`

Supplementary Online Material

The Long Non-coding RNA *MALAT1* Predicts Human Islet Isolation Quality

Wilson K.M. Wong[1], Guozhi Jiang[2], Anja E. Sørensen[3], Yi Vee Chew[4], Cody Lee-Maynard[1], David Liuwantara[4], Lindy Williams[4], Philip O'Connell[4], Louise T. Dalgaard[3], Ronald C. Ma[2], Wayne J. Hawthorne[4,✶], Mugdha V. Joglekar[1,✶] and Anandwardhan A. Hardikar[1,✶]

Running title: The lncRNA *MALAT1* predicts islet isolation outcome

1- Diabetes and Islet Biology Group, NHMRC Clinical Trials Centre, The University of Sydney, Level 6, Medical Foundation Building, 92-94 Parramatta Road, Camperdown, NSW 2050, Australia.
2- Department of Medicine and Therapeutics, The Chinese University of Hong Kong, Prince of Wales Hospital, Hong Kong.
3- Department of Science and Environment, Roskilde University, Universitetsvej 1, DK-4000 Roskilde, Denmark
4- Centre for Transplant and Renal Research, Westmead Institute for Medical Research, Sydney University, 176 Hawkesbury Rd., Westmead, NSW 2145, Australia.

✶ Equal senior authors; address for correspondence:

✶Wayne J Hawthorne,
Centre for Transplant and Renal Research, Westmead Institute for Medical Research, Sydney University, 176 Hawkesbury Rd., Westmead, NSW 2145, Australia.
E-mail: wayne.hawthorne@sydney.edu.au

✶Mugdha V. Joglekar or ✶Anandwardhan A. Hardikar
Diabetes and Islet Biology Group, NHMRC Clinical Trials Centre,
The University of Sydney, Level 6, Medical Foundation Building,
92-94 Parramatta Road, Camperdown,
NSW 2050, AUSTRALIA
Phone: +61 2 9562 5071 | Fax: +61 2 9565 1863
E-mails: mugdha.joglekar@sydney.edu.au; anand.hardikar@sydney.edu.au
Web: http://www.isletbiology.com/ | http://www.ctc.usyd.edu.au/

	Group 1		Group 2		Group 3		Adjusted P-value (ANOVA)		
	Average	SD	Average	SD	Average	SD	1 vs. 2	1 vs. 3	2 vs. 3
Donor gender: Female (F)/ Male (M)	2/3		2/4		2/5				
Donor Age(yrs)	49.8	5.6	50.3	9.4	43.9	18.4	0.95	0.77	0.77
Donor BMI kg/m2	29.0	1.4	30.4	13.4	26.0	4.2	0.80	0.80	0.73
Islet quality score	0.4	0.1	0.7	0.1	1.2	0.5	0.13	0.001	0.011
Edmonton Donor Points	81.5	8.7	67.3	15.8	58.7	13.8	0.19	0.04	0.28
NAIDS	75.0	4.2	71.2	11.3	57.4	10.5	0.52	0.02	0.04
Cold Ischemia (mins)	319.6	148.9	464.4	156.3	447.6	140.3	0.37	0.37	0.85
Pancreas wt. (g)	94.0	10.8	102.8	20.8	79.1	18.1	0.42	0.31	0.08
Pancreas undigested (g)	13.7	6.5	21.7	14.2	13.6	6.9	0.42	0.98	0.42
Pancreas wt digested(g)	80.3	15.2	68.8	10.5	65.5	16.7	0.38	0.28	0.69
Total Time in chamber(min)	63.6	9.6	51.0	7.6	55.4	13.5	0.20	0.38	0.47
Time to 37'C (min)	9.3	3.8	9.6	3.6	10.0	3.7	0.98	0.98	0.98
Total IEQ	434542.6	138126.5	372350.2	306254.5	284580.1	240077.6	0.64	0.16	>0.9999
Pure IEQ	374322.2	171793.0	217493.3	115840.3	167738.4	98388.2	0.12	0.04	0.49
Stimulation Index	1.6	0.7	1.7	1.1	1.4	0.4	0.91	0.91	0.88

Supplementary Table 1: Details and characteristics of cadaveric donor islets used in the discovery set (*n*=18) for this study. Different clinical parameters of the donor, pancreas and isolation steps are presented in this table for the 18 islet samples that were used as discovery set. A Chi-square test was applied on the donor sex across the groups in the discovery set (P-values = 0.92). One-way ANOVA multiple comparison was then used to assess significant differences across groups. The P-value adjusted for multiple comparisons is presented. P values less than 0.05 are highlighted in red.

	Group 1		Group 2		Group 3		Adjusted P-value (ANOVA)		
	Average	SD	Average	SD	Average	SD	1 vs. 2	1 vs. 3	2 vs. 3
Donor gender: Female (F)/ Male (M)	21/19		7/11		5/12				
Donor Age(yrs)	46.9	8.8	46.7	13.4	41.0	14.7	0.95	0.23	0.28
Donor BMI kg/m2	30.2	6.5	28.1	5.1	27.0	4.2	0.35	0.16	0.57
Islet quality score	0.5	0.1	0.7	0.1	1.6	0.9	0.06	<0.0001	<0.0001
Edmonton Donor Points	73.2	10.5	70.1	11.2	66.5	15.6	0.60	0.19	0.60
NAIDS	67.3	15.2	67.8	11.5	62.9	17.8	0.89	0.68	0.68
Cold Ischemia (mins)	429.7	154.3	437.9	135.0	418.2	128.3	0.97	0.97	0.97
Pancreas wt. (g)	91.6	30.1	95.0	17.7	84.8	18.0	0.64	0.58	0.56
Pancreas undigested (g)	17.6	11.0	12.9	6.7	22.0	14.0	0.25	0.25	0.05
Pancreas wt digested(g)	74.0	30.4	82.1	16.6	62.8	24.6	0.28	0.28	0.10
Total Time in chamber(min)	58.3	6.7	56.1	11.1	58.2	13.1	0.81	0.98	0.81
Time to 37'C (min)	11.0	4.2	8.8	4.0	9.6	4.1	0.20	0.44	0.58
Total IEQ	403265.3	244714.9	405281.6	249660.6	243918.2	147965.5	0.98	0.05	0.08
Pure IEQ	328521.1	237911.5	299803.8	214721.0	155396.2	83995.6	0.63	0.02	0.09
Stimulation Index	2.0	1.2	1.6	0.7	1.5	0.7	0.30	0.29	0.76

