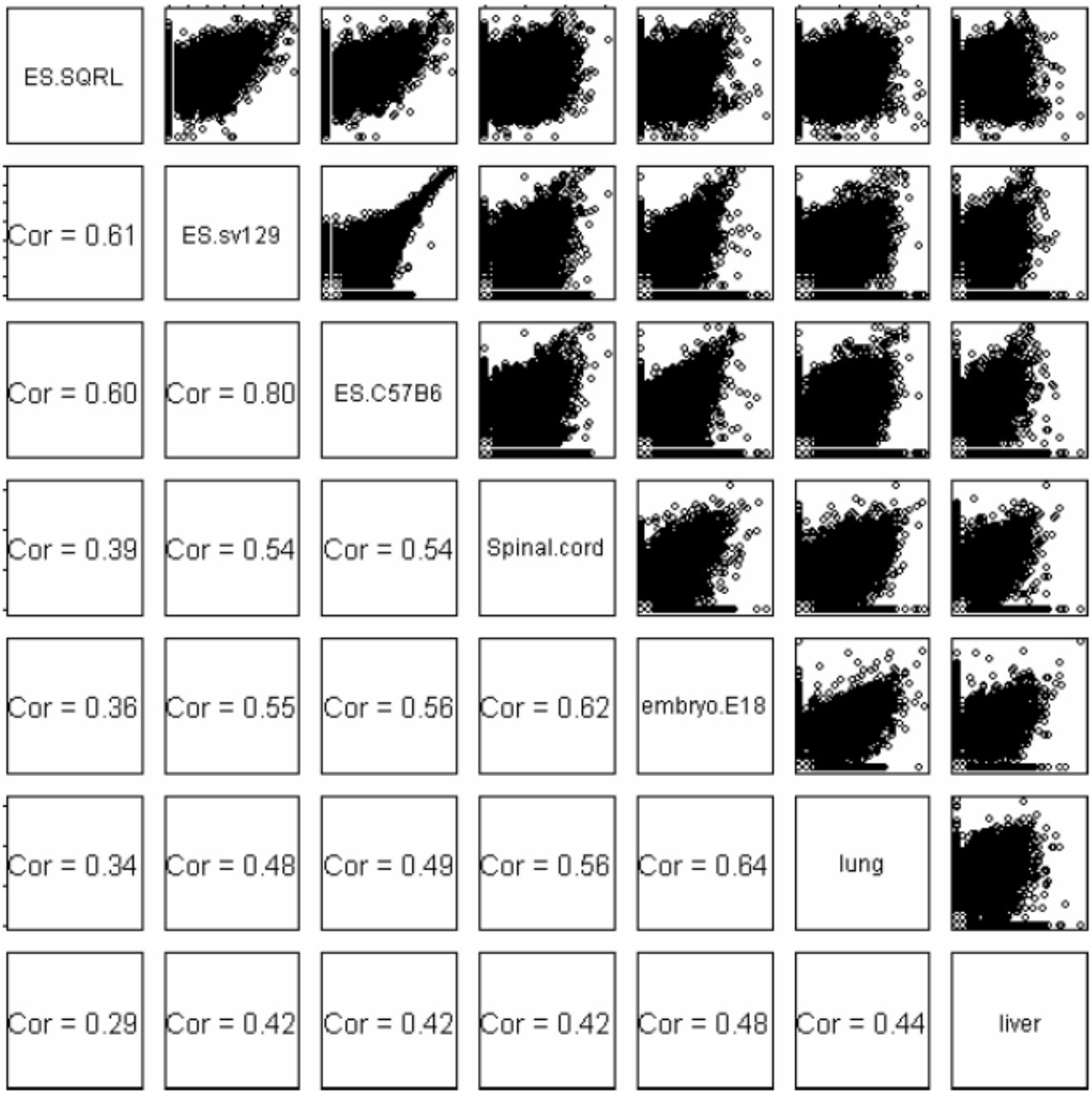
Supplementary Figure 6. High level of correlation for SQRL libraries. Correlation pairs plot showing relationship between Refseq transcripts (19,005 genes) compared by exons (a) and introns (b). Each scatter plot above the diagonal shows the relationship between individual libraries (as named on the diagonal). The Pearson correlation for each library comparison is shown below the diagonal. The X and Y axis are normalised $\log_2$ (tag count).

Supplementary Figure 7.



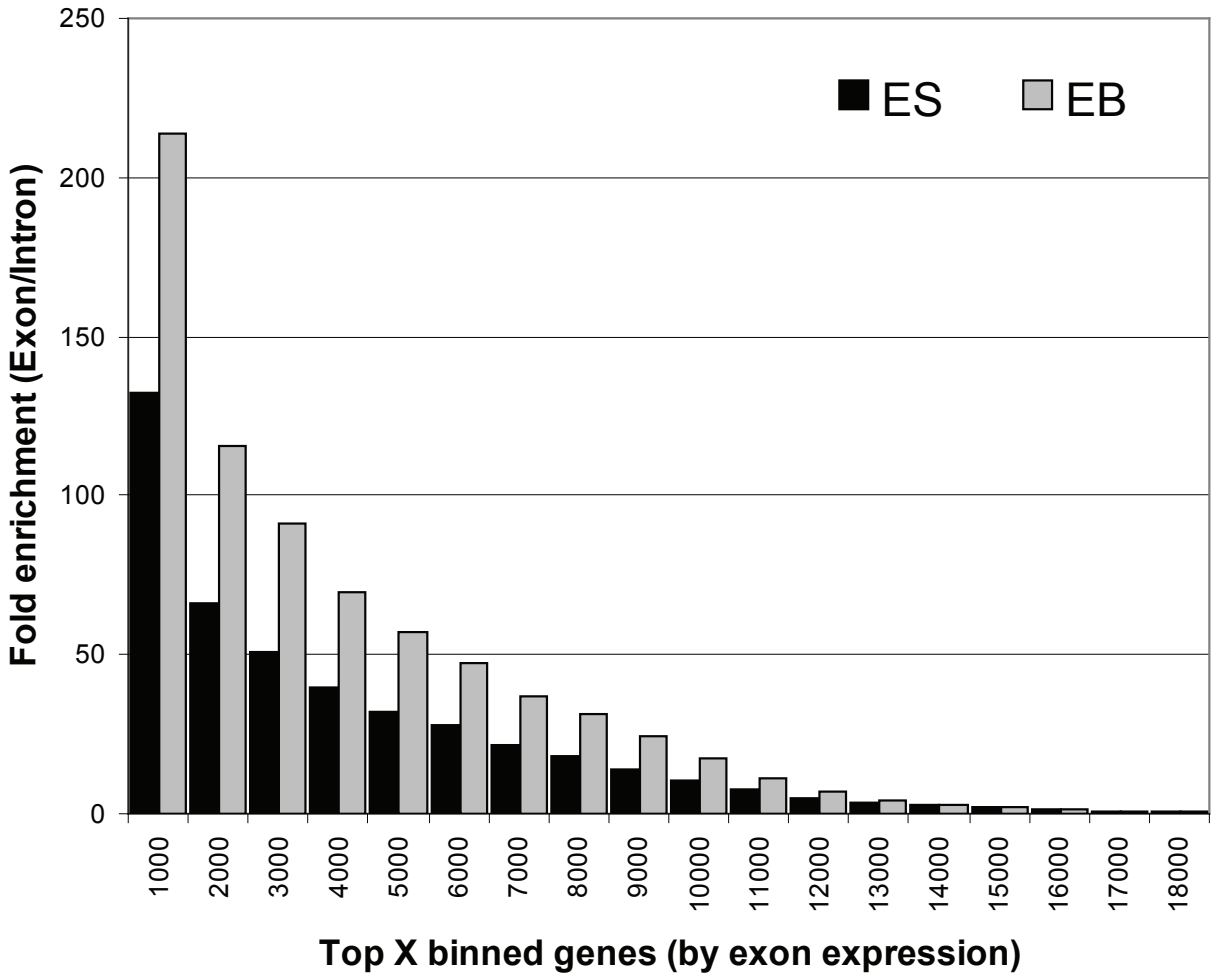
Supplementary Figure 7. Strong concordance between SQRL-SOLiD and Illumina in the detection of differential expression. Top panel: Differentially expressed genes were coloured on the correlation scatter plot for Illumina (blue: gene present or marginal in ESC or EB and > 2 fold up-regulated in EB; yellow: gene present or marginal and > 2 fold up-regulated in ES) and SQRL-SOLiD (blue: gene > 50 counts [normalised] and > 2 fold up-regulated in EB; yellow: >50 counts and > 2 fold up-regulated in ES). Pearson correlation for each comparison is shown (17,325 genes total). Lower panel. Genes > 2 fold in Illumina or SQRL-SOLiD experiment overlayed onto the converse scatter plot. Trend of differential expression is consistent between Illumina and SOLiD, except for low expressing genes. Note that many genes above detection and up- or down-regulated by SQRL-SOLiD are below detection level on the Illumina array.

Supplementary Figure 8.



Supplementary Figure 8. Comparison to the MPSS files line from the mouse transcriptome project. To determine correlation with existing data, SOLiD sequencing of ESC was compared to MPSS data (mouse transcriptome project, http://www.ncbi.nlm.nih.gov/geo/info/mouse-trans.html). Tag signature sequences from MPSS (GEO:GPL1010) were matched to all Refseq sequences using Vmatch. Those which matched to the unique Refseq set (19,005) were filtered. Any tags that matched to more than 1 unique Refseq as an ambiguous MPSS were excluded. Doing this, there was MPSS evidence for 14,579 genes from the 19,005 gene starting set. Note, the major MPSS tag for a given gene may be excluded if this tag is ambiguous. For each MPSS sample in the mouse transcriptome project, signal was determined by the raw number of counts for a given gene. For the 14,579 gene set, the correlation of these genes between the SOLiD ESC sample and each MPSS sample was determined by using Pearson correlation on the raw data signal, using the Cor function in R. Most striking is the correlation between SOLiD ESC and the MPSS SV129 ESC sample (0.65), and to C57BL6 (0.63) and correlation to other samples is lower than 0.4, showing good specificity for “stemness”.

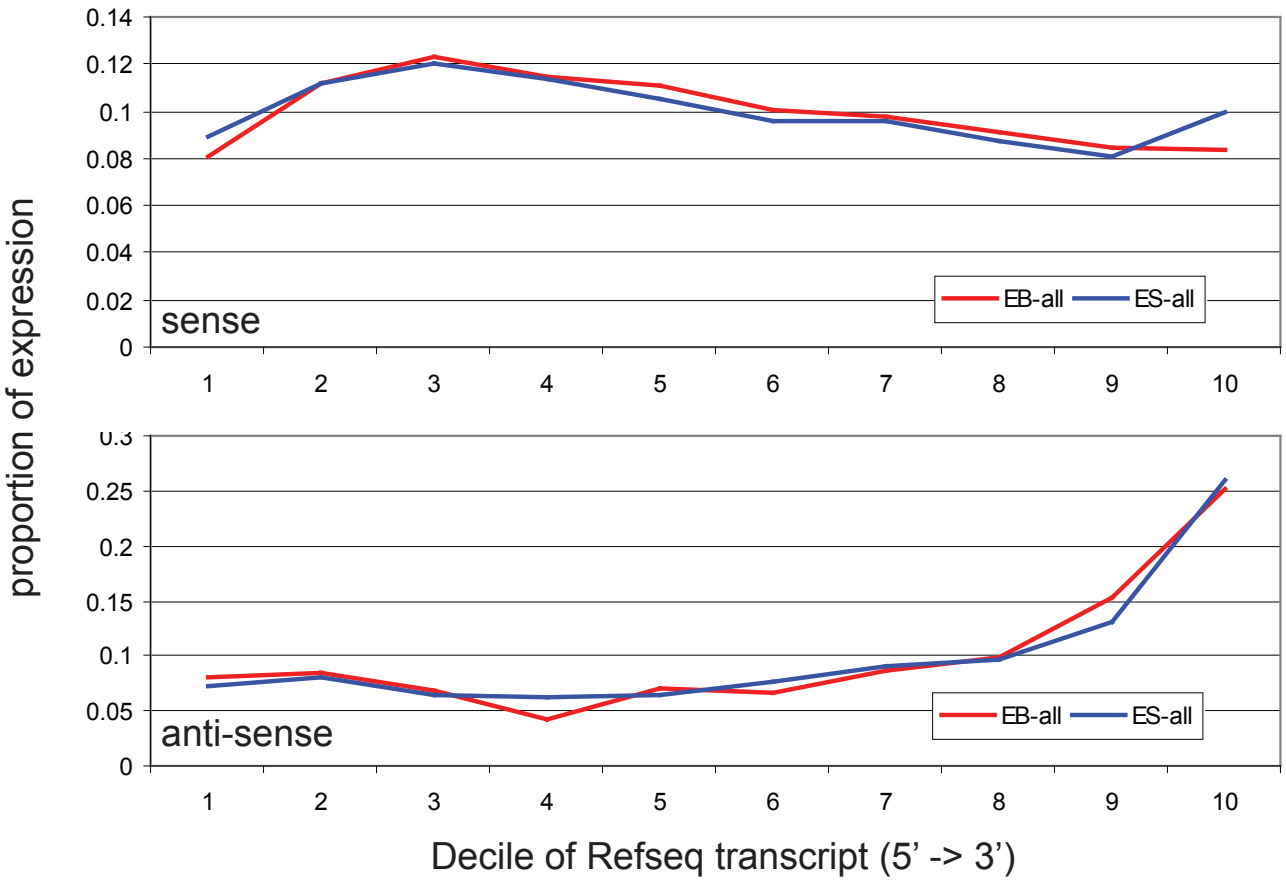
Supplementary Figure 9.



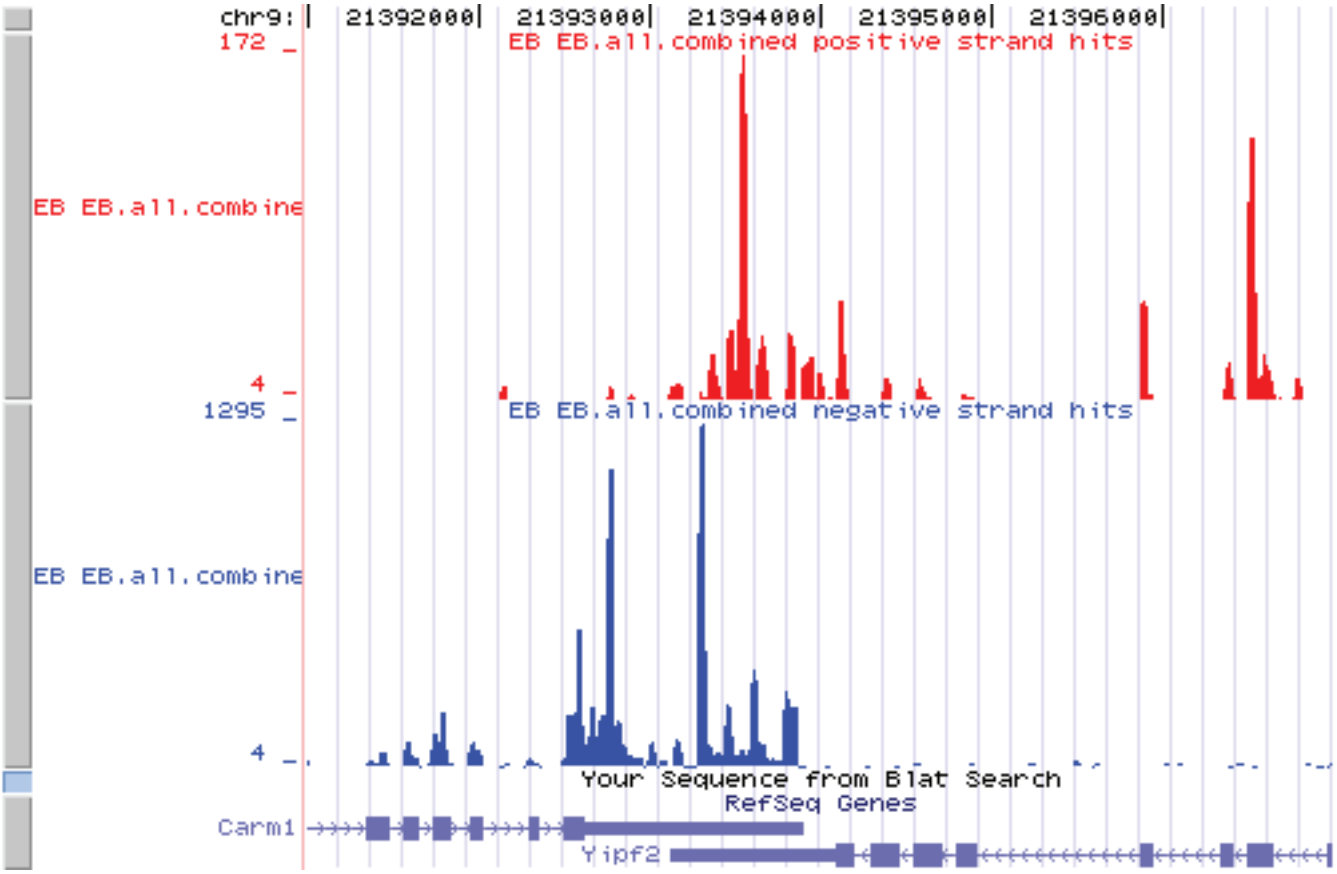
Supplementary Figure 9. Enrichment of exon versus intron expression for detected Refseq genes. To determine the extent of intron-retention evens or genomic DNA contamination in our sequencing results, we computed the difference between expression of exonic sequence versus the surrounding intronic sequence. We would expect that the majority of expression is within known exons, and we therefore an enrichment of exon expression. Representative Refseq genes with tag counts of 50 or more were selected for this analysis (11,669 genes for ESC and 11,621 genes for EB). Expression for exons was determined by summing all tags matching to a Refseq and dividing this number by the length of the transcript. This gives a per-base score (i.e. the average number of tags starting at each nt within the transcript). The same score is calculated for the tags matching to introns within a Refseq gene (total tags within introns / sum of intron lengths). For each gene, the enrichment score is: exon score/intron score. The mean enrichment for ESC was 48.7, and for EB was 62.1. To examine the proportion of enrichment in genes with differing levels of expression, the results were ranked, and then binned according to their expression.

Supplementary Figure 10.

a

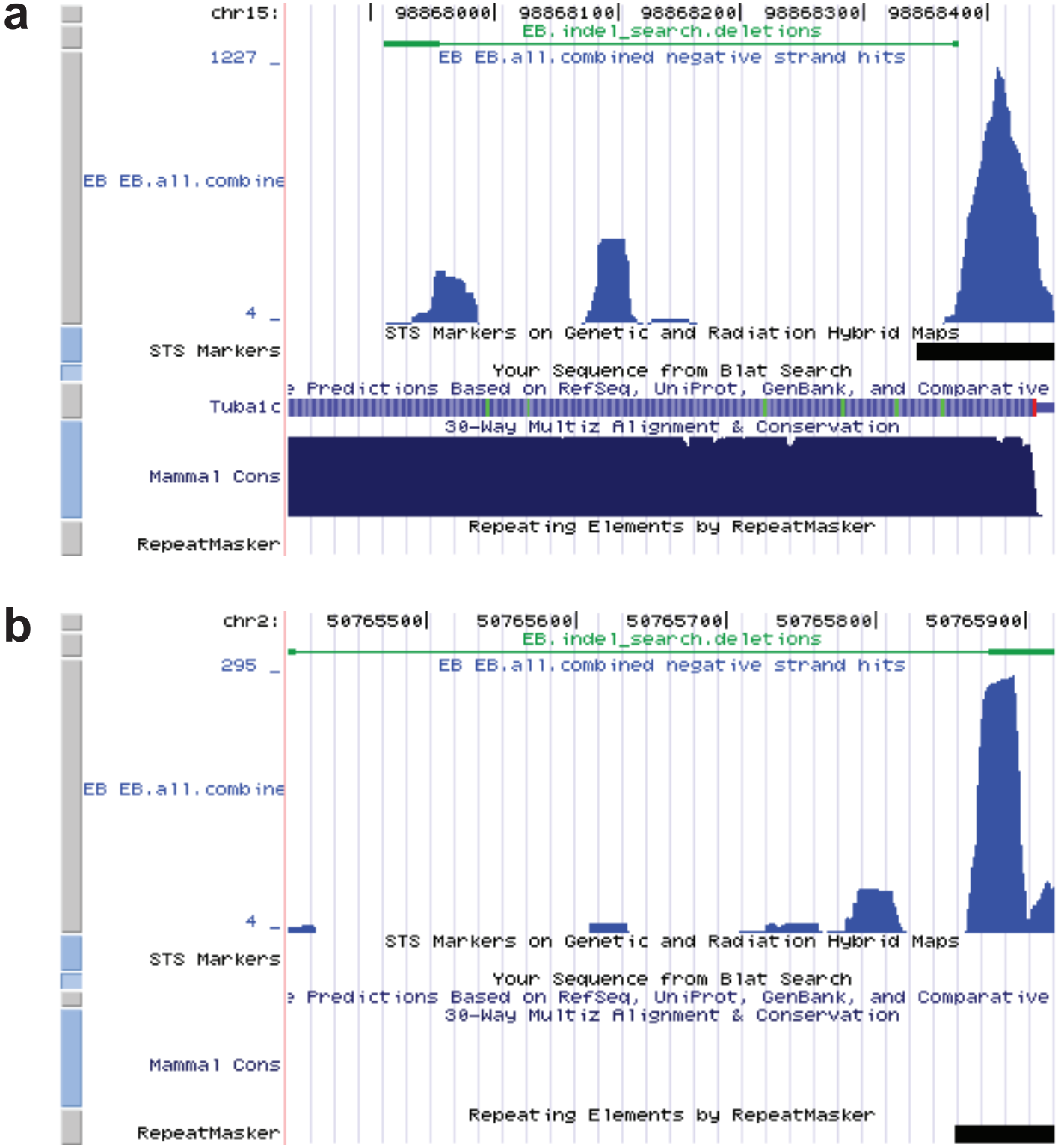


b



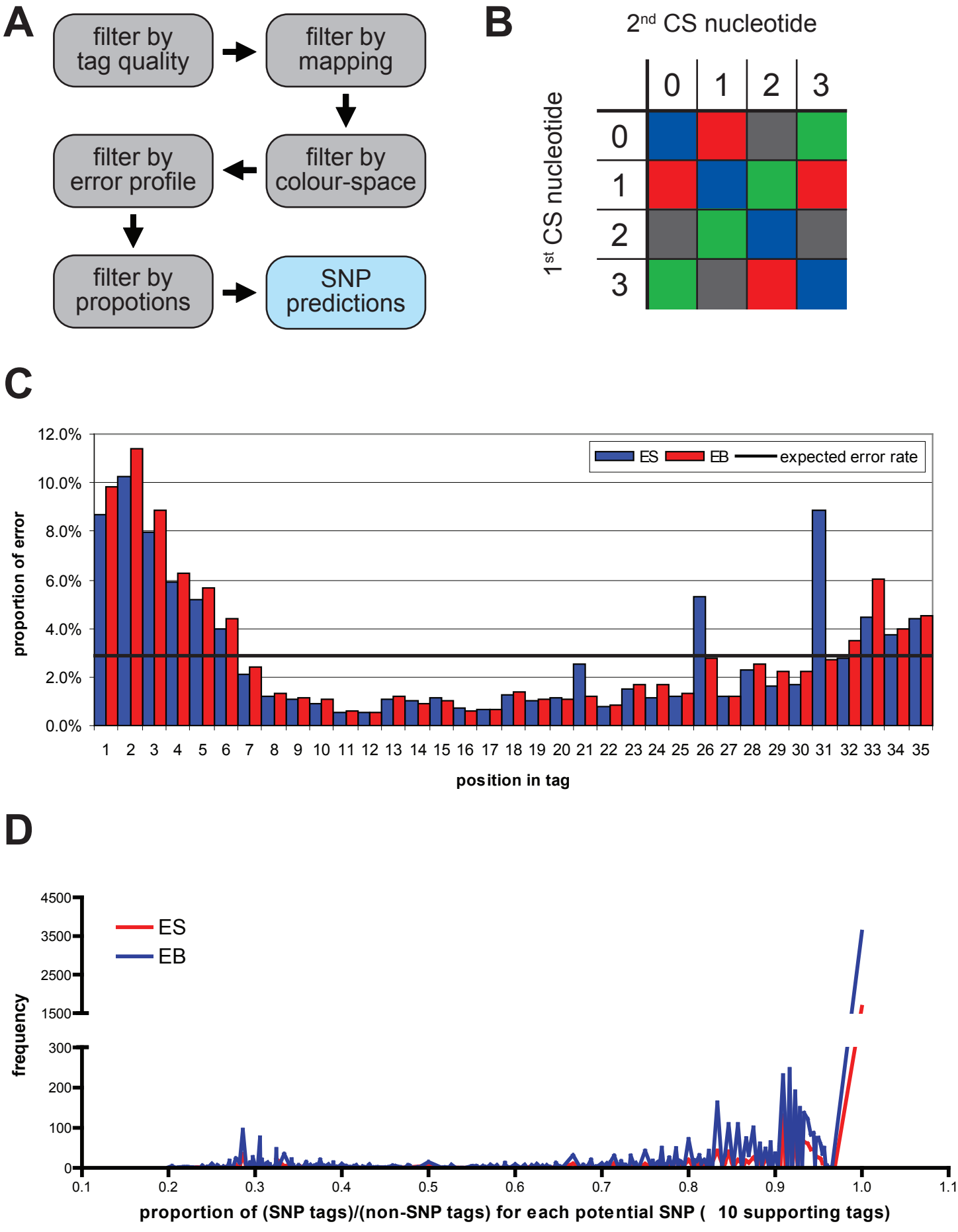
Supplementary Figure 10. Bias in tags across the length of transcripts. (A) Graph showing the percentage of filtered tags mapping to each decile of Refseq transcripts for sense strand (top panel) and the antisense strand (bottom panel). All tags were filtered to ≥5x coverage, and anti-sense tags were additionally filtered proportionally for those ≥1% of the corresponding sense-expression (based on directionality of library described in supplementary Table S16). The top panel reveals no significant transcript bias (either 5' or 3') in the library synthesis, and demonstrates that this filtering does not change the bias from unfiltered data presented in Supplementary Figure S4. The bottom panel reveals that there is strong enrichment of antisense transcripts that map to the 3' ends of sense transcripts. Further analysis shows that these are likely due to a strong enrichment of convergent (tail-tail) natural sense-antisense transcripts in the polyA+ fraction of ESC and EB cells. (B) Example of a tail-tail natural sense-antisense transcript pair that is detected by SQRL-SOLiD, showing clear and specific expression in both strands where these two genes overlap.

