

Supplementary Figure 1: Box-plots of read length distributions across seven tissues targeted by 5' and 3' standard ("std") and nested RACE-Seq.

testis



Supplementary Figure 2: Genome mapping statistics. Bar plot showing the number of 454 reads that were unmapped (blue) mapped uniquely (green) and multiple times (coral) to the reference human genome (GRCh37).



Supplementary Figure 3: Flowchart explaining the CAGE enrichment analysis and summarized results.



Supplementary Figure 4: Read mapping statistics. Number of uniquely mapped 454 reads overlapping GENCODEannotated loci (green), vs. in intergenic regions (coral).



Supplementary Figure 5: Number of amplified non-targeted loci in nested and standard RACE-Seq in the seven tissues assayed.



Supplementary Figure 6: Venn diagram comparing the sets of targeted loci that could be amplified in standard, primary RACE (blue) and nested RACE (orange)



Supplementary Figure 7: Genome browser screenshot of the PIGL locus. The original locus (highlighted, bottom left) was subsumed into the coding locus PIGL. The RNA-seq signal plot (light blue, all 454 data bumped) shows the density of reads mapping to exons. The RACE

А								В							
10	% Share	ed trans	cripts						% Share	d genes					
Brain	100 (516/515)	37.4 (193/516)	<b>43.8</b> (226/516)	32.2 (166/516)	<b>39.1</b> (202/516)	<b>43.8</b> (226/516)	57 (294/516)	Brain	100 (151/051)	60.3 (91/151)	60.9 (92/151)	45.7 (69/151)	60.3 (91/151)	60.9 (92/151)	85.4 (129/151)
Heart	68 (193/284)	1 <b>00</b> (284/284)	63.4 (180/284)	45.1 (128/284)	51.4 (146/284)	62 (176/284)	69.7 (198/284)	Heart	82 (91/111)	100 000000	74.8 (83/111)	54.1 (60/111)	67.6 (75/111)	76.6 (85/111)	90.1 (190/111)
Kidney	62.1 (226/364)	49.5 (180/364)	100 (364/364)	45.6 (166/364)	50.3 (183/364)	50.5 (184/364)	61.8 (225/364)	Kidney	76.7 (92/120)	69.2 (83/120)	100	56.7 (68/120)	67.5 (81/120)	64.2 (77/120)	82.5
Liver	69.2 (166/240	53.3 (128/240)	69.2 (166/240)	100 (240/240)	56.2 (135/240)	<b>57.1</b> (137/240)	74.2	Liver	78.4 (69/88)	68.2 (60/88)	77.3	100 (00000	75 (66/88)	69.3 (61/88)	87.5 (77/88)
Lung	60.1 (202/336)	<b>43.5</b> (146/336)	54.5 (183/336)	40.2	100 (336/336)	<b>53</b> (178/336)	65.8 (221/336)	Lung	74.6 (91/122)	61.5 (75/122)	66.4 (81/122)	54.1 (66/122)	100 (12.0132)	66.4 (81/122)	84.4 (103/122)
Spleen	67.7 (226/334)	52.7 (176/334)	<b>55.1</b> (184/334)	<b>41</b> (137/334)	<b>53.3</b> (178/334)	100 (334/334)	66.2 (221/334)	Spleen	80.7 (92/114)	74.6 (85/114)	67.5 (77/114)	53.5 (61/114)	71.1 (81/114)	100 (1971))	86.8 (99/114)
Testes	18.2	12.2	13.9 (225/1618)	11	13.7 (221/1618)	13.7 (221/1618)	100 (1618/1618)	Testes	43.6 (129/296)	33.8 (100/296)	33.4 (99/296)	26 (77/296)	34.8 (103/296)	33.4 (99/296)	100
	Brain	Heart	Kidney	Liver	Lung	Spleen	Testes		Brain	Неал	Kidney	Liver	Lung	Spleen	Testes

Supplementary Figure 8: Detection of IncRNAs in targeted tissues. Heat map showing the number of detected IncRNAs in each tissue and their proportion shared across other targeted tissues at the (A) transcript level (B) gene level. Transcripts were reconstructed directly from the read alignments, using transcript structure compatibility as merging criterion. A given locus was considered as expressed in a given tissue if at least one of its transcript was successfully reconstructed.