Supplementary Figure 11.



Supplementary Figure 11. Identification of putative alternative splicing in SQRL-SOLiD data. Tags that previously did not match to the genome or junction library were clustered by VCAKE and clusters (>50nt) were mapped to the genome. Shown here are two examples of such clusters on the UCSC genome browser: (a) novel splicing within a known gene, Tuba1c; and (b) splicing within a novel gene. The cluster is represented by the green line, where the thick line on either end reveals the regions that the cluster maps to and the thin line is the region between these. Mapping of tags to the genome is shown in blue for completeness.

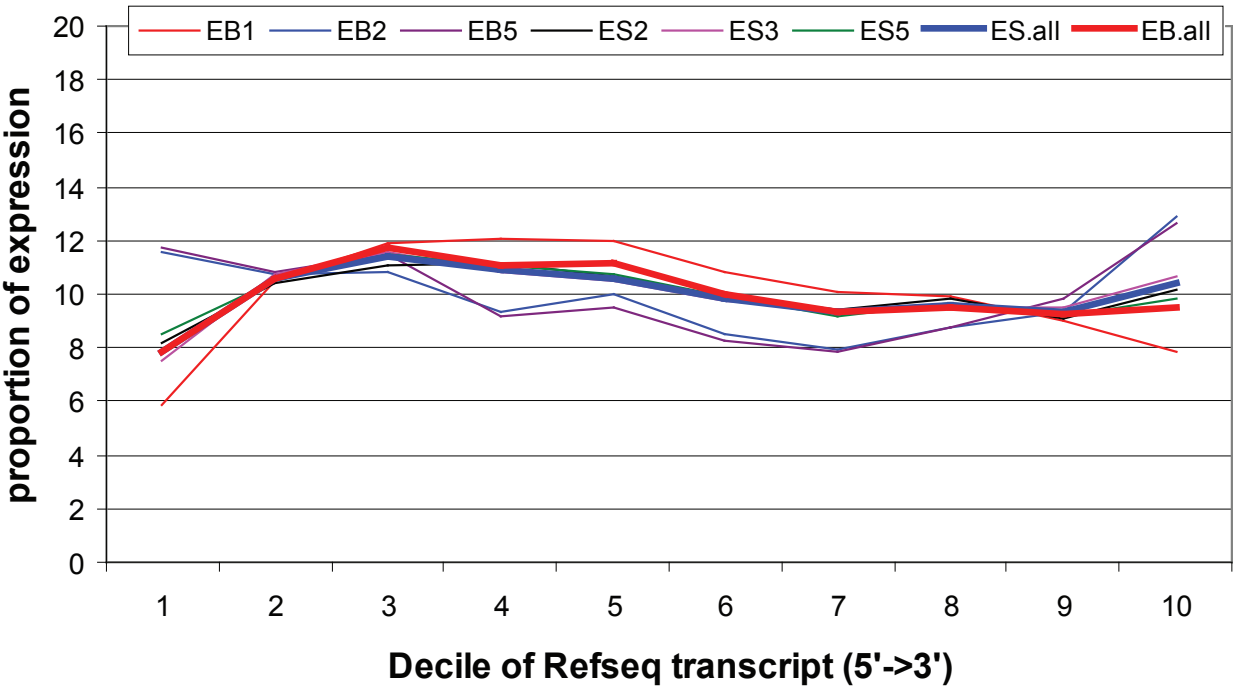
Supplementary Figure 12.



Supplementary Figure 12. SNP prediction using SQRL tags. (A) Extensive filtering is applied to the data before final SNP predictions are made. Firstly, tags must be full length (35nt) and must have passed QC (no more than 4 colour-space basecalls with Phred scores <10). Next, the data is filtered by mapping quality. No masks should be applied to the mapping, and no rescued multi-mapping tags are considered. The data is then filtered to by colour-space errors, which must be both adjacent and valid (see B). The data is then filtered by the error profile of the tags, only considering polymorphisms occurring in the reliable positions (see C). Finally, each polymorphism is filtered by both the number of unique independent tags (tags with independent start sites) supporting the potential SNP (a minimum 10 unique, independent tags are required), and by the ratio of tags that support the potential SNP to the number of tags that do not (see D). (B) Filtering by colour-space errors. SOLiD sequencing measures the relationships between nucleotides (referred to as colour-space), rather than the nucleotides directly (base-space). Because any individual nucleotide has two relationships, one to the preceding nucleotide, and one to the subsequent nucleotide, every base-space nucleotide position is measured twice. This means that for any nucleotide change in base-space, an accurate colour-space read of this requires two adjacent changes. Additionally, because there are only three base-space changes that can be made for any given nucleotide, and there are twelve possible colour-space changes, colour-space errors must be filtered by validity. Only changes within the same validity group maintain the integrity of the subsequent colour-space relationships when decoding from colour-space to base-space. For example, the only valid changes for a 00 colour-space di-nucleotide are 11, 22, and 33. All other adjacent colour-space errors are discarded. (C) Potential polymorphisms are discarded if they fall within an unreliable

position in the tag. Typically this includes the last 5nt of the tag, due to poorer sequencing performance and the presence of shorter inserts, and any individual positions that have performed poorly. For SQRL tags, this also typically includes the first 6nt, due to error introduced by priming with the random hexamer. The error profiles for the ESC and EB libraries are shown. Note particularly the increased error in the ESC library for position 26 which results in a reduction of the number of predicted SNPs from this data. (D) Potential SNPs are discarded if the proportion of tags supporting a SNP is lower than 0.75X of the sum of tags supporting a polymorphism and tags that support no change. This is the smallest ratio possible before the SNP would be predicted to be heterozygous. As we would expect genuine SNPs from a cell line derived from an inbred mouse to be predominantly homozygous, SNPs below this ratio were discarded. Heterozygous SNPs or potential RNA editing events are likely to have a lower ratio. This graph shows the frequency of this ratio for the ESC and EB libraries, showing that the majority of potential SNPs are found above this cut-off.

Supplementary Figure 13.



Supplementary Figure 13. The proportion of tags mapping across the deciles of Refseq transcript length, showing equitable distribution across the length of the transcript. Importantly, we do not observe an enrichment of extreme 5' ends, showing that the template-switch protocol does not favour capped RNA fragments.

Supplementary Table 1. The number and volume of sequences comprising the individual and composite libraries. Note that sequences mapping more than 100 times to the genome are not included in the mapping statistics, and only tags that map once at their highest frequency are included in the junction mapping statistics.

Sample	Library	Number (volume) good quality tags	Number (Volume) tags mapping genome	Number (Volume) tags mapping junctions
ES	ES-2	32233538 (1.00 Gb)	17188745 (0.52 Gb)	467194 (0.02 Gb)
	ES-3	76779322 (2.45 Gb)	56799108 (1.75 Gb)	1370413 (0.05 Gb)
	ES-5	46597803 (1.46 Gb)	22314735 (0.66 Gb)	603597 (0.02Gb)
	ES all	155610663 (4.91 Gb)	96334528 (2.93 Gb)	2441204 (0.08 Gb)
EB	EB-1	95893107 (2.98 Gb)	57876542 (1.76 Gb)	2407871 (0.08 Gb)
	EB-2	12783982 (0.39 Gb)	6807410 (0.20 Gb)	205224 (0.01 Gb)
	EB-5	51007013 (1.60 Gb)	26466027 (0.80 Gb)	767657 (0.02 Gb)
	EB all	159684102 (4.98 Gb)	91202673 (2.77 Gb)	3380752 (0.11 Gb)

Supplementary Table 2. Anti-sense junction counts as a metric for directionality. Sense strand junctions identified by more than two tags were examined for anti-sense transcripts. The table describes the number of tags hitting sense-strand junctions, the number of tags hitting anti-sense strand junctions, and the proportion of tags mapping to anti-sense junctions.

	Sense Junctions	Anti-sense Junctions	Proportion of Anti-sense Junctions
ES-2	431893	390	0.09%
ES-3	1335801	1824	0.14%
ES-5	565839	726	0.13%
ES all	2394032	4806	0.20%
EB-1	2360877	2859	0.12%
EB-2	181081	537	0.30%
EB-5	731486	4840	0.66%
EB all	3329019	13009	0.39%

Supplementary Tables 3. URLs of wiggle plots and BED tracks to visualize SQRL data of ES and EB cell state transcriptomes.

Description	URL
Informative Junctions	http://grimmond.imb.uq.edu.au/files/nicole/mm9/general/mm9_60mer_informative_junctions.BED.bz2
Diagnostic exons	http://grimmond.imb.uq.edu.au/files/nicole/mm9/general/aceview.mm9.ucsc.diagnostic.exons.gz
Diagnostic junctions	http://grimmond.imb.uq.edu.au/files/nicole/mm9/general/aceview.mm9.ucsc.only.diagnostic.junctions.bed.gz
EB junction BED track	http://grimmond.imb.uq.edu.au/files/nicole/mm9/ESEB/EB.all.sense.junctions.track.gz
ES junction BED track	http://grimmond.imb.uq.edu.au/files/nicole/mm9/ESEB/ES.all.sense.junctions.track.gz
EB Wiggle 4+ tags positive strand	http://grimmond.imb.uq.edu.au/files/nicole/mm9/ESEB/EB.all.combined.positive.wig.rounded.condensed.4plus.gz
EB Wiggle 4+ tags negative strand	http://grimmond.imb.uq.edu.au/files/nicole/mm9/ESEB/EB.all.combined.negative.wig.rounded.condensed.4plus.gz
ES Wiggle 4+ tags positive strand	http://grimmond.imb.uq.edu.au/files/nicole/mm9/ESEB/ES.all.combined.positive.wig.rounded.condensed.4plus.gz
ES Wiggle 4+ tags negative strand	http://grimmond.imb.uq.edu.au/files/nicole/mm9/ESEB/ES.all.combined.negative.wig.rounded.condensed.4plus.gz

Supplementary Table 4. Number and proportion Refseq loci identified by SQRL. Statistics are shown as the number of tags which match to the sense strand of Refseq exons at multiple tag thresholds.

Tags	ES combined number	EB combined number	ES combined proportion (%)	EB combined proportion(%)
1 to 9	2436	2732	13.21%	14.93%
10 to 99	4800	4207	26.02%	22.99%
100 to 999	5736	4828	31.09%	26.38%
1000 to 9999	5077	5895	27.52%	32.21%
10000 to 99999	391	626	2.12%	3.42%
100000 +	7	8	0.04%	0.04%

Supplementary Table 5. Statistically significant differentially expressed genes. Significance was determined by the empirical Bayes method. Lists are ranked by B statistic.

Table 5a. Genes up-regulated in the ESC library.

Refseq UCSC_ID	Bstat	Fold change	Gene symbol	Refseq ID
577	10.26	21.43	Tcfcp2l1	NM_023755
5456	9.77	41.83	Tdh	NM_021480
10018	9.54	13.69	Gm1967	NM_001033452
15887	8.95	12.79	EG435970	NM_001034893
985	8.33	19.32	Ifi202b	NM_008327
10644	8.18	7.66	Mybl2	NM_008652
6807	8.07	22.29	Dppa2	NM_028615
21838	8.01	15.36	Fgf4	NM_010202
319	7.92	8.96	Mreg	NM_001005423
15897	7.83	7.27	4933405K07Rik	NM_028913
20848	7.82	20.54	Krt42	NM_212483
726	7.46	12.93	Nr5a2	NM_030676
5367	7.44	10.89	Tgm1	NM_019984
2133	7.42	14.45	Cobl	NM_172496
12114	7.38	23.83	Klf4	NM_010637
14257	7.34	7.26	Cav1	NM_007616
11017	7.26	8.65	Phf17	NM_172303
1041	7.23	7.13	Lefty2	NM_177099
1597	7.22	6.78	6330514A18Rik	NM_183152
13160	7.21	10.56	Cnpy1	NM_175651
12236	7.11	6.54	Bnc2	NM_172870
1043	7.09	23.63	Lefty1	NM_010094
15996	7.08	12.92	Hsd17b14	NM_025330
7789	7.00	12.23	1700061G19Rik	NM_030141
21186	6.86	8.41	E130014J05Rik	NM_001040400
17531	6.85	10.96	Clgn	NM_009904
16769	6.80	29.50	Mylpf	NM_016754
7565	6.79	7.14	Zfp57	NM_001013745
4267	6.65	5.80	Aoah	NM_012054
15544	6.58	10.16	Zfp296	NM_022409
5039	6.58	10.55	Myst4	NM_017479
10521	6.53	6.91	Hck	NM_010407
8473	6.49	4.16	Actn3	NM_013456
11146	6.42	5.64	Trim2	NM_030706
5533	6.41	5.85	Epb4.9	NM_013514
1055	6.40	5.19	Enah	NM_008680
20715	6.34	4.99	Fer1l3	NM_001099634
5603	6.30	6.79	Klf5	NM_009769
20363	6.29	16.09	Jam2	NM_023844
8216	6.28	7.90	Gm949	NM_001033446
18538	6.28	5.09	Rpp25	NM_133982

20206	6.25	7.45	Centa1	NM_172723
5646	6.20	4.74	Slc15a1	NM_053079
12233	6.13	5.48	1810054D07Rik	NM_027238
2120	6.08	5.98	Tns3	NM_001083587
11652	6.05	5.14	F3	NM_010171
2111	6.04	3.82	Myo1g	NM_178440
7787	5.99	5.28	Rfx2	NM_009056
8014	5.96	9.66	Rnf125	NM_026301
1897	5.95	5.24	Gli1	NM_010296
21050	5.81	3.53	Atp11a	NM_015804
18845	5.80	6.87	Trf	NM_133977
4500	5.80	4.01	Edn1	NM_010104
5534	5.75	7.29	Fgf17	NM_008004
4669	5.67	4.46	EG630579	NM_001039239
15405	5.66	4.04	Gm397	NM_001013765
15785	5.64	4.44	Kirrel2	NM_172898
10485	5.58	8.59	Tcf15	NM_009328
666	5.57	4.07	Mdm4	NM_008575
3122	5.53	5.09	Mllt6	NM_139311
5566	5.40	6.25	Enox1	NM_172813
15198	5.36	14.28	Emp1	NM_010128
19270	5.33	9.53	Klhl13	NM_026167
11897	5.30	6.08	Ttpa	NM_015767
2544	5.23	3.50	Specc1	NM_001029936
13979	5.23	5.81	Zp3	NM_011776
5876	5.22	4.10	Wisp1	NM_018865
1853	5.18	6.51	Irak3	NM_028679
21793	5.12	3.72	2410081M15Rik	NM_028603
20707	5.11	6.73	Dnahc8	NM_013811
15618	5.11	3.56	Atp1a3	NM_144921
5388	5.08	3.98	Gzme	NM_010373
5782	5.08	3.39	Ubr5	NM_001081359
18823	5.07	4.36	Mras	NM_008624
1960	5.04	3.35	Si	NM_021882
770	5.04	4.47	Niban	NM_022018
14187	4.99	4.03	Slc7a1	NM_007513
9630	4.98	6.08	Slc25a12	NM_172436
8314	4.98	3.51	Impa2	NM_053261
20454	4.96	3.47	Klk1	NM_010639
11429	4.93	4.50	Hist2h2bb	NM_175666
4213	4.92	4.30	Pfkp	NM_019703
4673	4.90	4.42	Fbp2	NM_007994
6101	4.90	5.12	1700029P11Rik	NM_025503
14018	4.89	5.16	Pcolce	NM_008788
370	4.82	3.21	Tuba4a	NM_009447
3010	4.81	3.35	Akap1	NM_009648
780	4.79	3.66	Nmnat2	NM_175460
20873	4.76	29.31	Calcoco2	NM_001100177

19327	4.69	5.53	Rhox10	NM_001024850
21701	4.68	3.32	Gast	NM_010257
6323	4.68	4.19	Galnt6	NM_172451
5524	4.65	4.16	Piwil2	NM_021308
13595	4.64	2.95	Aff1	NM_133919
10964	4.61	3.10	Skil	NM_001039090
2529	4.61	6.45	Aldh3a1	NM_007436
7401	4.57	5.02	Myo1f	NM_053214
15455	4.55	3.64	Obox6	NM_145710
8758	4.53	3.06	Tle4	NM_011600
6674	4.52	3.62	Bdh1	NM_175177
19125	4.50	11.13	Tcstv3	NM_153523
10193	4.46	4.54	Itpka	NM_146125
3133	4.44	4.50	Arl5c	NM_207231
4443	4.44	3.80	Serpinb1a	NM_025429
9154	4.43	3.56	Sfmbt2	NM_177386
4124	4.43	3.14	Hsp90aa1	NM_010480
4142	4.42	7.12	Ckb	NM_021273
14626	4.41	4.48	Gm1070	NM_001033462
3982	4.37	43.18	Esrrb	NM_011934
21210	4.33	48.65	Zfp42	NM_009556
7931	4.32	3.36	Msh6	NM_010830
11884	4.32	2.90	Tmem64	NM_181401
10642	4.32	3.47	Sgk2	NM_013731
4001	4.29	13.00	EG435337	NM_001013824
1925	4.28	2.97	Baz2a	NM_054078
12276	4.25	5.25	Cdkn2b	NM_007670
16764	4.24	3.09	AI467606	NM_178901
12820	4.22	2.95	Arhgef10l	NM_172415
19546	4.17	14.05	Nr0b1	NM_007430
14235	4.16	4.29	Asns	NM_012055
6157	4.14	4.22	Wnt7b	NM_009528
1398	4.14	4.14	Aifm2	NM_178058
11913	4.13	4.74	Fut9	NM_010243
12613	4.12	7.18	Dlgap3	NM_198618
758	4.12	3.90	Ptgs2	NM_011198
3917	4.11	3.28	Smoc1	NM_022316
22235	4.10	8.49	Bcl3	NM_033601
1207	4.09	3.31	Sgk	NM_011361
13903	4.08	6.07	Phkg1	NM_011079
16622	4.07	3.41	Arntl	NM_007489
20654	4.06	3.16	Lamc1	NM_010683
19007	4.01	5.96	Lrrc2	NM_028838
8178	4.01	3.21	Fgf1	NM_010197
8446	4.00	3.43	Cabp4	NM_144532
9347	3.96	2.83	Setx	NM_198033
884	3.95	3.00	Dusp27	NM_001033344
4166	3.95	3.72	Jag2	NM_010588

16061	3.93	3.47	E2f8	NM_001013368
20266	3.91	6.49	Pim1	NM_008842
916	3.90	3.75	Fcgr2b	NM_001077189
4261	3.89	3.80	Sfrp4	NM_016687
10986	3.88	2.93	Atp11b	NM_029570
10490	3.84	4.41	Trib3	NM_175093
1784	3.83	16.26	Nts	NM_024435
12688	3.83	3.05	Smpdl3b	NM_133888
15004	3.82	4.85	LOC634428	NM_001080945
18662	3.78	4.22	6430514L14Rik	NM_029784
803	3.77	2.67	Tor1aip2	NM_172843
10079	3.74	3.27	Bdnf	NM_007540
8417	3.72	7.21	Cpt1a	NM_013495
6990	3.72	2.99	Morc3	NM_001045529
7528	3.72	2.91	Dhx16	NM_026987
3071	3.69	4.79	Ngfr	NM_033217
21953	3.69	6.14	Manba	NM_027288
9137	3.68	3.47	Ccdc3	NM_028804
15626	3.68	3.15	Erf	NM_010155
9636	3.67	3.42	Itga6	NM_008397
8090	3.64	2.54	Slc23a1	NM_011397
14020	3.63	5.13	Sap25	NM_001081962
15677	3.62	7.11	Ltbp4	NM_175641
1432	3.61	5.77	Arid5b	NM_023598
3015	3.59	2.67	Trim25	NM_009546
20401	3.58	2.85	Cars	NM_013742
13479	3.57	3.42	Csn3	NM_007786
10888	3.55	5.30	Stmn2	NM_025285
7712	3.54	6.68	Tcfeb	NM_011549
12949	3.52	2.91	Errfi1	NM_133753
20833	3.52	4.86	Ddx58	NM_172689
3468	3.52	2.92	Llgl2	NM_145438
22476	3.51	2.94	Pml	NM_178087
22070	3.48	2.68	C79267	NM_183148
19609	3.47	3.43	Slc7a3	NM_007515
1219	3.47	3.98	Vnn1	NM_011704
12322	3.47	4.43	Dnajc6	NM_198412
11920	3.45	2.65	Mdn1	NM_001081392
6174	3.44	2.90	Pim3	NM_145478
13608	3.42	4.74	C230055K05Rik	NM_001039231
17820	3.42	2.77	Aars	NM_146217
2548	3.41	2.78	Ncor1	NM_011308
16335	3.40	11.94	Folr1	NM_008034
8221	3.40	2.69	Dmxl1	NM_001081371
18699	3.38	2.63	Mapk6	NM_027418
5814	3.38	3.04	Trps1	NM_032000
8065	3.35	3.02	Epb4.1l4a	NM_013512
17682	3.35	2.61	Ndrg4	NM_145602

13606	3.31	42.99	Spp1	NM_009263
20134	3.29	16.19	Sep-01	NM_017461
10353	3.27	2.52	Slc4a11	NM_001081162
3375	3.26	2.98	Pecam1	NM_008816
6968	3.26	2.54	Tmem50b	NM_030018
2726	3.26	5.89	Tekt1	NM_011569
17215	3.23	3.06	Ash2l	NM_011791
13252	3.22	3.61	Zfyve28	NM_001015039
3373	3.21	2.41	Tex2	NM_198292
17856	3.20	5.93	Cdyl2	NM_029441
11118	3.19	3.39	Serpini1	NM_009250
18794	3.18	3.80	Pcolce2	NM_029620
12966	3.18	2.92	Plekhg5	NM_001004156
3121	3.18	5.77	E130012A19Rik	NM_175332
16724	3.17	2.59	Nfatc2ip	NM_010900
3245	3.16	2.87	Hsd17b1	NM_010475
12112	3.15	2.70	Zfp462	NM_172867
20892	3.11	11.26	Zscan4d	NM_001100186
3085	3.11	3.97	Ttll6	NM_172799
17776	3.10	3.50	Sntb2	NM_009229
5818	3.10	4.94	Aard	NM_175503
21169	3.10	2.32	Brwd1	NM_145125
17433	3.09	4.03	Gdf15	NM_011819
12766	3.09	4.52	Tcea3	NM_011542
15190	3.08	7.11	Gprc5a	NM_181444
22082	3.07	7.19	Inpp5d	NM_001110193
2555	3.06	3.69	Trim16	NM_053169
2137	3.05	3.29	Egfr	NM_207655
2201	3.04	2.36	Spnb2	NM_175836
13228	3.03	2.80	Yes1	NM_009535
3427	3.01	2.40	Dnaic2	NM_001034878
20871	2.99	3.73	Atrx	NM_009530
13960	2.99	3.06	Mlxipl	NM_021455
12845	2.97	7.21	Fblim1	NM_133754
20117	2.97	7.12	Mkrn1	NM_018810
17521	2.95	9.64	Gab1	NM_021356
20493	2.95	3.00	Tmem92	NM_001034896
12223	2.94	2.77	Jmjd2c	NM_144787
3636	2.94	3.97	Ubxd4	NM_145441
16032	2.93	6.17	Saa3	NM_011315
15289	2.93	10.46	AU018091	NM_001004153
1595	2.91	3.47	Reep6	NM_139292
9602	2.88	2.52	Lass6	NM_172856
15904	2.88	4.50	Klk10	NM_133712
9547	2.87	3.99	Rif1	NM_175238
6368	2.86	4.50	Itgb7	NM_013566
12321	2.85	2.41	Ak3l1	NM_009647
10305	2.84	3.57	Mal	NM_010762

124	2.83	2.68	BC050210	NM_201365
14275	2.83	3.24	Aass	NM_013930
2532	2.83	4.54	Slc47a1	NM_026183
8860	2.81	2.42	Acta2	NM_007392
15652	2.81	3.20	Axl	NM_009465
15214	2.81	2.51	Eps8	NM_007945
8082	2.79	2.28	Egr1	NM_007913
22296	2.76	2.58	Xrcc5	NM_009533
13524	2.75	2.62	Vdp	NM_019490
17806	2.73	3.55	Calb2	NM_007586
21059	2.70	2.55	Cnnm2	NM_033569
21938	2.69	2.47	Atxn1	NM_009124
13982	2.69	2.63	Rasa4	NM_133914
17621	2.68	3.95	Chd9	NM_177224
2793	2.68	3.32	Serpinf1	NM_011340
6100	2.67	2.93	Pmm1	NM_013872
6329	2.67	2.92	Grasp	NM_019518
2229	2.67	3.57	Stk10	NM_009288
13427	2.67	3.62	Kit	NM_021099
20177	2.66	2.67	Ctcfl	NM_001081387
7568	2.65	3.64	Gabbr1	NM_019439
6969	2.64	2.88	Gart	NM_010256
13443	2.64	2.74	Rest	NM_011263
13743	2.63	22.01	Pla2g1b	NM_011107
15254	2.63	5.85	Bcat1	NM_007532
11663	2.62	2.31	Usp53	NM_133857
19416	2.62	2.24	Zic3	NM_009575
17830	2.62	2.62	Ldhd	NM_027570
6917	2.62	2.33	Gabpa	NM_008065
8801	2.62	2.35	Pip5k1b	NM_008846
13896	2.60	3.28	Mmp17	NM_011846
12784	2.60	2.31	Usp48	NM_130879
3272	2.59	2.26	Vat1	NM_012037
2802	2.57	2.39	Myo1c	NM_008659
20749	2.56	2.55	Alms1	NM_145223
14217	2.54	4.99	Col1a2	NM_007743
5152	2.54	3.95	Mmrn2	NM_153127
7711	2.53	3.80	Pgc	NM_025973
8921	2.53	2.32	Zfp518	NM_028319
12461	2.52	3.12	Plk3	NM_013807
2234	2.52	3.08	Ranbp17	NM_023146
12012	2.52	2.36	Rusc2	NM_001037709
2737	2.51	3.27	Camkk1	NM_018883
5733	2.49	2.66	Myo10	NM_019472
20522	2.49	2.54	Lnpep	NM_172827
5717	2.48	2.53	Npr3	NM_001039181
4461	2.48	3.92	Tubb2b	NM_023716
13784	2.48	2.36	Oas2	NM_145227

15256	2.47	2.41	Lrmp	NM_008511
1467	2.46	3.48	Susd2	NM_027890
17598	2.46	2.65	Gpt2	NM_173866
19144	2.44	2.69	B020031M17Rik	NM_001033769
843	2.43	2.23	Bat2d	NM_001081290
638	2.42	2.36	Dyrk3	NM_145508
18195	2.42	2.91	Zbtb44	NM_172765
11711	2.39	2.89	Npnt	NM_033525
18805	2.39	3.72	Zbtb38	NM_175537
10958	2.38	5.33	Lrrc34	NM_027941
4837	2.38	2.58	Hexb	NM_010422
6673	2.37	5.15	Apod	NM_007470
22009	2.34	2.87	Wdtc1	NM_199306
5466	2.33	2.31	Kif13b	NM_001081177
3685	2.33	3.09	Grhl1	NM_145890
15784	2.33	4.28	Aplp1	NM_007467
5165	2.32	2.54	Gcap14	NM_027045
11076	2.31	2.14	Mbnl1	NM_020007
12305	2.30	8.49	L1td1	NM_001081202
18360	2.29	3.13	Pou2f3	NM_011139
10554	2.25	2.19	Cbfa2t2	NM_172860
19608	2.25	3.91	Tex11	NM_031384
18603	2.24	2.93	Lctl	NM_145835
7095	2.22	2.08	Mllt4	NM_010806
7742	2.18	4.96	Pcaf	NM_020005
13879	2.17	2.21	Ccdc92	NM_144819
5340	2.14	2.40	Zfhx2	NM_001039198
7089	2.14	2.17	Sod2	NM_013671
22367	2.14	2.81	Tjp2	NM_011597
15599	2.13	3.05	Cadm4	NM_153112
10350	2.13	2.51	Prosapip1	NM_197945
21093	2.13	2.96	Dclk2	NM_027539
5149	2.12	2.57	Glud1	NM_008133
2246	2.11	2.31	Wwc1	NM_170779
11862	2.10	2.32	Plekhf2	NM_175175
12106	2.09	2.99	AI427809	NM_001033454
19025	2.08	2.39	Glb1	NM_009752
341	2.08	4.77	Slc11a1	NM_013612
22378	2.06	3.29	Snta1	NM_009228
15423	2.06	2.15	Zfp110	NM_022981
16191	2.05	5.21	Slc28a1	NM_001004184
4511	2.05	2.80	Jarid2	NM_021878
17582	2.04	2.54	Hook2	NM_133255
10963	2.03	2.11	Prkci	NM_008857
19969	2.03	2.59	Mep1b	NM_008586
22464	2.02	2.36	Mpzl2	NM_007962
14998	2.02	2.66	Mfap5	NM_015776
15081	2.01	2.12	9630033F20Rik	NM_177003

15997	2.00	2.32	Bcat2	NM_009737
10173	2.00	2.30	Casc5	NM_029617
14599	1.98	2.43	Smarcad1	NM_007958
17520	1.97	2.31	Smarca5	NM_053124
15386	1.95	3.31	Aurkc	NM_001080965
21955	1.95	2.93	Ppef2	NM_011148
10073	1.95	2.22	2700007P21Rik	NM_173750
8429	1.95	3.38	Aldh3b1	NM_026316
11396	1.94	2.78	Mcl1	NM_008562
18958	1.94	2.07	Col7a1	NM_007738
13636	1.94	3.24	Mtf2	NM_013827
18346	1.93	2.57	Gramd1b	NM_172768
20710	1.93	2.69	Rnf39	NM_001099632
17954	1.91	2.35	Taf5l	NM_133966
7643	1.91	3.56	Enpp4	NM_199016
13111	1.90	2.23	AB112350	NM_178728
15251	1.90	2.68	St8sia1	NM_011374
9538	1.90	2.05	Kif5c	NM_008449
17220	1.90	2.22	Fnta	NM_008033
2586	1.88	4.33	Ntn1	NM_008744
3224	1.88	2.39	Klhl11	NM_172565
14973	1.85	2.74	EG406236	NM_001037155
2378	1.85	2.06	Hspa4	NM_008300
19769	1.85	2.33	Prps1	NM_021463
16317	1.84	2.44	Plekhb1	NM_013746
12119	1.84	2.37	Ctnnal1	NM_018761
16857	1.83	2.63	Ctbp2	NM_009980
12885	1.83	2.67	OTTMUSG00000010173	NM_001037926
6617	1.80	2.71	Liph	NM_001083894
17772	1.80	2.24	Tmco7	NM_173037
18487	1.79	2.52	Cul5	NM_027807
1658	1.79	3.67	Aes	NM_010347
7883	1.78	2.42	Cdc42ep3	NM_026514
9567	1.78	2.24	Tanc1	NM_198294
10959	1.76	4.38	4930558O21Rik	NM_026668
21198	1.75	2.33	Acacb	NM_133904
15857	1.74	7.83	2410004F06Rik	NM_001082581
3472	1.74	2.50	Itgb4	NM_001005608
15797	1.74	9.44	Zbtb32	NM_021397
2077	1.72	2.91	Nefh	NM_010904
5326	1.71	2.40	Acin1	NM_023190
5573	1.70	2.14	Tnfsf11	NM_011613
12554	1.70	2.47	Bmp8b	NM_007559
12105	1.70	4.30	Abca1	NM_013454
15457	1.70	5.75	Crxos1	NM_001033638
10491	1.70	2.66	Nrsn2	NM_001009948
12657	1.69	2.46	Tinagl	NM_023476
19488	1.69	2.15	Pdzd4	NM_001029868

15975	1.68	2.73	Cd37	NM_007645
4371	1.67	2.21	Hist1h3c	NM_175653
2145	1.64	2.07	Etaa1	NM_026576
12838	1.63	7.41	Epha2	NM_010139
17316	1.62	2.33	Acsl1	NM_007981
7171	1.62	3.11	Srrm2	NM_175229
2749	1.61	2.73	Aspa	NM_023113
569	1.61	2.19	Serpinb8	NM_011459
2864	1.60	2.26	A830091I15Rik	NM_172795
3062	1.59	5.24	Itga3	NM_013565
13446	1.57	4.42	Igfbp7	NM_008048
147	1.57	2.48	Chst10	NM_145142
7689	1.57	3.27	Gnmt	NM_010321
10154	1.56	2.67	Rasgrp1	NM_011246
14000	1.56	2.98	Vgf	NM_001039385
2888	1.56	2.34	Suz12	NM_199196
2689	1.56	2.29	Mink1	NM_001045964
14198	1.55	2.61	Brca2	NM_009765
2062	1.55	2.00	Tbc1d10a	NM_134023
113	1.55	2.59	Hs6st1	NM_015818
4094	1.54	14.57	Tcl1	NM_009337
21541	1.53	2.13	Fat1	NM_001081286
6845	1.52	2.27	Col8a1	NM_007739
17244	1.51	2.27	Rbm13	NM_026453
12975	1.50	2.55	Acot7	NM_133348
18920	1.50	2.47	Camkv	NM_145621
6370	1.49	3.47	Rarg	NM_011244
8366	1.48	2.37	Pias2	NM_008602
5150	1.48	24.46	2200001I15Rik	NM_183278
21995	1.48	12.80	Tex14	NM_031386
17294	1.47	21.66	Triml1	NM_177742
13537	1.47	2.06	Shroom3	NM_015756
16766	1.46	4.44	Spn	NM_001037810
21517	1.46	2.45	Capn3	NM_007601
15691	1.45	2.50	C030039L03Rik	NM_198417
2211	1.44	2.15	Cpeb4	NM_026252
12616	1.44	4.10	Gjb3	NM_008126
3009	1.44	2.03	Msi2	NM_054043
14808	1.43	2.30	V1rb3	NM_053226
16604	1.42	2.02	Ipo7	NM_181517
2275	1.42	2.37	Cyfip2	NM_133769
13255	1.42	2.84	Tnip2	NM_139064
4447	1.41	3.70	Serpinb9b	NM_011452
7322	1.41	2.06	Fkbp5	NM_010220
6509	1.40	2.12	Dnm1l	NM_152816
2894	1.39	1.98	Rhot1	NM_021536
18822	1.38	2.43	D9Ertd280e	NM_177775
455	1.37	3.37	Efhd1	NM_028889

14972	1.37	2.40	Wnk1	NM_198703
1299	1.36	2.30	A530089I17Rik	NM_133999
4838	1.36	2.14	Enc1	NM_007930
11963	1.35	8.54	Aqp3	NM_016689
21717	1.34	2.48	Gmpr	NM_025508
12070	1.33	2.45	Anp32b	NM_130889
5118	1.33	2.10	Ankrd28	NM_001024604
13664	1.33	3.11	Gm1679	NM_001033459
16258	1.32	2.48	Rab30	NM_029494
12201	1.32	2.19	Tnc	NM_011607
16736	1.32	12.50	Nupr1	NM_019738
22457	1.30	2.07	Ankrd25	NM_145611
4580	1.30	3.35	Cltb	NM_028870
13711	1.30	2.44	Foxn4	NM_148935
16539	1.29	16.72	Rbmxl2	NM_029660
20307	1.27	2.48	C77370	NM_001077354
2836	1.26	2.40	Nufip2	NM_001024205
11556	1.26	1.96	Rbm15	NM_001045807
18146	1.26	2.76	Rab3d	NM_031874
19858	1.24	2.68	Mtap7d2	NM_001081124
21488	1.22	2.51	C230081A13Rik	NM_172924
3115	1.21	2.17	Npepps	NM_008942
1271	1.21	2.09	Dse	NM_172508
5921	1.19	5.44	Mafa	NM_194350
21192	1.17	7.91	OTTMUSG00000011070	NM_001103158
2633	1.17	2.39	Senp3	NM_030702
17853	1.17	2.26	Maf	NM_001025577
3100	1.16	2.16	Nfe2l1	NM_008686
9169	1.15	3.35	St8sia6	NM_145838
12306	1.14	3.36	Ankrd38	NM_172872
19654	1.13	2.04	Taf9b	NM_001001176
8794	1.12	3.09	Klf9	NM_010638
8013	1.10	2.23	D030074E01Rik	NM_029491
175	1.10	2.76	1500015O10Rik	NM_024283
2131	1.09	3.27	Ddc	NM_016672
3816	1.09	2.20	Fancm	NM_178912
202	1.09	2.09	1700019D03Rik	NM_144953
1406	1.09	2.11	Hk1	NM_010438
2973	1.08	2.07	Thrap1	NM_001080931
17542	1.07	5.35	Cd97	NM_011925
12107	1.06	2.53	Slc44a1	NM_133891
20387	1.06	2.21	Nfib	NM_008687
13345	1.06	2.71	Rbpj	NM_001080928
14022	1.05	2.13	Hrbl	NM_178162
4671	1.05	3.51	6720457D02Rik	NM_175252
20693	1.04	2.81	Gzmd	NM_010372
20619	1.04	2.18	Wasf3	NM_145155
9415	1.03	2.14	Dnm1	NM_010065

8100	1.03	3.07	Pura	NM_008989
1389	1.02	4.12	Adamts14	NM_001081127
21968	1.02	11.00	Dmrt1	NM_015826
1167	1.01	3.71	Utrn	NM_011682
20877	1.01	3.32	Usp7	NM_001003918
19617	1.00	2.49	Zmym3	NM_019831
18872	0.98	2.17	Twf2	NM_011876
882	0.98	2.13	Pou2f1	NM_011137
90	0.96	1.90	Phf3	NM_001081080
17955	0.95	3.35	AK122209	NM_001029876
10285	0.95	2.64	Usp50	NM_029163
12143	0.94	2.00	Rod1	NM_144904
7654	0.94	2.92	Slc29a1	NM_022880
2907	0.93	4.11	Ccl7	NM_013654
7446	0.92	3.59	Btnl7	NM_001081663
12048	0.90	1.97	Zcchc7	NM_177027
2551	0.90	2.13	Trpv2	NM_011706
14721	0.90	2.80	Rab11fip5	NM_001003955
1973	0.89	2.34	Itga7	NM_008398
579	0.89	2.63	Inhbb	NM_008381
18113	0.89	4.26	Icam1	NM_010493
6018	0.89	2.93	Mfng	NM_008595
7676	0.88	2.87	Slc22a7	NM_144856
102	0.88	2.45	Dst	NM_134448
12907	0.88	2.42	Fbxo2	NM_176848
1169	0.88	4.57	Plagl1	NM_009538
5172	0.87	3.39	Sh2d4b	NM_177816
10200	0.86	2.24	Mga	NM_013720
13159	0.86	2.96	En2	NM_010134
9589	0.86	1.93	Scn3a	NM_018732
20194	0.86	2.11	Trim52	NM_198601
19070	0.86	2.76	Mobp	NM_001039365
21682	0.85	2.56	Col18a1	NM_009929
19635	0.85	2.44	4930519F16Rik	NM_029170
9327	0.84	3.04	Col5a1	NM_015734
11392	0.84	6.73	Hormad1	NM_026489
13900	0.82	2.10	Psph	NM_133900
8326	0.81	2.43	D18Ertd653e	NM_172631
20377	0.81	2.91	Lox	NM_010728
18820	0.80	2.51	Pik3cb	NM_029094
8187	0.80	2.01	Lars	NM_134137
2157	0.80	2.08	Vps54	NM_139061
3774	0.80	2.27	Arhgap5	NM_009706
20300	0.79	2.03	Tbc1d10b	NM_144522
17468	0.79	2.86	Jak3	NM_010589
7512	0.78	4.13	Pou5f1	NM_013633
19926	0.77	2.24	Ddx3y	NM_012008
10804	0.76	3.27	OTTMUSG00000016571	NM_001024825

12454	0.74	3.12	Hpdl	NM_146256
10619	0.73	2.63	Tgm2	NM_009373
9051	0.72	2.07	Dusp5	NM_001085390
20576	0.72	2.64	Syt9	NM_021889
10084	0.71	2.01	Lgr4	NM_172671
16773	0.71	1.92	Sephs2	NM_009266
15287	0.71	2.22	2810474O19Rik	NM_026054
6019	0.71	2.13	Card10	NM_130859
3464	0.71	2.00	Grb2	NM_008163
21944	0.70	3.35	Lima1	NM_023063
18741	0.69	2.55	4930486G11Rik	NM_175213
10463	0.69	2.11	OTTMUSG00000015743	NM_001034900
10831	0.67	1.93	Tcfl5	NM_178254
11021	0.66	2.77	Ccrn4l	NM_009834
14907	0.66	2.24	Irak2	NM_172161
17288	0.66	2.38	Mtus1	NM_001005863
7834	0.65	3.62	Lama1	NM_008480
12891	0.63	5.81	2610305D13Rik	NM_145078
3968	0.62	2.37	Mlh3	NM_175337
2074	0.62	2.09	Nf2	NM_010898
1042	0.61	2.07	Pycr2	NM_133705
12890	0.61	6.14	OTTMUSG00000010673	NM_001014397
19982	0.60	2.36	Lama5	NM_001081171
21595	0.60	2.63	D430041B17Rik	NM_172737
10777	0.60	7.15	Tcfap2c	NM_009335
2543	0.60	1.92	Akap10	NM_019921
20297	0.59	2.85	Pde8a	NM_008803
2838	0.59	1.84	Myo18a	NM_011586
16011	0.58	2.40	Sult2b1	NM_017465
10185	0.58	2.90	Chac1	NM_026929
22047	0.58	1.84	Cul4b	NM_028288
6237	0.58	1.96	Rapgef3	NM_144850
7149	0.58	2.51	2610036F08Rik	NM_029281
18144	0.57	3.99	Dock6	NM_177030
2047	0.57	2.19	Morc2a	NM_198162
19638	0.57	4.03	Tsx	NM_009440
12010	0.56	1.85	Unc13b	NM_001081413
22364	0.56	2.19	Capn1	NM_001110504
21676	0.56	1.89	Ahctf1	NM_026375
2965	0.54	1.78	Usp32	NM_001029934
4996	0.54	1.78	Slc4a7	NM_001033270
6808	0.53	20.31	Morc1	NM_010816
3239	0.53	2.18	Stat3	NM_011486
16709	0.52	2.52	Slc5a11	NM_146198
277	0.52	2.92	A430093A21Rik	NM_001081436
9570	0.52	1.96	Mar-07	NM_020575
3817	0.52	2.06	C79407	NM_172578
17357	0.52	5.10	Palld	NM_001081390

19126	0.51	2.71	LOC625360	NM_001037925
12642	0.51	3.64	Lck	NM_010693
1391	0.51	1.86	X99384	NM_013753
438	0.50	2.85	B3gnt7	NM_145222
6606	0.50	2.16	Clcn2	NM_009900
5574	0.50	2.34	Dgkh	NM_001081336
8886	0.47	2.81	Cep55	NM_028760
5318	0.46	1.89	Jub	NM_010590
2750	0.46	3.10	Spata22	NM_001045531
1269	0.46	2.25	BB146404	NM_178908
19616	0.46	3.22	Gjb1	NM_008124
4492	0.45	1.93	Pak1ip1	NM_026550
20813	0.45	2.11	Rad50	NM_009012
10068	0.45	2.65	Pax6	NM_013627
12844	0.44	2.49	Spen	NM_019763
2075	0.44	1.83	Nipsnap1	NM_008698
13042	0.43	1.88	Agrn	NM_021604
11883	0.43	2.81	Efcbp1	NM_178617
14637	0.42	2.00	Ptcd3	NM_027275
21846	0.42	2.19	Cdh3	NM_001037809
16259	0.42	2.16	4632434I11Rik	NM_001080995
16100	0.42	2.00	BB128963	NM_172742
5162	0.41	2.58	Wapal	NM_001004436
15752	0.41	2.53	Ppp1r14a	NM_026731
18733	0.40	1.94	Col12a1	NM_007730
19642	0.39	1.79	Rnf12	NM_011276
6124	0.39	2.10	Tcf20	NM_013836
7359	0.38	3.03	Ubash3a	NM_177823
13717	0.38	2.15	Trpv4	NM_022017
15001	0.37	11.02	Gdf3	NM_008108
18519	0.37	3.31	Pstpip1	NM_011193
762	0.37	2.14	Tpr	NM_133780
12913	0.36	1.93	Srm	NM_009272
4361	0.35	2.39	Hist1h3d	NM_178204
6322	0.35	2.68	Ela1	NM_033612
12711	0.35	3.92	Zdhhc18	NM_001017968
17345	0.34	2.49	Sap30	NM_021788
13618	0.34	1.90	Lrrc8d	NM_178701
1722	0.34	1.83	8030451F13Rik	NM_175418
18167	0.34	2.16	9530077C05Rik	NM_026739
14632	0.34	2.45	Jmjd1a	NM_173001
5420	0.33	33.69	D14Ertd668e	NM_199015
20673	0.33	1.99	Atad2b	NM_001099628
5014	0.33	2.33	Kcnk5	NM_021542
18188	0.32	2.11	Igsf9b	NM_001033323
5477	0.32	2.26	Esco2	NM_028039
22209	0.31	1.79	Klhl22	NM_145479
290	0.31	2.99	Fzd5	NM_001042659

1896	0.30	2.67	Arhgap9	NM_146011
1558	0.30	2.94	Fstl3	NM_031380
4885	0.30	1.80	Sgtb	NM_144838
9839	0.30	2.70	Pramel6	NM_178249
10730	0.29	3.96	Ncoa3	NM_008679
6971	0.29	1.89	Son	NM_178880
14597	0.29	2.01	Grid2	NM_008167
6009	0.29	3.17	Tst	NM_009437
15730	0.28	6.39	5830482F20Rik	NM_177158
4236	0.28	3.99	Nid1	NM_010917
13324	0.27	1.84	Lap3	NM_024434
17714	0.27	1.89	D230025D16Rik	NM_145604
17203	0.24	2.43	Tacc1	NM_177089
13964	0.24	4.07	Fkbp6	NM_033571
956	0.23	2.12	Casq1	NM_009813
4695	0.22	3.20	Nlrp4f	NM_175290
17938	0.22	2.51	Spire2	NM_172287
15635	0.21	2.24	Ceacam1	NM_001039185
808	0.21	1.84	Abl2	NM_009595
19859	0.19	3.60	A830080D01Rik	NM_001033472
1961	0.19	2.18	Dgka	NM_016811
12631	0.19	1.94	Yars	NM_134151
5284	0.18	2.34	Zfp219	NM_027248
8892	0.18	1.92	Tmem20	NM_175507
12739	0.16	2.60	Rhd	NM_011270
655	0.16	2.06	Pctk3	NM_008795
17943	0.16	9.58	Tubb3	NM_023279
1611	0.16	1.93	9030607L17Rik	NM_027829
12941	0.15	2.81	Spsb1	NM_029035
1412	0.15	1.86	Ddx21	NM_019553
12815	0.14	1.89	Aldh4a1	NM_175438
17116	0.13	1.95	Arhgef10	NM_172751
13135	0.13	1.87	Gbx1	NM_015739
781	0.12	3.36	Lamc2	NM_008485
17471	0.12	3.32	Fcho1	NM_028715
5692	0.11	1.94	Nipbl	NM_027707
18163	0.11	2.08	Zfp809	NM_172763
227	0.10	1.84	Sgol2	NM_199007
14680	0.10	1.94	Sema4f	NM_011350
11620	0.10	1.83	Col11a1	NM_007729
21993	0.10	1.83	Sparc	NM_009242
12086	0.10	1.77	Tex10	NM_172304
7159	0.10	2.52	BC038613	NM_153784
11741	0.09	1.82	Adh4	NM_011996
6376	0.09	1.99	Sp1	NM_013672
16235	0.08	2.37	Nox4	NM_015760
6282	0.07	2.04	Kcnh3	NM_010601
14188	0.07	1.78	Ubl3	NM_011908

18808	0.06	1.92	Slc25a36	NM_138756
8269	0.06	4.04	Cdx1	NM_009880
5523	0.06	2.46	Slc39a14	NM_144808
2724	0.06	1.92	Slc13a5	NM_001004148
5775	0.06	1.69	4631426E05Rik	NM_025712
9619	0.05	1.76	Zfp650	NM_001081548
1554	0.04	2.43	Hcn2	NM_008226
3137	0.04	2.15	Stac2	NM_146028
21118	0.03	2.29	Pou4f1	NM_011143
1839	0.03	2.42	Mdm2	NM_010786
11780	0.02	1.93	Ddah1	NM_026993
6892	0.02	2.30	Zfp654	NM_028059
1035	0.01	1.96	Parp1	NM_007415
19005	0.01	2.16	Als2cl	NM_146228

Table 5b. Genes up-regulated in the EB library.