Supplementary Figure 9: Distribution of detected splice junctions by RACE-Seq (bottom panel) and CaptureSeq (top panel) along targeted genes. Bar plots showing the location of splice junctions within targeted loci boundaries (+/-5kb). (a) All annotated and unannotated splice junctions within targeted loci boundaries (+/-5kb). (a) All annotated splice junctions detected in RNA CaptureSeq and RACE-Seq data sets. (B) The top 25% of annotated and unannotated canonical splice junctions ranked by read coverage.



Supplementary Figure 10: Detection of pre- and post-RACE-Seq targets by SLR-Seq. Number of loci covered (blue) and not covered (green) by SLR-Seq reads.

Dataset	Total # TSS (clustered)	# Clustered TSS +/- 50 bp from CAGE tag	% Clustered TSS +/- 50 bp from CAGE tag
Targets (pre-RACE)	527	241	46%
Targets (post-RACE, all)	873	415	48%
Targets (post-RACE, novel)	615	252	41%
CaptureSeq transcript models (all)	343	203	59%
CaptureSeq transcript models (novel)	70	37	53%

Supplementary Table 1: Table summarizing TSS discovery and CAGE coverage statistics in both RACE-Seq and Clark et al.'s CaptureSeq.

	Pre RA	CE-Seq data)	(GTEx	Post RACE-Seq (standard)				Post RACE-Seq (nested)			
Tis sue	Total # mappe d reads	# reads within target ed transc ripts	% reads within target ed transc ripts	Total # map ped read s	# reads within target ed transc ripts	% reads within target ed transc ripts	On- target fold enrich ment	Total # map ped read s	# reads within target ed transc ripts	% reads within target ed transc ripts	On- target fold enrich ment
brai n	37,162 ,929	330,9 59	0.9%	746,8 89	41,01 9	5.5%	6.2	655,5 24	328,6 66	50.1%	56.3
hea rt	35,835 ,990	564,3 11	1.6%	857,5 52	14,47 5	1.7%	1.1	602,8 69	171,6 68	28.5%	18.1
kid ney	45,384 ,859	544,0 64	1.2%	923,7 80	28,10 4	3.0%	2.5	730,4 76	291,2 91	39.9%	33.3
live r	39,432 ,611	282,2 30	0.7%	652,5 92	18,34 5	2.8%	3.9	751,2 22	199,5 11	26.6%	37.1
lun g	48,531 ,622	792,6 44	1.6%	809,1 39	18,89 4	2.3%	1.4	834,9 09	256,8 59	30.8%	18.8
spl een	33,051 ,498	475,4 24	1.4%	807,7 53	16,77 5	2.1%	1.4	866,2 28	272,3 86	31.4%	21.9
test is	43,888 ,261	545,5 28	1.2%	757,1 08	76,25 0	10.1%	8.1	679,7 35	345,7 09	50.9%	40.9
TO TAL	283,28 7,770	3,535, 160	1.2%	5,554 ,813	213,8 62	3.9%	3.1	5,120 ,963	1,866, 090	36.4%	29.2

Supplementary Table 2: On-target read enrichment statistics. 5' and 3' RACE datasets were merged in each tissue.

# Targets	RACE direction	RACE type	# Targets successfully RACE'd	% Targets successfully RACE'd
		Standard	248	62%
	5'	Nested	314	79%
398		Standard + Nested	341	86%
		Standard	255	64%
	3'	Nested	293	73%
		Standard + Nested	326	82%

Supplementary Table 3: Number and proportion of successfully RACE-amplified targets.

Dataset	Total # unique splice junctions	# Supported by RNA CaptureSeq	% Supported by RNA CaptureSeq
Targets (pre-RACE)	1,093	903	83%
Targets updated (post-RACE)	3,664	2,211	60%

Supplementary Table 4: Detection of pre- and post-RACE-Seq targets by RNA CaptureSeq. Proportion of annotated splice junctions in pre- and post-RACE-Seq targets supported by RNA CaptureSeq.

Dataset	Total # unique splice junctions	# Supported by 454-RACE- Seq	% Supported by 454-RACE- Seq	# Supported by SLR-Seq	% Supported by SLR-Seq
Targets (pre-RACE)	1,093	817	74.8%	226	20.68%
Targets updated (post-RACE)	3,664	3,277	89.4%	281	9.11%

Supplementary Table 5: Comparison of splice junction support by 454 RACE-seq and SLR-seq.