Refseq UCSC ID	**Bstat**	**Fold change**	**Common name**	**Refseq ID**
16427	11.24	245.51	Hbb-bh1	NM_008219
2964	9.91	16.76	Car4	NM_007607
14226	9.47	43.17	Asb4	NM_023048
21211	9.11	28.00	Hand2	NM_010402
8843	8.77	20.70	Dkk1	NM_010051
6847	8.53	7.94	St3gal6	NM_018784
11406	8.40	42.28	Car14	NM_011797
19444	7.51	13.57	Hmgb3	NM_008253
18996	7.42	7.22	Pthr1	NM_011199
10966	7.35	7.65	Cldn11	NM_008770
2620	6.81	6.41	Tmem88	NM_025915
2957	6.69	22.21	Lhx1	NM_008498
16428	6.62	8.00	Hbb-y	NM_008221
2677	6.60	7.64	Alox15	NM_009660
11726	6.55	5.51	Slc39a8	NM_026228
21817	6.50	10.40	Evx1	NM_007966
10049	6.46	9.75	Lmo2	NM_008505
18640	6.40	5.32	Dapk2	NM_010019
2140	6.36	9.51	1500041B16Rik	NM_029861
19723	6.30	5.86	Nxf7	NM_130888
13035	6.30	5.88	C1qdc2	NM_026125
11132	6.23	20.82	Tdo2	NM_019911
12231	6.20	43.36	Cer1	NM_009887
3097	6.15	5.24	Skap1	NM_001033186
6586	6.14	6.38	Klhl6	NM_183390
3257	6.13	7.34	Ramp2	NM_019444
11243	5.83	6.17	Efna1	NM_010107
7976	5.83	4.00	AK220484	NM_001083628
713	5.80	9.94	Tnni1	NM_021467

4083	5.74	4.83	Gsc	NM_010351
1319	5.68	4.17	Cd24a	NM_009846
4939	5.58	7.33	Emb	NM_010330
17307	5.49	4.86	Pdlim3	NM_016798
1259	5.49	7.49	Hey2	NM_013904
21552	5.40	29.63	Capn6	NM_007603
5789	5.33	4.68	Cthrc1	NM_026778
19696	5.32	4.29	Tspan6	NM_019656
5827	5.31	6.03	Enpp2	NM_015744
5329	5.23	4.20	Slc7a8	NM_016972
19641	5.12	4.53	Slc16a2	NM_009197
19176	5.09	7.20	Pim2	NM_138606
5974	5.04	6.91	Foxh1	NM_007989
52	4.98	5.91	Crispld1	NM_031402
1825	4.96	6.81	Kcnmb4	NM_021452
5453	4.94	8.65	Gata4	NM_008092
10778	4.79	4.40	Bmp7	NM_007557
4937	4.78	6.78	Isl1	NM_021459
22500	4.75	7.45	Has2	NM_008216
5688	4.72	6.34	Egflam	NM_178748
5541	4.72	4.06	Rcbtb2	NM_134083
1255	4.59	8.56	Rspo3	NM_028351
11134	4.51	3.87	Gucy1b3	NM_017469
20568	4.45	6.51	Grrp1	NM_001099296
11057	4.44	2.90	Pfn2	NM_019410
19679	4.38	6.14	Zfp711	NM_177747
3	4.35	13.28	Sox17	NM_011441
15705	4.31	4.20	Dll3	NM_007866
19771	4.27	4.98	Tsc22d3	NM_001077364
518	4.26	2.92	Kif1a	NM_008440
17626	4.21	6.76	Irx3	NM_008393
2219	4.19	125.48	Hba-x	NM_010405
7428	4.18	4.16	H2-Oa	NM_008206
8885	4.16	9.51	Cyp26a1	NM_007811
19154	4.14	3.19	Dgkk	NM_177914
9600	4.12	4.92	Stk39	NM_016866
22431	4.07	4.88	Cdh5	NM_009868
18924	4.07	4.81	6230427J02Rik	NM_026597
19656	4.06	4.14	Gm784	NM_001007580
2625	4.04	4.13	Atp1b2	NM_013415
1244	3.99	3.01	C030003D03Rik	NM_029881
10884	3.99	3.35	Pxmp3	NM_008994
6646	3.93	8.55	Leprel1	NM_173379
11517	3.91	3.65	Slc16a1	NM_009196
621	3.91	34.74	Cxcr4	NM_009911
16652	3.91	3.11	Tmc7	NM_172476
1122	3.90	4.41	Camk1g	NM_144817
14212	3.89	3.36	Calcr	NM_007588

8991	3.88	6.26	Fgf8	NM_010205
19796	3.87	9.89	Amot	NM_153319
16824	3.86	2.96	Fgfr2	NM_010207
22113	3.84	3.24	Adssl1	NM_007421
380	3.81	2.78	Accn4	NM_183022
4518	3.76	6.88	Cap2	NM_026056
19350	3.75	2.73	Smarca1	NM_053123
15263	3.74	5.26	Bhlhb3	NM_024469
21011	3.71	2.64	Dab2	NM_023118
4942	3.70	10.99	Fgf10	NM_008002
10712	3.69	4.51	Cd40	NM_011611
9035	3.67	2.71	Obfc1	NM_175360
18229	3.67	2.61	Pknox2	NM_001029838
13321	3.66	4.59	Ldb2	NM_001077398
14334	3.66	11.55	Podxl	NM_013723
20831	3.65	20.75	Foxa2	NM_010446
11821	3.65	4.47	Slc44a5	NM_001081263
18170	3.64	8.64	Bmper	NM_028472
5758	3.63	3.27	Laptm4b	NM_033521
6664	3.61	4.77	Hes1	NM_008235
2430	3.55	16.01	Hand1	NM_008213
15315	3.53	3.96	Cdc42ep5	NM_021454
20704	3.51	3.40	Rshl2b	NM_001083945
19898	3.49	2.49	Gpm6b	NM_023122
18707	3.45	8.13	Bmp5	NM_007555
13593	3.40	4.35	Ptpn13	NM_011204
19741	3.38	3.05	Tceal1	NM_146236
4517	3.37	6.40	Rbm24	NM_001081425
5415	3.36	2.65	Efha1	NM_028643
19577	3.36	5.69	Asb12	NM_080858
13139	3.34	4.93	Smarcd3	NM_025891
19839	3.26	19.53	Spin2	NM_001005370
16775	3.23	3.09	Zfp768	NM_146202
4445	3.23	5.25	Serpinb6b	NM_011454
9615	3.22	2.76	Klhl23	NM_177784
15063	3.21	2.69	Ltbr	NM_010736
8237	3.20	3.24	Zfp608	NM_175751
2243	3.20	3.91	Slit3	NM_011412
22352	3.19	3.41	Smoc2	NM_022315
21500	3.19	2.74	Snn	NM_009223
20275	3.14	3.19	Ppp1r3c	NM_016854
10476	3.13	2.88	Snph	NM_198214
8218	3.10	2.48	Sema6a	NM_018744
107	3.10	38.28	Cfc1	NM_007685
4932	3.09	5.15	Fst	NM_008046
6168	3.08	2.64	AW049604	NM_134096
15070	3.06	4.10	Ntf3	NM_008742
11667	3.05	12.39	Prss12	NM_008939

21753	3.04	27.64	Ttr	NM_013697
13367	3.01	4.09	Ugdh	NM_009466
16985	3.00	2.71	Tnnt3	NM_011620
3959	2.96	2.32	Npc2	NM_023409
13565	2.95	14.24	Fgf5	NM_010203
11236	2.94	2.31	Thbs3	NM_013691
10329	2.92	2.60	Sirpa	NM_007547
7454	2.91	2.65	Ppt2	NM_019441
3929	2.90	4.21	Dpf3	NM_058212
9709	2.90	2.43	Calcrl	NM_018782
19392	2.90	11.86	Cxx1c	NM_028375
10009	2.89	4.04	Chst1	NM_023850
4265	2.89	2.41	Elmo1	NM_080288
21683	2.87	4.08	Gamt	NM_010255
6182	2.86	2.98	Mapk12	NM_013871
5214	2.86	2.72	Peli2	NM_033602
1372	2.85	3.01	Ddit4	NM_029083
12402	2.84	2.59	2310026E23Rik	NM_026279
2534	2.78	4.28	Mfap4	NM_029568
5295	2.78	3.44	Sall2	NM_015772
7470	2.75	2.45	Zbtb12	NM_198886
3778	2.73	3.38	1110002B05Rik	NM_134054
2901	2.68	2.76	Tmem98	NM_029537
3162	2.66	2.61	Thra	NM_178060
7869	2.66	5.93	Rasgrp3	NM_207246
8234	2.66	20.57	Prdm6	NM_001033281
21054	2.65	2.74	Alas2	NM_001102446
11131	2.64	3.79	Ctso	NM_177662
902	2.63	3.88	Rgs4	NM_009062
3176	2.61	2.46	Smarce1	NM_020618
14496	2.61	9.55	Tmem176b	NM_023056
19394	2.60	3.08	Cxx1b	NM_001018063
11925	2.58	2.46	Rragd	NM_027491
13190	2.55	2.58	Khk	NM_008439
1246	2.53	3.16	L3mbtl3	NM_172787
365	2.52	2.28	BC038286	NM_170755
2969	2.51	9.02	Tbx2	NM_009324
19184	2.51	4.72	Gata1	NM_008089
19832	2.46	3.37	Mageh1	NM_023788
19524	2.46	2.50	Fundc2	NM_026126
14578	2.44	2.53	Pigy	NM_025574
9700	2.43	17.70	Frzb	NM_011356
11279	2.41	2.39	Slc27a3	NM_011988
13247	2.41	3.19	Gm1673	NM_001033458
20199	2.39	5.97	Tmem176a	NM_025326
19585	2.38	4.13	Heph	NM_010417
6498	2.38	7.06	Snai2	NM_011415
1540	2.37	2.51	2610008E11Rik	NM_001004362

11864	2.37	2.21	Ccne2	NM_001037134
18049	2.37	2.45	BC017612	NM_133214
6217	2.35	2.52	Pphln1	NM_146062
18342	2.34	2.86	Zfp202	NM_030713
13428	2.33	18.10	Kdr	NM_010612
19733	2.32	2.44	Tceal8	NM_025703
21785	2.31	2.73	Tspan5	NM_019571
6905	2.28	4.30	Samsn1	NM_023380
16288	2.27	5.78	Mogat2	NM_177448
8071	2.26	4.24	Wnt8a	NM_009290
4568	2.24	10.75	Msx2	NM_013601
21677	2.23	2.57	Dusp10	NM_022019
16265	2.17	2.31	Alg8	NM_199035
19352	2.16	2.97	Apln	NM_013912
21152	2.16	3.42	Col23a1	NM_153393
3957	2.14	9.07	7420416P09Rik	NM_001033776
11723	2.13	2.08	Zcd2	NM_025902
8642	2.11	2.39	2810441K11Rik	NM_026798
19217	2.10	2.77	Tspan7	NM_019634
21940	2.10	7.81	Otx2	NM_144841
17663	2.10	7.57	Dok4	NM_053246
7308	2.10	2.82	Spdef	NM_013891
20508	2.09	20.62	Tnnc1	NM_009393
2180	2.09	2.21	Bcl11a	NM_016707
19074	2.08	2.14	Eif1b	NM_026892
5551	2.03	3.31	Lcp1	NM_008879
2538	2.02	2.67	Grap	NM_027817
6996	2.01	3.29	Pigp	NM_019543
793	2.01	2.55	Ier5	NM_010500
957	1.97	2.28	Atp1a2	NM_178405
9183	1.97	2.65	Commd3	NM_147778
12063	1.96	2.41	Tmod1	NM_021883
10300	1.96	2.29	Kcnip3	NM_019789
4734	1.95	3.46	Irx1	NM_010573
3360	1.95	2.53	Limd2	NM_172397
10145	1.94	3.74	Actc1	NM_009608
15800	1.93	6.93	Etv2	NM_007959
9566	1.92	2.98	Dapl1	NM_029723
6214	1.92	2.30	Yaf2	NM_024189
13141	1.90	2.94	Wdr86	NM_001081441
4653	1.87	6.18	Gas1	NM_008086
18204	1.86	2.90	Kcnj5	NM_010605
22077	1.86	2.24	Cdkn1b	NM_009875
9592	1.85	2.05	Ttc21b	NM_001047604
20476	1.84	3.08	Gulp1	NM_027506
15642	1.82	2.71	B3gnt8	NM_001036740
1347	1.81	2.41	Gja1	NM_010288
14736	1.80	2.75	Cml1	NM_023160

2849	1.78	2.47	Tlcd1	NM_026708
105	1.78	3.72	Ptpn18	NM_011206
21919	1.77	10.20	Fgf3	NM_008007
2712	1.75	4.13	6330403K07Rik	NM_134022
14764	1.75	3.41	Prokr1	NM_021381
22249	1.71	2.97	Foxb1	NM_022378
11784	1.71	2.50	Mcoln3	NM_134160
9737	1.71	16.39	Agtrl1	NM_011784
9261	1.70	3.50	BC029214	NM_153557
49	1.68	2.89	Jph1	NM_020604
7947	1.67	6.40	Bambi	NM_026505
10179	1.67	2.48	Zfyve19	NM_028054
19387	1.67	13.43	Plac1	NM_019538
17620	1.67	3.11	Tox3	NM_172913
19759	1.66	3.29	Mum1l1	NM_175541
11790	1.66	3.32	Gng5	NM_010318
98	1.66	2.42	Bag2	NM_145392
11926	1.65	2.08	Ube2j1	NM_019586
1953	1.64	2.29	Rps26	NM_013765
14242	1.61	2.19	Ndufa4	NM_010886
12456	1.61	2.03	Urod	NM_009478
4803	1.60	2.25	Serinc5	NM_172588
14437	1.60	2.23	3321401G04Rik	NM_029930
7484	1.60	2.09	Ddah2	NM_016765
9248	1.59	2.50	Ssna1	NM_023464
7971	1.59	2.14	Colec12	NM_130449
21945	1.58	2.09	Lsm2	NM_001110101
8802	1.56	2.18	2900009I07Rik	NM_026520
14971	1.55	2.02	Rad52	NM_011236
13601	1.53	3.49	Sparcl1	NM_010097
7662	1.52	2.29	1700027N10Rik	NM_029338
17012	1.50	3.27	Tmem16a	NM_178642
13742	1.50	3.31	Msi1	NM_008629
18460	1.50	3.84	Pih1d2	NM_028300
8672	1.50	12.77	Ms4a4d	NM_025658
13129	1.50	2.13	Abcb8	NM_029020
9671	1.48	2.18	Ttc30b	NM_028235
19823	1.47	2.04	Hsd17b10	NM_016763
13995	1.44	2.45	Rabl5	NM_026073
21747	1.43	2.19	Ripk4	NM_023663
16161	1.41	28.59	Mesp1	NM_008588
10889	1.41	2.85	Hey1	NM_010423
3762	1.41	2.67	Stxbp6	NM_144552
1090	1.40	3.88	Prox1	NM_008937
22548	1.40	2.60	Gdap1l1	NM_144891
12414	1.40	5.77	Tal1	NM_011527
5134	1.38	7.03	Lrrc18	NM_026253
18042	1.38	2.24	Panx1	NM_019482

7018	1.37	6.87	Pcp4	NM_008791
178	1.35	1.96	Kdelc1	NM_023645
3260	1.33	2.10	Ccdc56	NM_026618
12829	1.32	3.65	Mfap2	NM_008546
18022	1.32	4.48	9230110C19Rik	NM_199017
1098	1.32	3.40	Nenf	NM_025424
11886	1.31	2.55	Decr1	NM_026172
18877	1.30	1.96	Rpl29	NM_009082
6194	1.29	3.21	Cpt1b	NM_009948
1937	1.28	2.00	Tmem4	NM_019953
13296	1.27	2.26	Stx18	NM_026959
9269	1.27	2.37	Phpt1	NM_029293
17896	1.26	2.03	Mthfsd	NM_172761
4204	1.25	3.60	Akr1c19	NM_001013785
20434	1.24	2.99	2810030E01Rik	NM_028317
14618	1.24	2.04	Rpia	NM_009075
14948	1.23	2.70	Cxcl12	NM_001012477
1273	1.22	2.62	Tspyl4	NM_030203
5330	1.22	1.88	Homez	NM_183174
13241	1.21	2.05	Tmem129	NM_026698
14930	1.20	2.03	Ift122	NM_031177
11660	1.19	2.30	Pde5a	NM_153422
18934	1.19	1.99	Tcta	NM_133986
7899	1.18	2.00	AW548124	NM_134117
901	1.17	13.76	Rgs5	NM_009063
3826	1.17	2.15	Klhdc2	NM_027117
7837	1.14	1.96	Epb4.1l3	NM_013813
1819	1.14	1.90	Thap2	NM_025780
4462	1.14	2.86	3110004L20Rik	NM_001033167
8800	1.13	2.13	Fxn	NM_008044
899	1.09	4.33	Pbx1	NM_183355
18547	1.08	2.88	Cyp1a1	NM_009992
807	1.05	2.40	Soat1	NM_009230
19578	1.05	2.12	AK129302	NM_001003916
3095	1.05	10.82	Hoxb2	NM_134032
3107	1.03	2.40	Sp6	NM_031183
7026	1.01	2.23	5830404H04Rik	NM_174847
7853	1.00	4.40	Lbh	NM_029999
12376	0.99	2.84	Gpx7	NM_024198
8812	0.97	2.75	Smarca2	NM_011416
12095	0.97	2.85	2810432L12Rik	NM_025944
11523	0.96	2.00	St7l	NM_153091
16314	0.96	2.01	Chchd8	NM_183270
5518	0.95	2.38	Bin3	NM_021328
11032	0.95	1.96	Lhfp	NM_175386
19816	0.94	2.57	Maged2	NM_030700
6036	0.93	3.27	1700088E04Rik	NM_138581
9330	0.93	4.26	Olfm1	NM_001038613

17606	0.93	2.23	Cbln1	NM_019626
9450	0.93	3.22	Pbx3	NM_016768
15648	0.92	3.93	Tgfb1	NM_011577
86	0.91	1.89	Lmbrd1	NM_026719
17676	0.91	3.06	Mmp15	NM_008609
9408	0.90	1.88	Ppapdc3	NM_145521
235	0.90	1.90	Ppil3	NM_027351
19248	0.89	2.29	Slc9a7	NM_177353
19683	0.88	2.41	Klhl4	NM_172781
22116	0.87	2.18	1700010I14Rik	NM_025851
21756	0.86	3.59	Ppic	NM_008908
6181	0.85	2.37	Hdac10	NM_199198
19730	0.85	1.99	Bex2	NM_009749
6851	0.85	2.34	Cldnd1	NM_171826
13089	0.84	2.17	Sema3a	NM_009152
10010	0.84	2.21	Syt13	NM_030725
17006	0.83	2.10	Mrgpre	NM_175534
2213	0.82	3.61	Nsg2	NM_008741
16862	0.80	1.82	Uros	NM_009479
18388	0.80	2.27	Rps25	NM_024266
19298	0.78	2.52	Sep-06	NM_019942
19094	0.78	1.96	Tmem16k	NM_133979
11873	0.78	1.82	Tmem67	NM_177861
17696	0.78	1.89	Cmtm3	NM_024217
22370	0.78	2.18	Pnliprp2	NM_011128
2978	0.77	2.29	Ptrh2	NM_175004
6740	0.76	2.32	Slc15a2	NM_021301
18851	0.75	3.08	Tmem108	NM_178638
18693	0.74	2.15	Arpp19	NM_021548
16786	0.74	2.86	Zfp629	NM_177226
14870	0.74	2.48	Lrrn1	NM_008516
2961	0.71	2.42	Pigw	NM_027388
2718	0.71	2.11	6720460F02Rik	NM_144526
10385	0.70	5.62	B430119L13Rik	NM_177303
19860	0.69	2.03	Sh3kbp1	NM_021389
6430	0.69	2.03	Nmral1	NM_026393
20480	0.67	2.41	Lrrfip1	NM_008515
14353	0.66	3.58	Slc13a4	NM_172892
11706	0.65	2.31	Papss1	NM_011863
2396	0.64	5.23	Pdlim4	NM_019417
22279	0.64	2.19	Gcgr	NM_008101
15643	0.64	3.46	B3gnt8	NM_146184
17267	0.63	3.05	6430573F11Rik	NM_176952
8486	0.63	1.97	Yif1a	NM_026553
8634	0.62	1.97	Fen1	NM_007999
11629	0.61	1.87	Cdc14a	NM_001080818
19151	0.61	2.48	Nudt10	NM_001031664
3977	0.61	1.95	0610007P14Rik	NM_021446

9254	0.60	2.18	Uap1l1	NM_001033293
9523	0.59	1.99	Rpl35	NM_025592
19186	0.58	3.06	Suv39h1	NM_011514
19504	0.58	1.77	Taz	NM_181516
19383	0.58	6.00	Gpc3	NM_016697
3022	0.57	2.39	Tmem100	NM_026433
16986	0.57	3.64	Mrpl23	NM_011288
4964	0.56	2.70	LOC544988	NM_001024712
6659	0.55	2.08	Hrasls	NM_013751
5027	0.55	2.06	Fut11	NM_028428
687	0.55	2.55	Adora1	NM_001039510
7070	0.52	3.89	T	NM_009309
15043	0.52	1.85	Ptms	NM_026988
5829	0.48	1.88	2600005O03Rik	NM_183089
4025	0.47	2.25	2610021K21Rik	NM_030172
19719	0.47	2.64	Tceal6	NM_025355
18000	0.47	1.97	Gria4	NM_019691
20742	0.46	2.39	1500011H22Rik	NM_026883
1491	0.44	2.07	A130042E20Rik	NM_172550
18630	0.44	2.03	Rbpms2	NM_028030
21874	0.43	1.87	Rab39b	NM_175122
19399	0.43	2.26	Zfp449	NM_030139
752	0.43	2.08	Rgs2	NM_009061
9062	0.43	1.85	Vti1a	NM_016862
19812	0.42	2.15	Apex2	NM_029943
21872	0.42	2.13	Mospd1	NM_027409
14230	0.41	2.23	Shfm1	NM_009169
9096	0.41	2.65	Slc18a2	NM_172523
5617	0.41	2.02	2610206B13Rik	NM_026047
18571	0.41	2.30	Hexa	NM_010421
3729	0.41	2.87	Atxn7l4	NM_028139
4938	0.41	2.15	Parp8	NM_001081009
18634	0.40	3.42	2810417H13Rik	NM_026515
1922	0.40	1.87	Naca	NM_013608
19164	0.40	1.89	Syp	NM_009305
11529	0.38	1.90	Rap1a	NM_145541
16069	0.38	2.11	Tmem16e	NM_177694
2840	0.38	1.87	Sez6	NM_021286
1029	0.35	2.17	Sccpdh	NM_178653
21203	0.35	2.74	Prss1	NM_053243
10622	0.35	2.04	Lbp	NM_008489
19433	0.34	1.79	Ids	NM_001038990
19291	0.34	4.05	Zcchc12	NM_028325
3312	0.33	2.51	Ccdc103	NM_028492
14950	0.32	2.61	Zfp637	NM_177684
12168	0.29	2.07	Cdc26	NM_139291
10042	0.29	2.06	Apip	NM_019735
17985	0.29	11.80	Nrp1	NM_008737

19670	0.28	2.42	Sh3bgrl	NM_019989
15277	0.26	1.96	Ccdc91	NM_025911
14746	0.26	1.79	C87436	NM_146170
7483	0.25	2.36	Clic1	NM_033444
3761	0.24	2.19	Nrcam	NM_176930
11882	0.24	1.74	Tmem55a	NM_028264
18896	0.24	1.94	Tmem115	NM_019704
3661	0.23	1.83	Trib2	NM_144551
19538	0.22	1.86	Tmem47	NM_138751
21457	0.22	2.36	Pddc1	NM_172116
1193	0.21	3.06	Slc35d3	NM_029529
8781	0.20	2.08	1110059E24Rik	NM_025423
2220	0.19	13.24	Hba-a1	NM_008218
4793	0.19	3.42	Ssbp2	NM_024272
11845	0.17	2.23	Chchd7	NM_181391
4279	0.17	1.97	OTTMUSG00000000421	NM_001039115
19388	0.16	1.92	4632404H22Rik	NM_030167
7258	0.16	1.94	2610003J06Rik	NM_028101
14502	0.16	2.67	Igf2bp3	NM_023670
5101	0.15	2.43	Glt8d1	NM_029626
3441	0.15	2.10	Nat9	NM_025400
6895	0.14	2.11	Chmp2b	NM_026879
3654	0.12	4.32	Msgn1	NM_019544
9291	0.12	2.05	D2Bwg1335e	NM_026828
22064	0.12	3.24	Scand1	NM_020255
9267	0.12	1.77	Edf1	NM_021519
4799	0.08	1.93	Dhfr	NM_010049
9184	0.08	2.38	Bmi1	NM_007552
5005	0.07	2.61	Rpl15	NM_025586
18997	0.07	8.27	Myl3	NM_010859
18482	0.06	2.89	Kdelc2	NM_212445
17895	0.06	14.31	Foxf1a	NM_010426
4814	0.06	2.19	Lhfpl2	NM_172589
3831	0.05	2.30	Cdkl1	NM_183294
14580	0.03	2.18	AW146242	NM_146168
18466	0.03	3.44	1110032A03Rik	NM_023483
9198	0.02	2.18	Gad2	NM_008078
5202	0.02	2.93	Cgrrf1	NM_026832
4989	0.01	1.89	Thoc7	NM_025435
19659	0.01	1.80	Gpr23	NM_175271
9666	0.00	1.71	Mtx2	NM_016804
3564	0.00	2.71	Slc25a10	NM_013770
22084	0.00	2.21	Echdc1	NM_001110195
14821	0.00	3.41	Fbln2	NM_007992

Supplementary Table 6. Identification of full length transcripts by diagnostic sequences. Diagnostic exons are defined as exon regions which are unique to one transcript in Aceview. Diagnostic junctions are defined as a junction coordinate (donor-acceptor pair) which is unique to one Aceview transcript. (A). Total number of transcripts and Loci in Aceview which have either a diagnostic junction or diagnostic exon. (B). Statistics for diagnostic exons or junctions from the ES and EB SQRL sequencing data.

A

	Number of transcripts	% of total transcripts	Number of loci	% of total loci
Total diagnostic	160156	92.57%	65254	99.83%
Diagnostic junctions	78971	45.65%	31119	47.61%
Diagnostic exons	138156	79.86%	65254	99.83%
Combined (both junction and exon)	56971	32.93%	30130	46.10%

B

	Transcripts (ES)	Loci (ES)	Transcripts (EB)	Loci (EB)
Total diagnostic	53056	31872	50881	29606
Diagnostic junctions (2+ tags)	10454	7767	12884	9018
Diagnostic exons (10+ tags)	48017	31302	44071	28744
Combined (2+ junction, 10+ exon)	5415	5058	6074	5607

Supplementary Table 7. Volume and proportion of the mouse genome that falls into the known and predicted regions of transcription defined in Table 1. Genome size includes all "random" contigs available in the mm9 release, and as regions are defined as strand specific, the total number of nucleotides in the genome has been doubled. Note that each of these regions will overlap with other categories (eg. predicted alternative transcripts will contain some known exons, etc.).

	Volume	Proportion
Mouse genome (mm9):	5.45 Gb	100%
Known expression:		
Known exons	0.27 Gb	4.9%
Predicted Exons	0.20 Gb	3.6%
Novel Expression:		
Known regions	1.90 Gb	34.8%
Predicted regions	2.84 Gb	52.0%
Conserved regions	3.83 Gb	70.2%

Supplementary Table 8. Number and proportion of repeat associated tags detected in SQRL libraries, broken down by type of repeat.

Type of repeat	number of tags (ES)	% of all tags (ES)	number of tags (EB)	% of all tag (EB)
DNA transposons	421156	0.450805	313519	0.357134
LINE.CR1	30843	0.033014	27182	0.030963
LINE.L1	4390151	4.699212	3717364	4.234508
LINE.L2	244963	0.262208	203407	0.231704
LINE.RTE	11229	0.01202	11108	0.012653
Low_complexity	1256516	1.344973	1547150	1.762383
LTR.ERV1	281533	0.301353	215263	0.245209
LTR.ERVK	1369700	1.466125	1173019	1.336205
LTR.ERVL	449065	0.480679	378954	0.431673
LTR.MaLR	1646508	1.76242	1308817	1.490894
rRNA	2712218	2.903155	2575224	2.933478
Satellite	224832	0.24066	130037	0.148127
Simple_repeat	1124117	1.203253	1343761	1.530699
SINE.Alu	1107182	1.185126	949980	1.082137
SINE.B2	3000583	3.21182	1457014	1.659708
SINE.B4	1180692	1.263811	985337	1.122413
SINE.ID	90351	0.096712	71422	0.081358
SINE.MIR	256764	0.27484	201363	0.229376
Other	540462	0.57851	502549	0.572462
Total	**20338865**	**21.7707**	**17112470**	**19.49309**

Supplementary Table 9. Analysis of natural sense-antisense transcript pairs in ES and EB. The gene set used for this analysis contains 1573 genes with overlap and coding information from Kiyosawa et al., 2006 (PMID: 15781571). To be selected, the tag count must be greater than average of 1 tag start per 6 bases (ie. above the pervasive transcription threshold set in Table 1) within the transcript and expressed in both sense and antisense strand.

Table S28a – nature of the overlap.

	All	Convergent (tail to tail)		Divergent (head to head)		Full (overlap)	
		number	*proportion*	*number*	*proportion*	*number*	*proportion*
Total Pairs	1573	617	39.22%	582	37.00%	374	23.78%
ES pairs	124	93	75.00%	29	23.39%	1	0.81%
EB pairs	141	99	70.21%	39	27.66%	3	2.13%

Table S28b – coding status of each pair.

	All	Coding-Coding		Coding-Noncoding		Noncoding-Noncoding	
		number	*proportion*	*number*	*proportion*	*number*	*proportion*
Total pairs	1573	685	43.55%	719	45.71%	169	10.74%
ES pairs	124	110	88.71%	12	9.68%	1	0.81%
EB pairs	141	121	85.82%	19	13.48%	1	0.71%

Supplementary Table 10. Exon-junction discovery in SQRL ESC and EB libraries. Exon-junction sequences (30 nt of donor exon sequence concatenated with 30 nt of acceptor exon sequence) are not expected to map to the genome, unless there have been retro-transposition events – almost 9000 junctions with EST support fall into this category. As we would not be able to accurately map the origin of RNA tags mapping to these junctions sequences, they were removed from the exon-junction library as "uninformative". Almost half of the "novel" exons match exactly these previously known but uninformative junctions, serving as a positive control for our approach. Novel junctions were then classified into those deriving from known transcripts, and those deriving from novel transcripts.

	Known (uninformative) junctions	**Novel junctions in known transcripts**	**Novel junctions in novel transcripts**	**Total**
ES	12	7	3	22
EB	30	18	17	65
Total	42	25	20	87

Supplementary Table 11. Single nucleotide polymorphism (SNP) detection in ESC and mEB samples. High quality SNP predictions were made for each chromosome in both the ESC and EB libraries (See supplementary Figure S21). Of the total SNPs common to both libraries, 20.3% overlap with dbSNP 128. However, when confined to Refseq transcripts, 83.9% of SNPs overlap with dbSNP 128.

	ES-all	EB-all	Common to ES and EB	In dbSNP (%)	Present in Refseq	Refseq in dbSNP (%)
chr1	172	448	162	24.07	33	81.82
chr2	182	461	171	25.15	44	81.82
chr3	164	399	154	21.43	34	79.41
chr4	174	432	166	19.88	46	80.43
chr5	166	426	153	22.22	45	80.00
chr6	146	387	140	17.14	27	96.30
chr7	133	350	130	25.38	41	73.17
chr8	161	376	154	22.08	60	85.00
chr9	129	316	122	24.59	51	84.31
chr10	145	360	137	20.44	12	91.67
chr11	145	372	141	24.82	75	72.00
chr12	118	290	115	17.39	19	94.74
chr13	144	335	135	17.04	13	100.00
chr14	143	345	136	16.91	36	91.67
chr15	131	292	122	22.13	25	92.00
chr16	135	295	127	14.96	15	46.67
chr17	162	364	153	18.30	26	92.31
chr18	149	373	141	15.60	21	76.19
chr19	106	226	95	25.26	10	90.00
chrX	174	455	160	12.50	10	10.00
chrY	12	20	9	0.00	0	0.00
chrM	0	1	0	0.00	0	0.00
Total	2991	7323	2823	20.3	643	83.90

Supplementary Table 12. Summary of SNP validation results. Genomic regions were amplified using a high fidelity Taq. The resulting PCR products were both sequenced directly and cloned using TOPO-TA (Invitrogen). Individual colonies were picked and sequenced by Sanger-capillary sequencing. Validation was preferentially carried out on Refseq genes where the SNP was predicted to result in a non-synonymous substitution, and then targeted changes in the UTR regions.

Chromosome	Position	Strand	Gene UTRme	Predicted SNP	Amino acid change	Validated
chr1	174437452	+	Tagln2	C->T	*UTR*	Yes
chr1	34501748	+	Imp4	C->T	*UTR*	Yes
chr11	84777618	+	Car4	C->T	A->V	Yes
chr14	22616785	+	Samd8	A->G	*UTR*	Yes
chr18	38460154	+	Rnf14	C->T	A->V	Yes
chr19	41992404	+	Pgam1	C->T	*UTR*	Yes
chr2	119453391	+	Nusap1	A->G	K->R	Yes
chr3	121978860	+	Dnttip2	T->A	L->Q	Yes
chr3	93330080	+	S100a11	C->T	*UTR*	Yes
chr4	15903284	+	Nbn	C->G	Y->F	Yes
chr5	138548192	+	Zkscan1	A->G	*UTR*	Yes
chr5	138548273	+	Zkscan1	A->G	*UTR*	Yes
chr7	38752418	-	C80913	T->C	N->D	Yes
chr7	50209670	-	EG668668	C->G	G->R	Yes
chr7	5076892	-	Rasl2-9	T->A	E->D	No
chr9	123371057	+	Lars2	T->G	*UTR*	Yes
chr9	57681529	-	Arid3b	G->T	Q->K	Yes
chr9	64083463	+	Uchl4	A->T	D->E	No

Supplementary Table 13. Summary of variants isolated from the SQRL data from pathways involved in ES cell pluripotency and differentiation. Numbers show transcripts or loci which are identified by at least 2 diagnostic junctions.

TGF-beta pathway (Ingenuity pathways analysis).

Sample	Loci detected	All variants detected	Transcripts per loci
ES	25	38	1.52
EB	25	41	1.64

Wnt pathway (Ingenuity pathways analysis)

Sample	Loci detected	All variants detected	Transcripts per loci
ES	31	46	1.48
EB	33	55	1.67

FGF pathway (Ingenuity pathways analysis)

Sample	Loci detected	All variants detected	transcripts per loci
ES	20	30	1.50
EB	21	36	1.71

Mouse regulatory network: (PMID: 17940043).

Sample	Loci detected	All variants detected	transcripts per loci
ES	26	38	1.46
EB	27	45	1.67

Supplementary Table 14. Details of splice-variants isolated from the SQRL data from pathways involved in ES cell pluripotency and differentiation. The TGFB, Wnt, and FGF pathways were derived from Ingenuity Pathway Analysis (www.ingenuity.com), while the Mouse Regulatory Network was derived from PMID: 17940043.

Table 14a. The TGFB pathway.

Gene Name	Ace Name	Number of total Diagnostic junctions per variant	Number of Diagnostic junctions present in ES	Counts per Junction - ES	Number of Diagnostic junctions present in EB	Counts per Junction - EB
Acvr2a	Acvr2a.aSep07	3	2	,6,4	2	,19,10
Acvr2a	Acvr2a.bSep07	1	0	0	1	,1
Axl	Axl.dSep07	1	1	,1	0	0
Axl	Axl.aSep07	1	0	0	1	,1
Axl	Axl.fSep07	1	0	0	1	,1
Bmpr1a	Bmpr1a.aSep07	2	2	,3,3	2	,12,3
Crebbp	Crebbp.aSep07	8	7	,6,2,1,5,1,8,6	8	,16,13,3,12,4,2,17,11
Crebbp	Crebbp.bSep07	1	0	0	1	,6
Ddr1	Ddr1.gSep07	1	1	,1	1	,2
Ddr1	Ddr1.aSep07	1	1	,2	1	,2
Ddr1	Ddr1.iSep07	1	1	,1	0	0
Egfr	Egfr.aSep07	6	6	,1,2,1,2,5,12	4	,3,2,1,9
Egfr	Egfr.fSep07	1	2	,1,1	0	0
Ep300	Ep300.aSep07	10	9	,3,2,7,1,48,64,1,1,85	8	,9,2,13,87,56,1,2,141
Ep300	Ep300.fSep07	1	1	,5	1	,3
Fgfr1	Fgfr1.jSep07	1	1	,4	1	,34
Fgfr1	Fgfr1.fSep07	1	0	0	1	,2
Grb2	Grb2.bSep07	1	1	,252	1	,98
Grb2	Grb2.dSep07	1	1	,5	1	,1
Hnf4a	Hnf4a.bSep07	1	1	,1	1	,1
Hnf4a	Hnf4a.gSep07	1	1	,1	0	0
Hras1	Hras1.eSep07	2	2	,1,1	1	,1
Hras1	Hras1.dSep07	2	1	,3	1	,3
Kras	Kras.aSep07	2	2	,2,3	2	,2,8
Kras	Kras.fSep07	1	1	,2	1	,2
Map2k4	Map2k4.eSep07	1	1	,4	1	,4
Map2k4	Map2k4.cSep07	1	0	0	1	,3
Map3k7	Map3k7.aSep07	1	1	,14	1	,28

Map3k7	Map3k7.iSep07	1	1	,1	1	,1
Map3k7ip1	Map3k7ip1.aSep07	1	1	,3	1	,9
Map3k7ip1	Map3k7ip1.cSep07	1	1	,1	0	0
Mapk1	Mapk1.dSep07	1	1	,2	1	,2
Mapk1	Mapk1.aSep07	1	1	,1	1	,5
Mras	Mras.cSep07	1	1	,1	1	,1
Mras	Mras.bSep07	1	1	,4	0	0
Nras	Nras.eSep07	3	3	,2,4,1	2	,6,2
Nras	Nras.cSep07	1	1	,1	1	,11
Nras	Nras.aSep07	1	1	,13	1	,10
Nras	Nras.fSep07	1	1	,4	1	,2
Nras	Nras.gSep07	1	0	0	1	,1
Ntrk1	Ntrk1.aSep07	12	4	,5,1,1,1	4	,1,2,1,1
Ntrk1	Ntrk1.bSep07	2	1	,2	0	0
Pdgfrb	Pdgfrb.aSep07	2	1	,1	1	,1
Pdgfrb	Pdgfrb.eSep07	1	0	0	1	,1
Ptk7	Ptk7.aSep07	11	9	,21,3,12,4,130,5,58,5,4	10	,4,54,17,30,16,389,23,433,10,20
Ptk7	Ptk7.dSep07	1	0	0	1	,1
Raf1	Raf1.aSep07	1	1	,5	1	,5
Raf1	Raf1.jSep07	1	0	0	1	,1
Smad2	Smad2.eSep07	1	1	,2	0	0
Smad2	Smad2.cSep07	1	0	0	1	,1
Smad2	Smad2.dSep07	2	0	0	1	,4
Smad4	Smad4.bSep07	1	1	,5	1	,26
Smad4	Smad4.dSep07	2	1	,2	1	,4
Smad4	Smad4.eSep07	1	0	0	1	,3
Smad7	Smad7.bSep07	1	1	,3	1	,2
Smad7	Smad7.aSep07	1	1	,2	1	,9
Smurf1	Smurf1.aSep07	1	1	,2	1	,4
Smurf1	Smurf1.fSep07	1	1	,1	1	,2
Tcfe3	Tcfe3.lSep07	1	1	,48	1	,10
Tcfe3	Tcfe3.fSep07	1	1	,5	1	,2
Tcfe3	Tcfe3.jSep07	1	1	,3	1	,1
Tcfe3	Tcfe3.iSep07	2	0	0	1	,1
Tgfb1	Tgfb1.aSep07	3	3	,14,2,5	3	,115,31,41
Tgfb1	Tgfb1.bSep07	1	0	0	1	,5
Tgfb2	Tgfb2.aSep07	2	2	,11,1	1	,5
Tgfb2	Tgfb2.bSep07	1	1	,4	1	,1
Tgfbr1	Tgfbr1.cSep07	1	1	,1	1	,1
Tgfbr1	Tgfbr1.bSep07	1	0	0	1	,1
Tgfbr2	Tgfbr2.aSep07	1	1	,1	0	0

Tgfbr2	Tgfbr2.bSep07	1	0	0	1	,1

Table 14b. The Wnt Pathway.

Gene Name	Ace Name	Number of total Diagnostic junctions per variant	Number of Diagnostic junctions present in ES	Counts per Junction - ES	Number of Diagnostic junctions present in EB	Counts per Junction - EB
Acvr2a	Acvr2a.aSep07	3	2	,6,4	2	,19,10
Acvr2a	Acvr2a.bSep07	1	0	0	1	,1
Akt2	Akt2.aSep07	2	1	,1	1	,1
Akt2	Akt2.qSep07	2	1	,1	1	,1
Akt3	Akt3.aSep07	4	4	,2,166,39,26	5	,3,329,65,96,3
Akt3	Akt3.bSep07	3	3	,1,2,1	2	,5,15
Akt3	Akt3.eSep07	1	1	,1	1	,3
Akt3	Akt3.cSep07	1	0	0	1	,2
Apc	Apc.aSep07	2	2	,6,7	2	,5,10
Apc	Apc.fSep07	1	1	,1	1	,3
Axin1	Axin1.aSep07	5	3	,4,1,14	4	,1,10,2,17
Axin1	Axin1.bSep07	2	2	,1,24	2	,3,40
Axin2	Axin2.aSep07	5	4	,7,1,1,12	4	,3,7,1,22
Axin2	Axin2.bSep07	1	1	,2	1	,15
Cd44	Cd44.cSep07	1	1	,1	1	,2
Cd44	Cd44.eSep07	1	1	,2	1	,14
Cdkn2a	Cdkn2a.bSep07	1	1	,5	1	,2
Cdkn2a	Cdkn2a.aSep07	1	1	,2	0	0
Csnk1a1	Csnk1a1.aSep07	1	1	,12	1	,22
Csnk1a1	Csnk1a1.kSep07	1	1	,16	1	,13
Csnk1d	Csnk1d.pSep07	2	0	0	2	,1,2
Csnk1d	Csnk1d.iSep07	2	0	0	1	,1
Csnk1e	Csnk1e.dSep07	2	1	,6	2	,3,1
Csnk1e	Csnk1e.fSep07	1	1	,2	0	0
Csnk1e	Csnk1e.hSep07	1	1	,1	0	0
Csnk1e	Csnk1e.eSep07	2	0	0	1	,1
Csnk1g1	Csnk1g1.aSep07	2	1	,5	1	,14
Csnk1g1	Csnk1g1.bSep07	2	1	,1	1	,1
Csnk1g2	Csnk1g2.eSep07	1	1	,4	1	,11
Csnk1g2	Csnk1g2.fSep07	1	1	,5	0	0
Csnk2a2	Csnk2a2.dSep07	1	1	,6	1	,6
Csnk2a2	Csnk2a2.cSep07	2	1	,1	0	0
Csnk2a2	Csnk2a2.hSep07	1	1	,1	0	0

Csnk2b	Csnk2b.hSep07	1	1	,1	1	,6
Csnk2b	Csnk2b.dSep07	1	1	,1	0	0
Ctnnb1	Ctnnb1.dSep07	2	0	0	1	,5
Ctnnb1	Ctnnb1.bSep07	1	0	0	1	,1
Ctnnb1	Ctnnb1.jSep07	1	0	0	1	,2
Dkk3	Dkk3.bSep07	1	1	,1	1	,2
Dkk3	Dkk3.dSep07	1	1	,1	1	,1
Ep300	Ep300.aSep07	10	9	,3,2,7,1,48,64,1,1,85	8	,9,2,13,87,56,1,2,141
Ep300	Ep300.fSep07	1	1	,5	1	,3
Gnao1	Gnao1.eSep07	1	1	,3	1	,10
Gnao1	Gnao1.aSep07	1	0	0	1	,1
Gsk3b	Gsk3b.bSep07	3	2	,1,1	1	,2
Gsk3b	Gsk3b.aSep07	1	1	,5	1	,11
Gsk3b	Gsk3b.dSep07	1	1	,2	1	,1
Kremen1	Kremen1.aSep07	1	1	,1	1	,5
Kremen1	Kremen1.bSep07	2	1	,3	0	0
Map3k7	Map3k7.aSep07	1	1	,14	1	,28
Map3k7	Map3k7.iSep07	1	1	,1	1	,1
Map3k7ip1	Map3k7ip1.aSep07	1	1	,3	1	,9
Map3k7ip1	Map3k7ip1.cSep07	1	1	,1	0	0
Myc	Myc.bSep07	1	1	,9	1	,23
Myc	Myc.cSep07	1	1	,7	1	,28
Myc	Myc.dSep07	1	1	,1	1	,6
Nlk	Nlk.aSep07	6	3	,4,17,2	4	,2,36,5,7
Nlk	Nlk.cSep07	1	1	,2	1	,1
Ppp2r4	Ppp2r4.lSep07	1	0	0	1	,2
Ppp2r4	Ppp2r4.hSep07	2	0	0	1	,1
Ppp2r4	Ppp2r4.iSep07	1	0	0	1	,1
Ppp2r5c	Ppp2r5c.eSep07	2	1	,2	1	,3
Ppp2r5c	Ppp2r5c.aSep07	1	0	0	1	,2
Ppp2r5d	Ppp2r5d.bSep07	1	1	,8	1	,12
Ppp2r5d	Ppp2r5d.cSep07	1	0	0	1	,5
Ppp2r5e	Ppp2r5e.fSep07	1	1	,1	0	0
Rara	Rara.fSep07	1	1	,7	1	,1
Rara	Rara.dSep07	1	1	,1	0	0
Rara	Rara.aSep07	2	0	0	1	,1
Rarg	Rarg.bSep07	1	1	,10	1	,8
Rarg	Rarg.dSep07	1	1	,2	0	0
Sox13	Sox13.bSep07	2	1	,2	1	,3
Sox13	Sox13.aSep07	2	1	,1	1	,1
Src	Src.aSep07	2	2	,5,2	2	,10,3

Src	Src.hSep07	1	1	,1	1	,1
Tcf4	Tcf4.cSep07	1	1	,3	1	,2
Tcf4	Tcf4.vgSep07	1	1	,1	1	,1
Tcf4	Tcf4.gSep07	1	1	,1	0	0
Tcf7l2	Tcf7l2.eSep07	1	1	,1	1	,2
Tcf7l2	Tcf7l2.iSep07	1	1	,1	0	0
Tcf7l2	Tcf7l2.aSep07	1	0	0	1	,2
Tcf7l2	Tcf7l2.gSep07	1	0	0	1	,2
Tcf7l2	Tcf7l2.kSep07	1	0	0	1	,1
Tcf7l2	Tcf7l2.nSep07	1	0	0	1	,1
Tgfb2	Tgfb2.aSep07	2	2	,11,1	1	,5
Tgfb1	Tgfb1.bSep07	1	0	0	1	,5
Tgfb2	Tgfb2.bSep07	1	1	,4	1	,1
Tgfbr1	Tgfbr1.cSep07	1	1	,1	1	,1
Tgfbr1	Tgfbr1.bSep07	1	0	0	1	,1
Tgfbr2	Tgfbr2.aSep07	1	1	,1	0	0
Tgfbr2	Tgfbr2.bSep07	1	0	0	1	,1
Tle1	Tle1.fSep07	1	1	,3	1	,1
Tle1	Tle1.gSep07	3	1	,2	0	0
Tle3	Tle3.gSep07	1	1	,1	1	,1
Tle3	Tle3.aSep07	1	0	0	1	,3
Wnt5a	Wnt5a.bSep07	1	0	0	1	,3
Wnt5a	Wnt5a.fSep07	1	0	0	1	,3
Wnt5b	Wnt5b.aSep07	1	1	,3	1	,2
Wnt5b	Wnt5b.dSep07	1	0	0	1	,1

Table 14c. The FGF pathway.

Gene Name	Ace Name	Number of total Diagnostic junctions per variant	Number of Diagnostic junctions present in ES	Counts per Junction - ES	Number of Diagnostic junctions present in EB	Counts per Junction - EB
Akt2	Akt2.aSep07	2	1	,1	1	,1
Akt2	Akt2.qSep07	2	1	,1	1	,1
Akt3	Akt3.aSep07	4	4	,2,166,39,26	5	,3,329,65,96,3
Akt3	Akt3.bSep07	3	3	,1,2,1	2	,5,15
Akt3	Akt3.eSep07	1	1	,1	1	,3
Akt3	Akt3.cSep07	1	0	0	1	,2

Atf2	Atf2.fSep07	1	1	,2	1	,4
Atf2	Atf2.qSep07	2	1	,3	1	,2
Creb3	Creb3.aSep07	1	1	,6	1	,5
Creb3	Creb3.eSep07	1	0	0	1	,2
Crk	Crk.eSep07	1	1	,6	1	,15
Crk	Crk.aSep07	1	1	,16	1	,25
Crk	Crk.dSep07	1	1	,1	0	0
Fgf1	Fgf1.bSep07	1	1	,1	0	0
Fgf1	Fgf1.fSep07	1	1	,2	0	0
Fgf10	Fgf10.aSep07	1	1	,2	1	,43
Fgf10	Fgf10.bSep07	2	0	0	1	,1
Fgf13	Fgf13.iSep07	1	1	,1	0	0
Fgf13	Fgf13.bSep07	1	0	0	1	,3
Fgf17	Fgf17.aSep07	2	1	,60	1	,11
Fgf17	Fgf17.bSep07	1	1	,14	0	0
Fgf5	Fgf5.aSep07	1	0	0	2	,4,21
Fgf5	Fgf5.bSep07	1	0	0	1	,6
Fgfr1	Fgfr1.jSep07	1	1	,4	1	,34
Fgfr1	Fgfr1.fSep07	1	0	0	1	,2
Gab1	Gab1.aSep07	2	1	,6	2	,2,1
Gab1	Gab1.bSep07	1	1	,84	1	,14
Grb2	Grb2.bSep07	1	1	,252	1	,98
Grb2	Grb2.dSep07	1	1	,5	1	,1
Hras1	Hras1.eSep07	2	2	,1,1	1	,1
Hras1	Hras1.dSep07	2	1	,3	1	,3
Itpr1	Itpr1.aSep07	14	12	,1,10,5,27,1,3,9,7,6,5,2,1	11	,16,6,77,5,14,3,18,3,7,7,6
Itpr1	Itpr1.dSep07	1	1	,15	1	,19
Itpr1	Itpr1.iSep07	1	1	,1	1	,2
Itpr1	Itpr1.eSep07	1	1	,3	0	0
Itpr1	Itpr1.gSep07	1	0	0	1	,1
Map3k1	Map3k1.aSep07	12	11	,2,5,2,3,1,2,1,4,16,2,8	11	,5,10,2,2,1,3,3,4,29,4,9
Map3k1	Map3k1.bSep07	0	0	0	1	,1
Mapk1	Mapk1.dSep07	1	1	,2	1	,2
Mapk1	Mapk1.aSep07	1	1	,1	1	,5
Mapk14	Mapk14.bSep07	1	2	,1,1	1	,2
Mapk14	Mapk14.cSep07	1	1	,1	1	,1
Pik3c2a	Pik3c2a.aSep07	3	3	,14,8,5	3	,21,6,12
Pik3c2a	Pik3c2a.bSep07	1	1	,3	2	,3,2
Pik3c2a	Pik3c2a.eSep07	1	1	,6	1	,3

Pik3r1	Pik3r1.aSep07	9	5	,1,1,3,4,4	8	,2,9,1,14,8,3,6,5
Pik3r1	Pik3r1.bSep07	1	0	0	1	,14
Pik3r3	Pik3r3.aSep07	1	1	,1	1	,3
Pik3r3	Pik3r3.cSep07	1	0	0	1	,1
Raf1	Raf1.aSep07	1	1	,5	1	,5
Raf1	Raf1.jSep07	1	0	0	1	,1
Stat3	Stat3.bSep07	1	1	,3	1	,4
Stat3	Stat3.cSep07	1	0	0	1	,1

Table 14d. The Mouse Regulatory network.

Gene Name	Ace Name	Number of total Diagnostic junctions per variant	Number of Diagnostic junctions present in ES	Counts per Junction - ES	Number of Diagnostic junctions present in EB	Counts per Junction - EB
Ahctf1	Ahctf1.aSep07	25	24	,88,2,7,8,1,20,25,113,15,10,9,2,17,15,16,3,8,2,116,15,17,117,9,9	22	,63,1,1,14,15,24,25,6,5,9,3,18,19,23,11,2,75,13,12,51,8,12
Ahctf1	Ahctf1.fSep07	1	1	,6	1	,11
Ahctf1	Ahctf1.dSep07	1	1	,1	0	0
Ahctf1	Ahctf1.hSep07	2	0	0	1	,1
Arid3a	Arid3a.bSep07	1	1	,12	1	,33
Arid3a	Arid3a.dSep07	1	1	,2	1	,1
Arid3a	Arid3a.cSep07	1	0	0	1	,1
Arid3bandClk3	Arid3bandClk3.kSep07	2	2	,1,1	2	,1,3
Arid3bandClk3	Arid3bandClk3.cSep07	1	1	,4	1	,16
Arid3bandClk3	Arid3bandClk3.sSep07	2	0	0	2	,2,2
Arid3bandClk3	Arid3bandClk3.gSep07	0	0	0	1	,1
Arid3bandClk3	Arid3bandClk3.mSep07	1	0	0	1	,2
Btbd14a	Btbd14a.dSep07	4	0	0	1	,1
Btbd14b	Btbd14b.bSep07	2	1	,1	0	0
Btbd14b	Btbd14b.dSep07	1	1	,1	0	0
Cdc2a	Cdc2a.aSep07	1	1	,114	1	,281
Cdc2a	Cdc2a.bSep07	1	1	,1	1	,8
Dpf2	Dpf2.aSep07	2	2	,13,7	2	,29,15
Gatad2b	Gatad2b.dSep07	1	1	,16	1	,30
Gatad2b	Gatad2b.aSep07	1	1	,2	0	0
Hdac2	Hdac2.aSep07	1	1	,9	1	,34
Hdac2	Hdac2.eSep07	2	1	,3	1	,6
Hdac2	Hdac2.cSep07	1	0	0	1	,1

Itga1andPelo	Itga1andPelo.aSep07	8	3	,1,1,2	7	,2,2,1,4,5,1,2
Itga1andPelo	Itga1andPelo.lSep07	0	1	,1	0	0
Itga1andPelo	Itga1andPelo.cSep07	2	0	0	1	,2
Itga1andPelo	Itga1andPelo.iSep07	2	0	0	1	,1
Mybbp1a	Mybbp1a.dSep07	1	1	,1	0	0
Nanog	Nanog.aSep07	1	1	,77	1	,53
Nanog	Nanog.cSep07	1	1	,1	0	0
Nr0b1	Nr0b1.aSep07	1	1	,1	0	0
Pou5f1	Pou5f1.aSep07	4	4	,486,367,1699,1300	4	,207,173,687,909
Rai14	Rai14.cSep07	2	2	,9,1	1	,4
Rai14	Rai14.fSep07	1	0	0	1	,2
Rai14	Rai14.hSep07	1	0	0	1	,1
Rif1	Rif1.bSep07	1	1	,72	1	,10
Rif1	Rif1.iSep07	2	1	,1	0	0
Rif1	Rif1.eSep07	1	1	,1	0	0
Rif1	Rif1.nSep07	1	0	0	1	,1
Rnf2and1200016B10Rik	Rnf2and1200016B10Rik.eSep07	2	2	,1,13	1	,6
Rnf2and1200016B10Rik	Rnf2and1200016B10Rik.dSep07	1	1	,4	1	,15
Rnf2and1200016B10Rik	Rnf2and1200016B10Rik.hSep07	1	1	,1	0	0
Rnf2and1200016B10Rik	Rnf2and1200016B10Rik.nSep07	1	1	,3	0	0
Rnf2and1200016B10Rik	Rnf2and1200016B10Rik.aSep07	1	0	0	1	,5
Rybp	Rybp.aSep07	2	1	,18	1	,20
Sall1	Sall1.aSep07	1	1	,6	2	,6,1
Sall1	Sall1.cSep07	0	0	0	1	,1
Sall4	Sall4.cSep07	2	1	,1	2	,1,1
Sall4	Sall4.aSep07	1	1	,16	1	,14
Sall4	Sall4.bSep07	1	0	0	1	,2
Sall4	Sall4.eSep07	1	0	0	1	,1
Smarcad1	Smarcad1.eSep07	3	1	,2	2	,3,1
Smarcad1	Smarcad1.lSep07	1	1	,2	1	,20
Smarcad1	Smarcad1.nSep07	1	0	0	1	,8
Smarcad1	Smarcad1.mSep07	2	0	0	1	,1
Smarcc1	Smarcc1.aSep07	1	1	,6	1	,8
Smarcc1	Smarcc1.fSep07	1	1	,1	0	0
Smarcc1	Smarcc1.hSep07	1	1	,1	0	0
Smarcc1	Smarcc1.iSep07	1	0	0	1	,1
Trim28	Trim28.aSep07	1	1	,54	1	,65
Wapal	Wapal.aSep07	3	3	,21,31,11	3	,19,37,24
Wapal	Wapal.cSep07	2	1	,35	1	,30
Wapal	Wapal.bSep07	1	1	,4	0	0

Wapal	Wapal.eSep07	1	1	,1	0	0
Wdr18	Wdr18.aSep07	2	2	,75,34	2	,159,66
Yy1	Yy1.aSep07	1	1	,31	1	,48
Zfp219	Zfp219.eSep07	1	1	,1	1	,2
Zfp219	Zfp219.aSep07	1	1	,1	1	,4
Zfp219	Zfp219.dSep07	1	1	,2	1	,1
Zfp219	Zfp219.cSep07	1	1	,1	0	0
Zfp219	Zfp219.kSep07	1	0	0	1	,1
Zfp281	Zfp281.bSep07	1	1	,1	0	0
Zfp281	Zfp281.aSep07	0	0	0	1	,1
Zfp42	Zfp42.eSep07	3	2	,73,3	1	,2
Zfp42	Zfp42.gSep07	2	1	,101	1	,2
Zfp609	Zfp609.aSep07	4	2	,12,1	3	,14,1,7
Zmym2	Zmym2.aSep07	6	6	,3,60,6,14,23,12	6	,5,51,14,27,63,26
Zmym2	Zmym2.eSep07	1	1	,2	1	,4
Zmym2	Zmym2.iSep07	2	1	,1	1	,1

Supplementary Table 15. Quantitative RT-PCR results. (a) Tablulated results of quantitative real-time PCR results presented in Figure 3. (b) Isoform specific identification by SQRL and semi-quantitative real-time PCR. Primers were designed to amplify a set of transcripts from loci containing multiple isoforms as identified by SQRL, where one of the primers spans the diagnostic junction. ES and EB counts are the number of tags assigned to the specific junction. Real time is the average cycle number (n=4) of real time PCR with Forward and Reverse primers. ND is not detected.

Table 15a.

Genbank	Common Name	ES cycle number	ES SOLiD tags	EB cycle number	EB SOLiD tags
NM_011949	Mapk1	22.13764	5310.254	21.63845	7347.606
NM_010937	Nras	22.96125	2642.121	22.81491	5019.941
NM_008228	Hdac1	22.57395	845.4121	22.3401	1174.684
NM_001079908	Fgfr1	23.2706	2451.191	21.69304	4816.217
NM_175168	Ptk7	23.64395	2357.776	22.48915	5463.843
NM_001079908	Fgfr1	23.67074	2451.191	22.22619	4816.217
NM_001079908	Fgfr1	23.68026	2451.191	21.68276	4816.217
NM_008927	Map2k1	23.71356	3476.775	21.90571	4294.512
NM_011952	Mapk3	23.91087	3287.23	22.59715	7637.099
NM_001025432	Crebbp	23.51923	1055.223	22.97768	1990.331
NM_175168	Ptk7	23.73324	2357.776	21.69664	5463.843
NM_009231	Sos1	23.98299	638.078	26.23452	392.2004
NM_023138	Map2k2	24.42755	2578.704	23.34689	4415.79
NM_029780	Raf1	24.53259	2620.086	23.97024	3899.741
NM_001042607	Ryk	24.67893	2965.917	23.49197	3556.81

NM_021501	Pias4	24.22248	1311.983	24.10765	1335.394
NM_009101	Rras	24.41513	2630.336	22.52206	4044.436
NM_016961	Mapk9	24.63528	2106.987	24.03271	1697.658
NM_172688	Map3k7	24.69545	1152.507	24.17391	1710.166
NM_009758	Bmpr1a	25.02532	1200.138	23.58968	2285.31
NM_019392	Tyro3	25.04825	665.6785	23.43235	1444.645
NM_007397	Acvr2b	25.17729	375.8531	24.1661	570.5452
NM_008284	Hras1	25.62532	4194.521	24.86961	7027.144
NM_025481	Smurf2	25.20405	692.5019	24.2972	832.7369
NM_009758	Bmpr1a	25.8889	1200.138	25.49339	2285.31
NM_001038627	Smurf1	25.58711	1024.889	24.90417	1180.154
NM_011577	Tgfb1	25.64276	177.3725	22.55482	1378.419
NM_007561	Bmpr2	25.65387	1084.54	24.53502	1249.418
NM_021284	Kras	25.73224	1887.186	24.99845	2171.217
NM_007584	Ddr1	26.22968	1000.135	27.96704	994.373
NM_025846	Rras2	26.29756	2632.496	25.06322	4166.659
NM_009370	Tgfbr1	26.49044	673.4312	24.50842	1205.12
NM_008541	Smad5	26.69509	563.1713	25.58061	1053.789
NM_009370	Tgfbr1	26.88518	673.4312	26.02576	1205.12
NM_010207	Fgfr2	27.10811	71.84449	25.40933	364.5034
NM_001110204	Acvr1	27.27563	120.6195	25.7714	308.8941
NM_009465	Axl	27.30406	783.9535	28.53082	215.0044
NM_016769	Smad3	27.3512	100.0847	25.10473	199.7649
NM_008587	Mertk	27.79649	59.37528	28.56632	37.44862
NM_007396	Acvr2a	27.41199	421.5695	25.80305	1033.823
NM_008871	Serpine1	27.56447	3207.413	27.74486	241.2575
NM_007395	Acvr1b	27.5534	541.3111	26.43666	1267.634
NM_007557	Bmp7	27.57368	355.1296	24.18815	2050.13

NM_009465	Axl	27.83104	783.9535	26.81194	215.0044
NM_008624	Mras	27.90132	796.8422	29.1894	180.883
NM_008011	Fgfr4	27.91065	57.41736	25.901	63.16913
NM_013690	Tek	27.9951	250.7572	30.51207	218.0649
NM_009371	Tgfbr2	28.09224	88.9194	28.64944	106.278
NM_007560	Bmpr1b	28.72691	59.7167	29.82869	40.49754
NM_009368	Tgfb3	28.79632	69.84049	29.4501	27.49605
NM_008279	Map4k1	28.5731	252.7605	27.92812	180.2125
NM_008380	Inhba	29.36635	139.3778	31.68163	61.83224
NM_013846	Ror2	28.94738	415.6349	27.32325	782.9453
NM_013845	Ror1	28.99812	55.58212	26.61767	203.7274
NM_025609	Map3k7ip1	29.01361	373.8021	27.86141	850.8656
NM_008261	Hnf4a	29.34712	48.43486	28.06075	111.1253
NM_008809	Pdgfrb	29.02357	330.789	27.10651	1949.803
NM_010234	Fos	29.41763	168.7299	27.73115	328.1763
NM_008542	Smad6	29.61984	120.8769	25.57063	1316.565
NM_207655	Egfr	29.74992	496.7688	27.92828	106.0937
NM_144547	Amhr2	29.93275	120.592	28.28257	330.6112
NM_011098	Pitx2	30.14793	470.5871	27.75162	1418.153
NM_010565	Inhbc	30.76214	46.94626	32.40724	10.3788
NM_203345	Ltk	30.61591	58.64773	30.07558	43.85042
NM_008010	Fgfr3	30.35706	43.34709	27.59249	364.4558
NM_009371	Tgfbr2	30.5693	88.9194	28.11753	106.278
NM_019732	Runx3	31.0753	40.68603	30.57396	56.52883
NM_011058	Pdgfra	30.83449	26.9458	25.4203	252.3289
NM_022563	Ddr2	31.54436	50.43414	36.1391	16.75731
NM_011587	Tie1	31.39734	31.39377	29.13842	57.20111
NM_009504	Vdr	32.03512	29.3673	33.36767	20.82744

NM_008700	Nkx2-5	32.1891	17.31292	31.6074	24.05443
NM_009367	Tgfb2	31.90848	678.9178	29.60443	1272.238
NM_011058	Pdgfra	32.12296	26.9458	26.15645	252.3289
NM_009820	Runx2	34.52617	35.12999	32.47466	21.54571
NM_007439	Alk	34.606	18.86252	31.8986	23.79985
NM_008380	Inhba	34.90037	139.3778	35.44247	61.83224

Table 15b.

Transcript	Aceview	Genomic coordinates	ES junction count	EB junction count	RT ES (cycle)	RT EB (cycle)	No RT (cycle)
Akt3	Akt3.aSep07	chr1_179033236_179039757	166	329	27.00703	25.90487	34.6682
Akt3	Akt3.bSep07	chr1_179039608_179039757	2	15	32.51142	31.60121	34.4815
Akt3	Akt3.bSep07	chr1_179033236_179039549	1	5	30.90337	30.29073	ND
Axin1	Axin1.bSep07	chr17_26327175_26330883	24	40	26.77519	27.60964	ND
Axin1	Axin1.aSep07	chr17_26329567_26330883	14	17	28.83183	29.51696	31.119
Axin1	Axin1.aSep07	chr17_26327175_26329460	1	2	28.00483	29.00448	ND
Axin2	Axin2.aSep07	chr11_108781874_108784465	7	3	ND	ND	ND
Axin2	Axin2.bSep07	chr11_108782506_108784465	2	15	29.08967	27.23876	29.7396
Crebbp	Crebbp.aSep07	chr16_4101637_4107427	1	3	25.84427	25.61635	27.7254
Crebbp	Crebbp.bSep07	chr16_4101667_4107427	0	6	28.11053	28.03597	ND
Csnk1a1	Csnk1a1.aSep07	chr18_61729215_61731527	51	53	24.86645	25.42199	ND

Csnk1a1	Csnk1a1.bSep07	chr18_61740140_61747079	36	98	24.72105	25.42184	ND
Csnk1a1	Csnk1a1.dSep07	chr18_61740104_61747079	19	40	23.94017	ND	ND
Csnk1a1	Csnk1a1.kSep07	chr18_61744057_61744818	16	13	28.10126	29.03959	32.0394
Csnk1a1	Csnk1a1.aSep07	chr18_61740104_61741689	12	22	34.23382	34.63918	ND
Csnk1a1	Csnk1a1.bSep07	chr18_61729215_61735063	4	2	23.60193	24.2035	26.6826
Csnk1g1	Csnk1g1.cSep07	chr9_65756974_65768810	5	0	28.07337	27.7596	34.3804
Csnk1g1	Csnk1g1.aSep07	chr9_65756974_65797822	2	1	29.01295	29.23306	ND
Mark2	Mark2.aSep07	chr19_7352575_7355487	8	22	33.13691	33.66124	35.1829
Mark2	Mark2.eSep07	chr19_7352575_7353600	3	10	31.14151	32.06858	ND
Mark2	Mark2.cSep07	chr19_7352575_7354475	0	4	35.61735	36.99471	ND
Myc	Myc.bSep07	chr15_61817505_61819061	9	23	27.20391	25.83068	37.4391
Myc	Myc.cSep07	chr15_61817505_61819064	7	28	27.87136	26.91339	34.3312
Myc	Myc.dSep07	chr15_61818857_61819061	1	6	34.12246	33.06245	ND
Pik3c2a	Pik3c2a.bSep07	chr7_123586191_123586817	3	2	ND	ND	ND
Pik3c2a	Pik3c2a.bSep07	chr7_123562094_123586104	0	3	27.83701	27.87773	ND
Pik3c2a	Pik3c2a.aSep07	chr7_123562094_123586817	24	8	25.9582	26.50988	ND
Ppp2r4	Ppp2r4.aSep07	chr2_30295283_30298809	6	15	23.24776	25.16518	ND
Ppp2r4	Ppp2r4.eSep07	chr2_30298916_30301896	5	5	24.27214	25.09736	ND
Ppp2r4	Ppp2r4.hSep07	chr2_30297497_30298809	0	1	36.36248	36.04673	ND
Ppp2r4	Ppp2r4.lSep07	chr2_30298525_30298809	0	2	ND	ND	ND
Ppp2r5d	Ppp2r5d.aSep07	chr17_46821325_46821454	15	44	24.85686	25.30659	ND
Tgfb2	Tgfb2.aSep07	chr1_188473776_188514599	11	5	34.19441	33.19714	ND
Tgfb2	Tgfb2.bSep07	chr1_188473776_188528307	4	1	28.34279	26.32685	ND
Zfp42	Zfp42.bSep07	chr8_44389912_44392194	561	14	21.09935	27.54184	32.9746
Zfp42	Zfp42.eSep07	chr8_44389912_44391033	73	2	26.87531	31.76694	ND
Zfp42	Zfp42.eSep07	chr8_44391295_44392194	3	0	28.03289	33.80009	ND

Supplementary Table 16. Primers used for experimental validation in this study, including quantitative real-time PCR, semi-quantitative real-time PCR, and amplification of genomic DNA for SNP validation.

Reference ID	Gene Name	Forward	Reverse
6519	Mapk1	ACATGAGAACATCATTGGCATCA	AAAGGTCCGTCTCCATGAGGT
11499	Nras	AAGTCCAAAAGCCTCCCGAG	CAACACCACCTGCTCCAACC
12641	Hdac1	TCGCTGCTGGACTTACGAAA	GCTTGAAATCCGGTCCAAAG
17205	Fgfr1	CATCGGGCTGGATAAGGACA	CCGAGATCAGATCCGACAGG
7680	Ptk7	CACCCAAGCCCACTGTGATC	CACACTATTGATGCGCAAGGTC
17205	Fgfr1	CATCGGGCTGGATAAGGACA	CCGAGATCAGATCCGACAGG
17205	Fgfr1	CATCGGGCTGGATAAGGACA	CCGAGATCAGATCCGACAGG
18606	Map2k1	GTGCAACTCCCCGTACATCG	CTTGATCCAAGGACCCACCA
16742	Mapk3	ACAGTAGAGGAAGCGCTGGC	TCCAGCTCCATGTCGAAGGT
6423	Crebbp	CTGGCAGACCTCGGAAAGAA	CTGGCGCCGCAAAAACT
7680	Ptk7	CACCCAAGCCCACTGTGATC	CACACTATTGATGCGCAAGGTC
7894	Sos1	GCACCCTAAGCCTCTTCCG	GGATGTCTCATGGTTCCTGGTC
21931	Map2k2	GCCTACCTCCGAGAGAAGCA	ACCCCGAAGTCACACAGCTTA
14924	Raf1	TACTCGTACGGCATCGTGCT	TACCCACGGCCTACCATGAA
18840	Ryk	AGCTTCGAGGTCTGCACCAC	CCCCCAATTCATGTATGGCA
21932	Pias4	CCAATACCCACCCAACATCG	AGGGCCTCTTAGGTTCCACAC
15960	Rras	GTCGATGAGGCATTTGAGCA	CCTCCATCCTTCTTCCTGGG
2325	Mapk9	CTCTGGAGCCCAAGGAATTGT	CGTGCGTTTGGTTCTGAAAA
11916	Map3k7	GTGATAACACGCCGGAAACC	AAATTTTTGATCAGTGGTGGTCG
5153	Bmpr1a	TCATTGAAAGACCTGATTGACCA	CGAACCATCTGAATCTGTTTGG
10199	Tyro3	GAGTTTGGATCAGTGCGGGA	TCGCTTGAGGCAATGATGTC
19055	Acvr2b	AGTTCATTGCTGCCGAGAAAC	CCCCTTGAGGTAATCCGTGA

16944	Hras1	GCAGCTATGGCATCCCCTAC	CAATTTATGCTGCCGAATCTCA
3380	Smurf2	ACTCCTCCAGACCTACCGGAA	GGATCATGC CATGTGCTCAC
5153	Bmpr1a	TCATTGAAAGACCTGATTGACCA	CGAACCATCTGAATCTGTTTGG
14141	Smurf1	CTTGCTGAAGCCCTTTGACC	AGTGTTTCAGCCGGGTGTTG
15648	Tgfb1	GCCTGAGTGGCTGTCTTTTGA	CACAAGAGCAGTGAGCGCTG
258	Bmpr2	ACTTTCCACCCCCTGACACA	AACAGCTAACACAGAAACTGATGCC
15259	Kras	AGCGCCTTGACGATACAGCT	GTTTCTCCATCAATTACTACTTGTTTCCT
7521	Ddr1	CTGGAGAACAAGGCCACTCAG	GCGATCTGGGCCCCC
16629	Rras2	TGTGTGATCGATGACCGAGC	CCCTCGCCTGTCCTCATGTA
12079	Tgfbr1	TTTCAGAGGGCACCACCTTAA	GGTCCTGGCAATTGTTCTTTG
4629	Smad5	AGTCAACCATGGATTCGAGGC	CGTCCTGTCGGTGGTACTCTG
12079	Tgfbr1	TTTCAGAGGGCACCACCTTAA	GGTCCTGGCAATTGTTCTTTG
16824	Fgfr2	TTTAAGCTGCTCAAAGAGGGACA	TGTGAGGGTACAGCATGCCA
22121	Acvr1	CTCCGGTCTTCCTTTCCTGG	AGCTGCCCCTCCATACTTCTC
15652	Axl	TGTTCCTTGTCCATCGGAGG	CTTTCGGACACGGTACCTGAC
18601	Smad3	CATCACCACGCAGAACGTG	AGATAACGTGAGGGAGCCCC
10314	Mertk	CCGTGTTAATGAAAAACCGGA	GTCCTTTGTCATTGTGGGCC
22512	Acvr2a	CTGGCAAGTCTGCAGGTGAC	AAATGCGTCCCTTTGGAAGTT
14002	Serpine1	CACCGTCTCTGTGCCCATG	GGTAGGGCAGTTCCACAACG
6328	Acvr1b	AGACGCTCCAGGATCTCGTC	TTGTAAAACAATGGTTCGGGC
10778	Bmp7	CAGCCACTTCCTCACTGACG	GGAACTCCCGATGGTGGTATC
15652	Axl	TGTTCCTTGTCCATCGGAGG	CTTTCGGACACGGTACCTGAC
18823	Mras	TCTACTCCGTCACCGACAAGG	CACGAGGATCATTGGGAATGA
4593	Fgfr4	CCATCCACTGGCTCAAGGA	CACACTTTCCATCACCAGGCT
12287	Tek	GGGTGTTCCTGTGCCACAG	TTGGTACAGTGGCACCTGAGC
19036	Tgfbr2	GCCTGTAACATGGAAGAGTGCA	ACACCCGTCACTTGGATAATGAC
11751	Bmpr1b	TCTCAGAGCTCGGGAAGTGG	CCATAGCGGCCTTTTCCAAT
3979	tgfb3	CCAAGCGCACAGAGCAGA	CCTTGTGGGCAGATTCTTGC

15739	Map4k1	CGCCTTTCTTTCATTGGGAC	CCAAGAGACCAGATGTCGCA
4249	Inhba	GAACTCATGGAGCAGACCTCG	GACAGGTCACTGCCTTCCTTG
4566	Ror2	TGTGGCTACCAACGGGCT	GATCGTCATCCTGAAAATTGTGG
12316	Ror1	AGAGCCCGGAAGCTGCA	CCGGTAGTCTACACCCGTGC
6069	Map3k7ip1	CCAGATGGTCAACGGCTCTC	GCGTGTTGGTGGACTGCAG
10649	Hnf4a	GTTACTGCAGGCTTAAGAAGTGCTT	GTCCTCGTAGCTTGACCTCCG
8270	Pdgfrb	AGCTCACGGTCT GAGCCATT	GCTCGGACATTAAGGCTTGC
3973	Fos	TGGTGAAGACCGTGTCAGGA	CCTTCGGATTCTCCGTTTCTC
18602	Smad6	GCTGTCTCCTCCTGACCAGTACA	CACCCGGAGCAGTGATGAG
2137	Egfr	GTGGAGGGACATCGTCCAAA	CATTGGGACAGCTTGGATCA
6377	Amhr2	GCTGCTTTGGGATCTGGAATC	GTCACAGTGGAGGGACTCACAG
11682	Pitx2	CCTTACGGAAGCCCGAGTC	CAAAGCCATTCTTGCACAGC
1899	Inhbc	GGAAACCCTGTTGGAGCATG	GGAACTCGAGCCGGGTCT
22224	Ltk	CAAGTGGATGCCGCCAG	ACCCCAGTGAGAAGATCTCCC
13243	Fgfr3	AGGCCCACCTTCAAGCAGTT	GAGTACTGCTCAAACGGCACG
19036	Tgfbr2	GCCTGTAACATGGAAGAGTGCA	ACACCCGTCACTTGGATAATGAC
12742	Runx3	CAACGCTTCCGCTGTCATG	CGGTGATTGTGAGCGTGAAA
13425	Pdgfra	GACCCCATGCAGTTGCCTTA	TCGAC CACTTTCCCAAATGC
21673	Ddr2	TCCTGGGTTCTGCAAAAGCT	ACTTGAGGCTGTGATGTCCTCA
12487	Tie1	GCCATGATCAAGAAGGACGG	TCTAGTTCACCTGCAAAGTCTCGA
6240	Vdr	AAGGACAACCGGCGACACT	TTACGCTGCACCTCCTCATCT
7287	Nkx2-5	CCTACGGTGACCCTGACCC	GCCATCCGTCTCGGCTTT
1078	Tgfb2	CGAGGAGTACTACGCCAAGGA	CCAATGAGCCAGAGGGTGTT
13425	Pdgfra	GACCCCATGCAGTTGCCTTA	TCGAC CACTTTCCCAAATGC
20368	Runx2	TAAAGTGACAGTGGACGGTCCC	GCCCTAAATCACTGAGGCGA
7851	Alk	CGGTGGCAAGTGAGGACC	TCACTGTTGCACCTTCAGAAGC
4249	Inhba	GAACTCATGGAGCAGACCTCG	GACAGGTCACTGCCTTCCTTG
11682	Pitx2	CCTTACGGAAGCCCGAGTC	CAAAGCCATTCTTGCACAGC

Akt3.aSep07	Akt3	TTCCGTCCACTCTTCTCTTTCC	ATGTCTTCAGTGGACCACTGTTATAGA
Akt3.bSep07	Akt3	TTTATCATTTTCCAGCCCTTTCC	TGTCTTCAGTGGACCACTGTTATAGAG
Akt3.bSep07	Akt3	CTCCTCCTCTTGCCTCTGCA	TCTTCTGAGAAAACGAGAAGAGTGG
Axin1.bSep07	Axin1	CCGCAGGCGTTTGGAAG	CCTTTCTTTGTGATTTTGTCCTCTG
Axin1.aSep07	Axin1	TTGGAACTCTCCGAGACAGAGAC	CAATGCTGTCACATGGTGGTG
Axin1.aSep07	Axin1	CAGCTGGAAGAGGCCCG	TGACTGCCTGCACATACCTCTG
Axin2.aSep07	Axin2	AAGCAGCCGTTCGCGAT	CATGTGAGCCTCCTCTCTTTTACAG
Axin2.bSep07	Axin2	CGCCAGCGGATCAATGA	TTCTTCCAGTTCCTCTCAGCAAT
Crebbp.aSep07	Crebbp	TGAAACACTTCTCACAGAAATGATACCT	CAGCTGTGTACAATTCCTCGTGAT
Crebbp.bSep07	Crebbp	GAAGGCCATACTTGGGTGAACT	GGAAAGCAGCTGTGTACAATTCCT
Csnk1a1.aSep07	Csnk1a1	GCATACAAAGAATTTTATACACAGAGACATT	CCACTGGAGATTCTAAACACTTATTACAG
Csnk1a1.bSep07	Csnk1a1	GACAAAACCAAGAGTAACATGAAAGGT	GCTCCAATCGTCTGCTCTGC
Csnk1a1.dSep07	Csnk1a1	CCAGCAGGCAGCCTCTTC	TTCCACAATTCATGCTTAGAAACCT
Csnk1a1.kSep07	Csnk1a1	AGGATTCATTCATAACAGGAACAGAGA	TCCAGAGCGCATCTTCACTG
Csnk1a1.aSep07	Csnk1a1	CCCAAACCCCCACAGGTATT	TCACTTTTCTTTTTCAAGGAAACATACT
Csnk1a1.bSep07	Csnk1a1	TTGGGCGTCACTGTAATAAGTTATTC	TTGCCTTGTCCTGTTGTCTCTG
Csnk1g1.cSep07	Csnk1g1	GCACCGCTGATTTAGGGATACA	TGTTACTCTACCTTTGAAGTCTTGTTTCTT
Csnk1g1.aSep07	Csnk1g1	GCACCGCTGATTTAGGAGGA	TCACTTACTCACCAGTTAAGCCCC
Mark2.aSep07	Mark2	AGGAGGCTGTGGCAGGC	CAAGTTTGTCCGCAGGAACC
Mark2.eSep07	Mark2	CCACTCGGTCTTTGCTTTCAG	TCCCTTCCTGGGACAGAGC
Mark2.cSep07	Mark2	AGGAGGCTGTGGCAGGC	TTTCAGGTTTGCCAGAAGGAAC
Myc.bSep07	Myc	CTCCTGAAAAGAGCTCCTCGAG	GTCGTGGCTGTCTGCGG
Myc.cSep07	Myc	CGTTGGAAACCCCGACAG	TCGTAGTCGAGGTCATAGTTCCTG
Myc.dSep07	Myc	TAAGTCCCTGCTCGAAGGAGG	CGTCGTGGCTGTCTGTATCAGT
Pik3c2a.bSep07	Pic	TGAATTCAAACACCCACTGGC	CCTTGGAGGGCGCGA
Pik3c2a.bSep07	Pic	CCAAAATACCAGGACCTGGAAA	GAAGGAAGCATCCCTGTTTGAA
Pik3c2a.aSep07	pic2a2a	TCCAAAATACCAGGACCTCACG	GGACCTTGGAGGGCGC
Ppp2r4.aSep07	Ppp2r4	AGGCGGTGAGCGAGAACC	GGGCCAGTCTTCATCTCAGTG

Ppp2r4.eSep07	Ppp2r4	GTGCTGTCCCCTCCTGGTC	GGGAACTTCTCCAGGCACTCT
Ppp2r4.hSep07	Ppp2r4	GGAACGCAGCGTCTATGAAGA	CCCTGGTTCACTTTAGACCAGG
Ppp2r4.lSep07	Ppp2r4	CCTTTCCTGTCCTCCAGATGAA	GGGACAGCACTGATGTTCCAC
Ppp2r5d.aSep07	Ppp2r5d	GTGGGCGTCTCTGTCTCCAT	GGCTCAATCCCCAGTACCCT
Tgfb2.aSep07	Tgfb2	TGAGACATCAAAGCGGACGAT	GCTCTGTGGGTACCTTGATGC
Tgfb2.bSep07	Tgfb2	TGAGACATCAAAGCGGACGA	CCACCTCCCCTCCGAAAAT
Zfp42.bSep07	zfp42	AGCTCTTAGTCCATTTCTCTAATGCC	TCAGAAAGGAAACCAAGGAGGA
Zfp42.eSep07	Zfp42	CCAGCTCTTAGTCCATTTCTCACA	GAGAAAGGATCCCTGTCTCGAA
Zfp42.eSep07	Zfp42	AGAAAAGTTGGTCGGTAGTATGGC	CATCGCTGTGGGCATTAGAA
SNP validation	EG668668	GGTCCTGACTTTCCCATCAA	TAAGCTGACTGGCGTGTTTG
SNP validation	Uchl4	ATGCCATTGCAAACAACAAA	GGTGCCTCAGTCTGACCTTC
SNP validation	Car4	GAAGCAATGAGCAGGTAGGC	GCAGGTCTCCTCCGATAATG
SNP validation	S100a11	TGCCATGATTCTTTCATCCA	CAGCTCCTGCTACCAGCTTC
SNP validation	Nusap1	TAGGAAGCAGTCGTCAGTGC	CATTCACCTGGGAGGTAGGA
SNP validation	Lars2	CGATCCTTCTGACCTTTTGG	ACGCTTGGTGAATTCTGCTT
SNP validation	Rasl2-9	AGGGGGCTAAATACGCAAGT	GACTGCTCTCCCAGATGAGG
SNP validation	Zkscan1	TCTAGAGCCAGGTTGGCAGT	CCCCTCACTGGTTCACTTGT
SNP validation	Tagln2	TGGGCTGTACTTACCACACG	AGCCAGAATCCATGCAAAAC
SNP validation	Samd8	AACCCTGCCGTTCACAATAG	TGACGAGAAACCCAAGGAAG
SNP validation	Imp4	GTCTTTGTCACCAGCAGCAA	TAACCAGCCATAGGGCACTC
SNP validation	Pgam1	CTCAGGGCAAGGTGAAGAAG	TGGGTGCAGGGGATAAAATA
SNP validation	Arid3b	TCATCCAGACTCCAGCTCGG	AGAGGCCCAGGCTGCTTC
SNP validation	C80913	AGCACCTTTCAAGTTCACCCA	CAGTCATGGCATTTCCTTTTTAAAC
SNP validation	Dnttip2	GTGTGTTCAGAGGACTCAGATGCT	AAGAGACACAACATTGGCTCGTT
SNP validation	Nbn	CCAAGCCTTTCCCAAGTCCT	CGCTCCATTCCATGCTTTGT
SNP validation	Rnf14	TCTCCCATTGTGTGCATTCTTT	AAATTCTGTGGCAGATCCAAATAGA

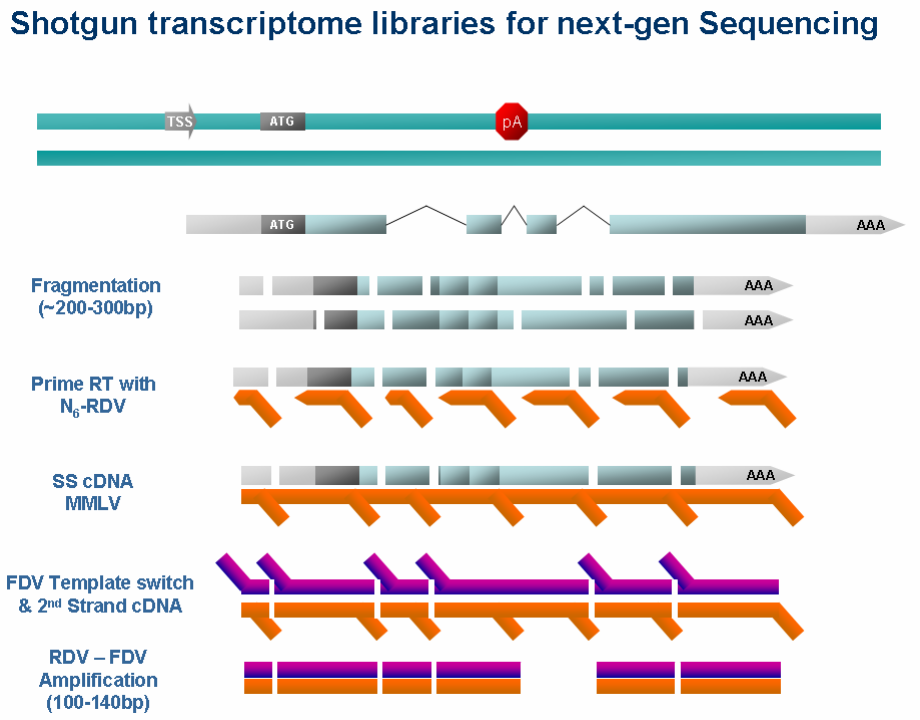
Supplementary Methods.

Expression Genomics Laboratory SQRL protocol

Background: The primary aim for developing this protocol is to massively shot gun sequence the transcriptome. The method was designed to meet the following essential criteria:

1) It has to be random primed to ensure transcriptome coverage.
2) It has to work of off a reasonable amount of RNA (deemed 50-100ug of total RNA).
3) It has to be directional (ie not lose strandedness).
4) It has to generate a library of small linear PCR ready products (<200bp) with different sequences at either end compatible with whatever sequencing tech is to be applied.

Several methods were applied to this approach. The most successful by far was the template switch-mediated library generation off fragmented RNA. The protocol has been applied primarily to mRNA but is equally amenable to studying non-adenylated material (underway).



The library method commences with RNA fragmentation followed by first strand cDNA synthesis which is primed using a tagged random hexamer. A modified MMLV Reverse Transcriptase is used in the cDNA synthesis to allow for non-templated Cs to added termini of the completed 1st strand cDNAs. A tagged ribo-G primer is then used to prime

off the non-template C tail to generate double stranded cDNAs (this process is known as template switching). The library is them PCR amplified using primers complementary to 5' and 3' tag sequences and size selected via electrophoresis to obtain optimally sized amplimer. ePCR and sequencing. Finally, samples are size selected by gel electrophoresis, and QCed before sequencing. The sequencing is carried out from the random hexamer to maximize random distribution of tags.

1. RNA preparation:

Notes: The work to date has focused on the polyA transcriptome. A series of different RNA purification methods, and different combinations of depletion and RNA preps were tested :-

*1 round of Oligotex polyA is insufficient (approx 10% ribosomal contamination, mainly 18s). *1 round of Ribominus is insufficient (approx 20% ribosomal contain, mainly 28s, also 5s and fragments)

* Single rounds of ployA prep followed by depletion are better than 2 rounds of polyA purification.

*Reversing the order of depletion and polyA preparation provided a similar quality product with a lower yield.

Briefly, we prepare total RNA via RNAeasy spin column purification exactly following the manufacturer's protocol (URL for protocol is:). The amount of RNA prepared has varied depending on the starting material. In every case, total RNA should be quantified by nanodrop spectrophotometry and checked for integrity and quality by bioanalysis prior to continuing to the next step.

50- 300ug of total RNA is then subjected to a single round of polyA selection using Oligotex mRNA Mni Kit. (Qiagen: Cat ID70042). using the spin column method. 100ug of total RNA generally yields sufficient material for 2 library preparations. (text below is taken directly from QIAGEN's protocol PDF with our comments added in RED.

Finally, the PolyA preparation is then subjected to 1 round of Ribominus ribosomal depletion (Invitrogen: Cat ID: K1550-02).

1.1 Oligotex mRNA Spin-Column Protocol For isolation of poly A+ mRNA from total RNA

Abridged text taken directly text from Qiagen's Oligotex mini protocol
http://www1.qiagen.com/literature/handbooks/literature.aspx?id=1000156).

Important notes before starting:

- **Ensure buffers are appropriately prepared.**
- **Heat Oligotex Suspension to 37°C. Vortexing prior to use, then place at RT.**
- **Heat a water bath or heating block to 70°C, and heat Buffer OEB.**
- **Review the introductory material on pages 12–19 before starting.**

• If working with RNA for the first time, please read Appendix A (page 76).
• Buffer OBB may form a ppt upon storage. If necessary, re-dissolve at 37°C, then RT
• Unless otherwise indicated, all steps, including centrifugation, should be performed at RT.
• All centrifugation steps should be performed in a micro-centrifuge at maximum speed.

1. Determine the amount of starting RNA. Do not use more than 1 mg. The initial volume of the RNA solution is not important so long as the volume can be brought up to the indicated amount with RNase-free water. Make up the total RNA to 250µl with H20 and add 250µl of Buffer OBB.

2. Add 15µl of Oligotex Suspension. Mix the contents thoroughly by pipetting or flicking the tube. Incubate for 3 min at 70°C in a heating block. This step disrupts secondary structure of the RNA.

4. Remove sample from the water bath/heating block, and place at 20 to 30°C for 10 min. This step allows hybridization between the oligo dT30 of the Oligotex particle and the poly-A tail of the mRNA.

5. Pellet the Oligotex:mRNA complex by centrifugation for 2 min at maximum speed (14,000–18,000 x g), and carefully remove the supernatant by pipetting. Loss of the Oligotex resin can be avoided if approximately 50µl of the supernatant is left in the microcentrifuge tube. The remaining solution will not affect the procedure. Note: Save the supernatant until certain that satisfactory binding and elution of poly A+ mRNA has occurred.

6. Resuspend the Oligotex:mRNA pellet in 400µl l Buffer OW2 by vortexing or pipetting, and pipette onto a small spin column placed in a 1.5 ml micro-centrifuge tube. Centrifuge for 1 min at max speed.

7. Transfer the spin column to a new RNase-free 1.5 ml micro-centrifuge tube, and apply 400µl Buffer OW2 to the column. Centrifuge for 1 min at maximum speed and discard the flow-through.

8. Transfer spin column to a new RNase-free 1.5 ml micro-centrifuge tube. Pipette 20µl hot (70°C) Buffer OEB onto the column, pipette up and down 3 or 4 times to resuspend the resin, and centrifuge for 1 min at maximum speed. Note: The volume of Buffer OEB used depends on the expected or desired concentration of poly A+ mRNA. Ensure that Buffer OEB does not cool significantly during handling. Remember that small volumes cool down quickly. With multiple samples, it may be necessary to place the entire micro-centrifuge tube (with spin column, Oligotex, and sample) into a 70°C heating block to maintain the temperature while preparing the next samples.

9. To ensure maximal yield, pipette another 20µl hot (70°C) Buffer OEB onto the column. Pipette up and down 3 or 4 times to resuspend the resin, and centrifuge for 1 min at maximum speed. To keep the elution volume low, the first eluate may be used for a second elution. Reheat the eluate to 70°C, and elute in the same micro-centrifuge tube. However, for maximal yield, the additional volume of Buffer OEB is recommended.

10: Check selection on Bioanalyzer pico chip (load 0.5-1ng per well). Quantify by nanodrop rather than relying on Bioanalyzer estimates. Figure 2 shows the expected profile from pico chip bioanalysis after each stage of RNA preparation. Expected yield of mRNA from 100ug of total RNA is in the order 300ng.

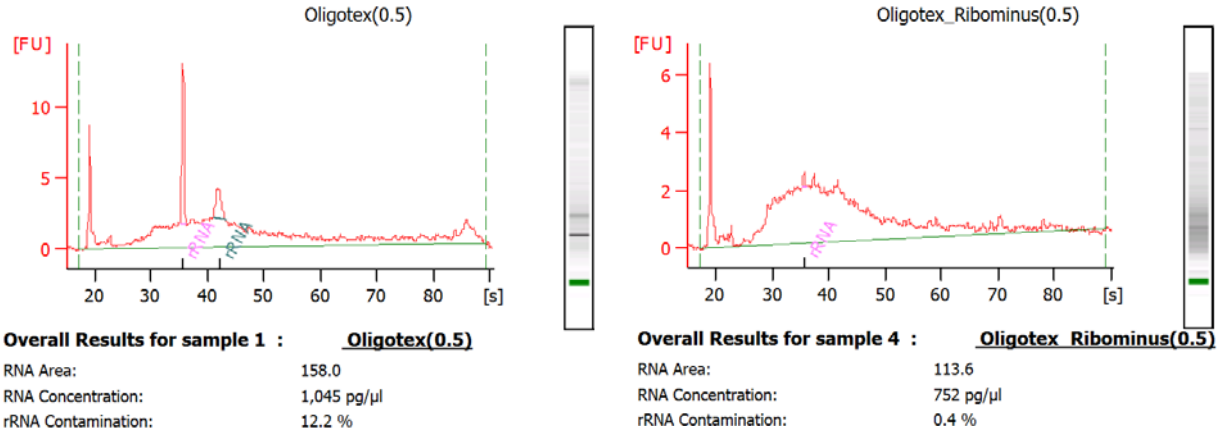


Figure2: Pico chip Bioanalyzer profiles of polyA prepared and ribo-depleted RNA.

1.2 Ribo-depletion of mRNA

Starting Material:

The protocol typically uses total RNA. We have found that it works better starting with mRNA for our sequencing purposes. We have tested the amount of starting material and up to 5ug of message can be used effectively.

Protocol is an abridged version of text taken directly from the manufacturer's manual
http://www.invitrogen.com/content/sfs/manuals/ribominus_human_mouse_man.pdf

Selective hybridization:
1. To a sterile, RNase-free 1.5 ml micro-centrifuge tube, add the following:
-RNA (2-10µg): <20µl
-RiboMinus™ Probe (100 pmol/µl): 8µl
-Hybridization Buffer (B5): 300µl

2. Incubate the tube at 70-75°C for 5 minutes to denature RNA. Allow the sample to cool to 37°C slowly over a period of 30 minutes by placing the tube in a 37°C water bath.

Prepare RiboMinus™ Magnetic Beads:
1. Resuspend the RiboMinus™ Magnetic Beads in its bottle by thorough vortexing.

2. Pipette 500µl of the bead suspension into a sterile, RNase-free, 1.5-ml microcentrifuge tube.

3. Place the tube with the bead suspension on a magnetic stand for 1 minute. Gently aspirate and discard the supernatant.

4. Add 500µl sterile, RNase-Free Water to the beads and resuspend beads. Place the tube on a magnetic stand for 1 minute. Gently aspirate and discard the supernatant.

5. Repeat Step 4 once.

6. Resuspend beads in 500µl Hybridization Buffer (B5). Place the tube on a magnetic stand for 1 minute. Gently aspirate and discard the supernatant.

7. Resuspend beads in 200µl Hybridization Buffer (B5) and keep the beads at 37°C until use.

Removing rRNA
1. Transfer ~328 µl of the cooled hybridized sample (from Step 3, previous page) to the prepared RiboMinus™ Magnetic beads from Step 7, previous page, and mix well.

2. Incubate the tube at 37°C for 15 minutes. During incubation, gently mix the contents occasionally.

3. Place the tube on a magnetic stand for 1 minute to pellet the rRNA-probe complex. The supernatant contains the RiboMinus™ RNA fraction.

4. Transfer the supernatant (~ 528 µl) to a tube capable of holding 3X volume of the supernatant.
Concentrating RiboMinus™ RNA using RiboMinus™ Concentration Module.

5. The Recovery Tube contains purified RiboMinus™ RNA. Store RiboMinus™ RNA at -80°C or place RiboMinus™ RNA on ice to proceed to desired downstream application.

2.1 Fragmentation of RNA

Notes: The fragmentation step is a necessary and critical part of this protocol. Generation of a directional shotgun cDNA library by template switching requires random fragmentation of the RNA. Next-gen sequencing also requires small amplimers for efficient and unbiased PCR of sequencing templates.

We currently use 2 different methods of RNA fragmentation. Both appear to give an unbiased tag representation in the libraries we have prepared to date. While the original work has relied on chemical fragmentation we are also using heat as an alternative, given the ability to control it and the avoidance of clean up steps. Protocols for both are provided.

We have found that both Concentration and Volume are important parameters in the fragmentation process and the amounts suggested here have been optimized to give the best size range of amplimers and minimal primer-dimer etc down stream.

It is strongly recommended you calibrate this step independently in your lab prior to starting on real samples. We have used a mixture of total and aRNA (made from the standard message Amp kit from Ambion) as a cost-effective proxy to mRNA for optimizing this step. ***Differences in fragmentation can effect the representation of sequence tag sampling and this can have serious consequences if you are looking to do comparative quantitative analyses across sequence runs.***

Note: We have found considerable batch to batch variation with chemical fragmentation reagents. We have also found that the heating step needs to be carried out in ***a PCR machine with a heated lid .*** The times and temps here have been optimized on a MJ Research Engine Tetrad. We also recommend accurate pipettors be used! The volumes have been kept small to ensure sample volumes and RNA requirements are kept to a minimum.

2.1.1 Fragmentation of RNA by Zinc Acetate protocol:

Chemical fragmentation is routinely used to break up long labeled aRNA molecules prior to hybridizing to microarrays. We used this approach as it minimized the volume needed to fragment (sonication needs a decent volume) and

1. The suggested amount of starting material 8 µl polyA depleted RNA (~300ng).

2. Add 0.9 µl fragmentation buffer: (buffered Zinc acetate, AMBION:8740)

3. Heat to 70^0C for a period of time that give you optimal fragmentation to >100 and <200bp. In our hands this equates to **2 mins, performed** in PCR machine with a heated lid). Suboptimal results have been observed in the absence of the hot lid. It is strongly suggested you carry out a panel of fragmentations for different periods and follow them through to library synthesis and review the overall size of template to ensure you are in

an optimal range. This material cannot be run directly on the Bioanalyzer prior to purification due to the fragmentation buffer.

4. Chill on ice and add 0.9µl STOP solution (supplied in AMBION kit)

5. Dilute to 500µl with ice cold RNAse free water.

6. Keep on ice until purification

NOTE =Even when the system has been calibrated, we recommend running an independent positive control RNA sample of the same amount, to show that the fragmentation and cDNA synthesis works efficiently.

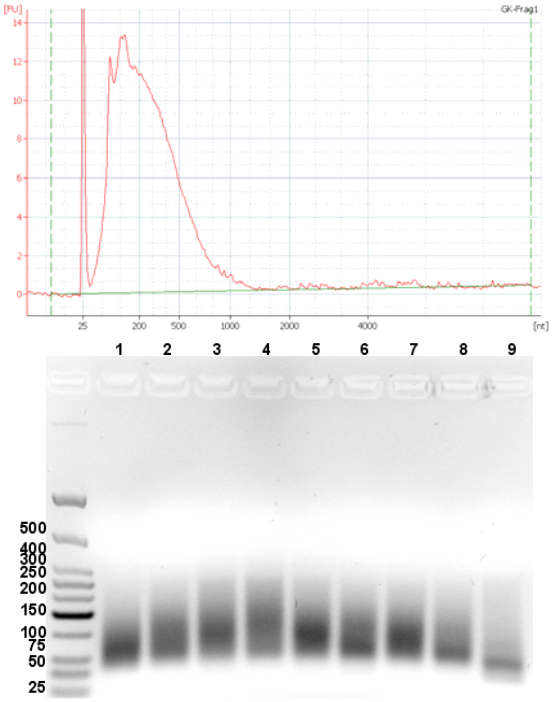
7. Purify fragmented RNA by ultra-filtration on a YM-30 column (Millipore) (CAT ID#)

8. Spin for 12 mins at 13,000 rpm. Check the volume retained on the filter. If you can see or almost see the filter, invert filter and collect retained sample via a pulse spin to 10,000rpm

9. Check the amount recovered. If < 1.7µl, make up to 1.7µl. If the amount exceeds >1.7 µl, concentrate by vacuum centrifugation and make up to 1.7µl. DO NOT OVER DRY.

10. Check fragmentation on Bioanalyzer pico chip (load 0.5-1ng per well). See Fig 3 for expected results:

A B



A: Bioanalysis of fragmented RNA after clean up. **B:** Examples of library generation from a spectrum of different fragmented RNAs: RNAs from a spectrum of 1: (300ng Hek total, 0.6ul frag reagent in 6ul 5min), 2: (300ng amplified, 0.6ul frag reagent in 6ul 5min), 3: (300ng Hek total 0.6ul frag reagent in 6ul 2min), 4: (300ng amplified 0.6ul frag reagent in 6ul 2min), 5: (300ng Hek total 0.06ul frag reagent in 6ul 5min), 6: (300ng amplified 0.06ul frag reagent in 6ul 5min), 7: (600ng Hek total 0.6ul frag reagent in 6ul 5min), 8: (600ng amplified 0.6ul frag reagent in 6ul 5min), 9: (300ng Hek total UNFRAGGED -,10: (300ng amplified UNFRAGGED -, 11: (NO RNA(, 12: (NO RNA/RT),

2.1.3: Alternate Protocol: Fragmentation of RNA by heat method:

1. Transfer 300ng of mRNA to a 200µl PCR tube at 100ng/µl concentration. Adjust volume to 4.5µl with nuclease free H2O.

2. Place on heating block (PCR block) at 95^0C for optimized period. In our hands this is 20-30mins. has proven optimal. (see figure 4).

3. Transfer to ice for minimum 2mins;

4. Concentrate fragmented RNA in speedi-vacuum (~5-10mins) to 1.5 µl.

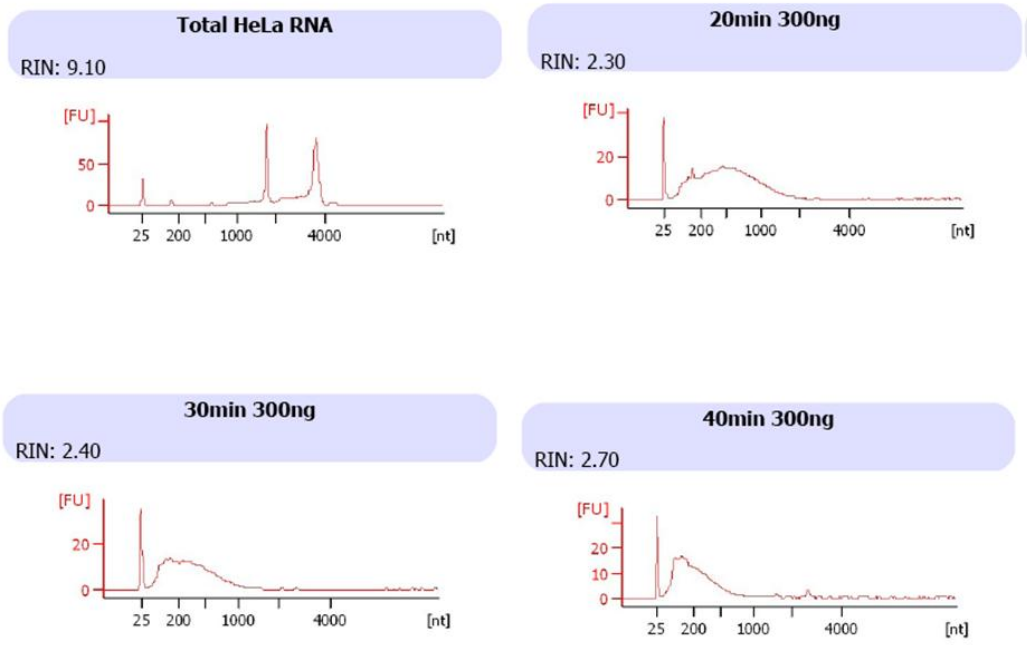


Figure 4: Pico Bioanalyzer profiles on 300ng total RNA subjected to 95^0C over 20min-40min.

3.1 cDNA synthesis with template switch.

Notes: This protocol uses template switching (which is used in Clontech's SMART protocol). It is dependent upon an modified MMLVs ability to add non-templated C's as a 3' overhang once reverse transcription reaches the end of a DNA template. These extra C's are used to hybridize a GGG stretch of the template switch primer. The RT then continues on to the end of the template switch primer. This generates a 1st strand fragment with FDV and RDV sequences at each end.

Since the design of this protocol, Clontech has ceased producing Powerscript reverse transcriptase. Superscript II can be used as a substitute. NOTE: Superscript-III will make good first strand cDNA but will NOT add non-templated bases.

Special comment needs to be made regarding controls. BOTH a no RNA, and a no RNA & no RT control should be included. The template negative control WILL generate primer dimers using the primers optimized for SOLiD sequencing. This product is ~100bp and efforts have been made to minimize its amplification. As the library is size selected post amplification, it is possible to avoid this if it is a minor contaminant.

3.1.1 Protocol:

1. Make up required amount of sFDV-hex primer at 2μM (Stock at 10μM)

2. Preheat one PCR block to 72^0C and a second PCR block to 20^0C

3. Prepare the following reagents on ice

Reagent	Control 1	Control 2	Test….
dH20	1.5	1.5	0
RNA (250ng/1.5μl)	0	0	1.5
sFDVhex (2μM)	0.5	0.5	0.5

4. Heat mix to 72^0C for 10min followed by chilling on ice for 2min. Note: This step is under scrutiny to determine the optimal temperature for minimizing hexamer mis-priming.

5. Add RT reagents (both Powerscript and Superscript II volumes are provided for a single RT reaction). A master mix should be made and added to primed RNA when multiple samples libraries are being prepared.

Powerscipt	μl	**Superscript II**
Powerscript RT Buffer	1.0	Superscript RT Buffer
1.0		

DTT (20mM – be careful to check) 0.5	0.5	DTT (20mM)
dNTP (10mM each nucleotide) 0.5	0.5	dNTP(10mM each)
Powerscript reverse transcriptase 0.5	0.5	Superscript II enzyme

6. Move samples to thermocycler. NOTE: DO NOT USE A HEATED LID. Including the lid affects cDNA synthesis and the amount of primer dimmer. Incubate the samples at 20^0C for 10min, followed by 37^0C for10min and then 42^0C for 45min. This is necessary to generate full length 1st strand cDNAs and template for each fragment.

7. Add 0.5µl of heat denatured 10uM RDV-GGG template switch primer to each tube and incubate for a further 15min at 42^0C. This is required to prime 2nd strand from non-templated Cs present on the end of the single cDNA strand. Heat Denaturing of RDV-GGG template switch primer prior to priming (72^0C for 5min) is required to maximize the efficiency of 2nd strand cDNA synthesis.

8. Stop the cDNA synthesis by > 95^0C for 5min.

4.1 PCR amplification of cDNA library (using Advantage polymerase from Clontech)

Set up the following

H2O	13.5ul
2mM dNTP each	2.5ul
10x Buffer	2.5ul
10uM AMPRDV	2.5ul (needed to minimze primer dimer)
2uM LAMPFDV	2.5ul
1st strand cDNA	1.0ul
Advantage DNA polymerase	0.5ul

FDV primers	sFDVhex	CTTTCCTCTCTATGGGCAGTCGGTGATNNNNNN
	LAmpFDV	CCACTACGCCTCCGCTTTCCTCTCTATGGGCAGTCGGTGAT
RDV primers	AmpRDV	AACTGCCCCGGGTTCCTCATTCTCT
	RDV-GGG	AACTGCCCCGGGTTCCTCATTCTCTrGrGrG

Thermocycle as follows: 1 x 94^0C -> 5 min (activate the polymerase)
20 x 94^0C -> 15s, 68^0C -> 15s }

NOTE for the AMPRDV concentration is it is at 5x normal primer concentration, to avoid concatamers

4.2 Size Selection of amplimers via gel electrophoresis

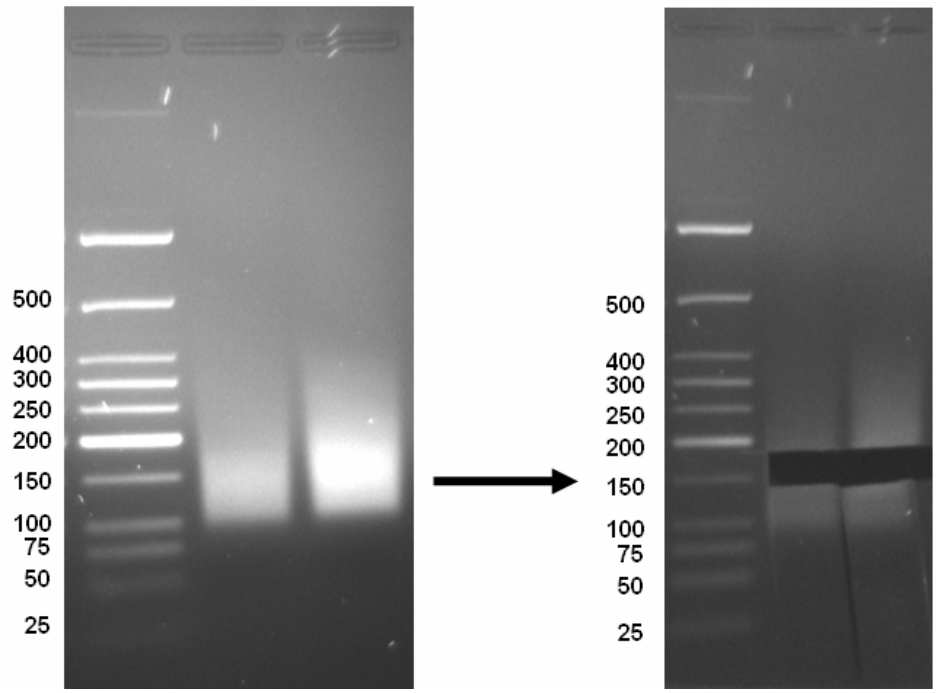
1. Load the entire library onto a 1xTAE 3% agarose gel and run at 100V, 55 minutes with LMW ladder.

2. Cut out band with a scalpel. Use a unique scalpel every time.

3.Purify DNA via Qiaquick agarose extraction:

4. Optional: Reload the purified material onto another 3% Agarose gel and repeat purification. (we've found this necessary to remove smaller amplicons).

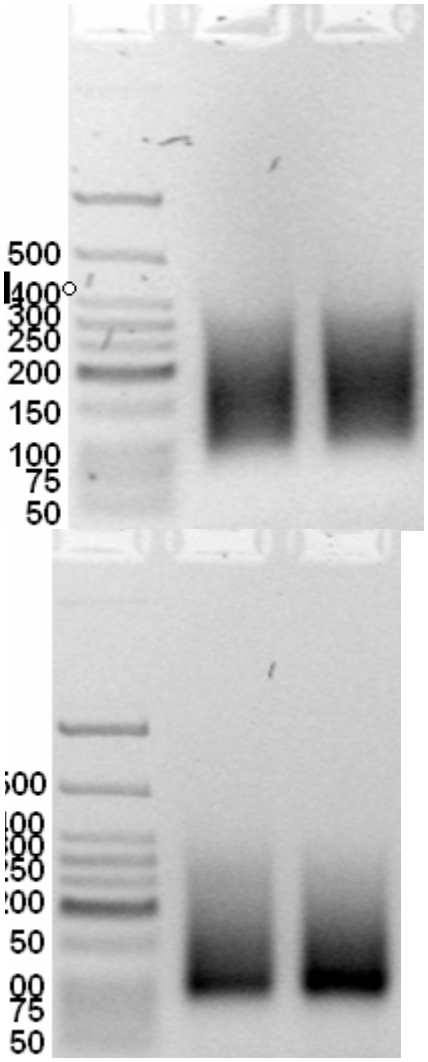
5. Quantify final amplier pool concentration by nanodrop.



4.3 Assessing library synthesis.

A good library should have a smear ranging from 100 to 300 with the majority of the product in the 120 – 180 range. (lanes 1 and 2 of the first gel below are good syntheses).

A failed or poor synthesis is generally characterized by a strong band just below 100 and a very faint smear if anything. The figs below show examples of good Vs poor product range.



5. Library QC by sequencing.

Notes: We QC libraries prior to SOLiD sequencing by both capillary sequencing of cloned amplimers and accurate quantification. To sequence, we take 1ul and 0.1ul of eluted library and clone into a TA cloning vector (TOPO-TA PCR-TOPO).

Pick 96 colonies and prepared plasmid DNA via for 96well miniprep. Samples are sequenced using standard AB sequencing via core.

Sequences are reviewed on the following criteria:
-Insert size
-Length of mappable sequence (BLAT UCSC and BLAST NCBI)
-Count of non-templated Gs at the 5' end.

Critical statistics:

1) Median amplicon size: Currently our preferred median insert size of 50bp. This size is preferred as it suits our read length (we're currently getting 35mers).

2). Range of amplicon size: In some cases we have had a good median size but too broad a size range. Ideally you want to ensure that all sequences are larger than the minimum read length. Insert size for us now should be a min of 35bp long and max of 70bp).

3) %Ribosomal RNA: Ribosomal RNAs are a common contaminant if depletion was poor.

4) Position relative to full length sequences: Once tags are mapped to their respective genes, we check to see if the tags are spread across the entire length of the transcript Vs biased towards the 3' or 5' end.
(check for 3'/5' biases)