Supplementary Table 2: Details and characteristics of cadaveric donor islets used in the validation set (*n*=75) in this study. Different clinical parameters of the donor, pancreas and isolation steps are presented in this table for the 75 islet samples that were used as validation set. A Chi-square test was applied on the donor sex across the groups in the validation set (P-values = 0.24). One-way ANOVA multiple comparison was then used to assess significant differences across groups. The P-value adjusted for multiple comparisons is presented. P values less than 0.05 are highlighted in red.

	Group 1		Group 3		P-value (Mann-Whitney test)
	Average	SD	Average	SD	1 vs. 3
Donor gender: Female (F)/ Male (M)	7/7	-	2/3	-	
Donor Age(yrs)	45.7	9.1	33.0	12.6	0.05
Donor BMI kg/m2	31.5	6.0	25.0	4.5	0.07
Islet quality score	0.4	0.1	1.3	0.9	**<0.0001**
Edmonton Donor Points	71.7	10.6	63.7	19.6	0.61
NAIDS	75.3	9.7	54.8	21.5	0.02
Total IEQ	495529.1	240182.4	211120.8	131062.8	0.01
Cold Ischemia (mins)	432.4	96.9	415.2	96.0	0.62
Pancreas wt. (g)	93.7	21.6	86.4	23.9	0.67
Pancreas undigested (g)	14.7	9.8	13.1	10.2	0.74
Pancreas wt digested(g)	79.0	17.6	73.3	18.2	0.55
Total Time in chamber(min)	58.4	6.2	55.8	5.1	0.43
Time to 37'C (min)	13.2	4.0	9.7	4.5	0.12
Total Islet No.	166987.9	46929.2	118666.6	57197.6	0.19
Pure IEQ	441379.8	245688.5	170797.8	84169.9	0.0007
Stimulation Index	3.1	1.6	1.5	0.3	0.01

Supplementary Table 3: Details and characteristics of cadaveric donor pancreas used in the prediction set (*n*=19 pancreas) in this study Different clinical parameters of the donor, pancreas and isolation steps are presented in this table for the 19 pancreas samples that were used as prediction set. A Fisher's exact test was performed on the donor sex across the groups in the pancreas predict set (P-value >0.9999). Mann-Whitney test was then used to assess significant differences across groups and resulting P-values are presented. P values less than 0.05 are highlighted in red.

LncRNA: Group 1 (N=5) vs Group 2 (N=6)

Transcript ID	Gene ID	Official Symbol	Gene type	Frequency (%)
ENST00000616691.1	ENSG00000251562.7	MALAT1	lincRNA	93.6
ENST00000419809.1	ENSG00000235495.1	AC010987.5	lincRNA	57.8
ENST00000619449.1	ENSG00000251562.7	MALAT1	lincRNA	45.3
ENST00000544034.1	ENSG00000256725.1	RP11-662M24.2	lincRNA	24.5
ENST00000583934.1	ENSG00000265401.1	RP11-138I1.4	antisense	1

LncRNA: Group 2 (N=6) vs Group 3 (N=7)

Transcript ID	Gene ID	Official Symbol	Gene type	Frequency (%)
ENST00000620902.1	ENSG00000251562.7	MALAT1	lincRNA	98.7
ENST00000530422.1	ENSG00000254911.3	SCARNA9	antisense	65
ENST00000414790.5	ENSG00000130600.15	H19	processed_transcript	63.9
ENST00000583934.1	ENSG00000265401.1	RP11-138I1.4	antisense	58.4
ENST00000418403.1	ENSG00000237643.1	RP11-462G2.1	lincRNA	29.7
ENST00000419809.1	ENSG00000235495.1	AC010987.5	lincRNA	25.8
ENST00000450589.5	ENSG00000234741.7	GAS5	processed_transcript	24.5
ENST00000414002.5	ENSG00000203875.10	SNHG5	processed_transcript	19
ENST00000563192.1	ENSG00000260260.1	SNHG19	lincRNA	0.1
ENST00000601801.2	ENSG00000245532.5	NEAT1	lincRNA	0.1

LncRNA: Group 1 (N=5) vs Group 3 (N=7)

Transcript ID	Gene ID	Official Symbol	Gene type	Frequency (%)
ENST00000616691.1	ENSG00000251562.7	MALAT1	lincRNA	96.3
ENST00000620902.1	ENSG00000251562.7	MALAT1	lincRNA	45.9
ENST00000619449.1	ENSG00000251562.7	MALAT1	lincRNA	31.5
ENST00000418403.1	ENSG00000237643.1	RP11-462G2.1	lincRNA	12.2
ENST00000414002.5	ENSG00000203875.10	SNHG5	processed_transcript	0.3
ENST00000414790.5	ENSG00000130600.15	H19	processed_transcript	0.3

Supplementary Table 4: Bootstrap frequency tables of the random sampling revalidation performed using penalized logistic regression analysis. Bootstrap revalidation was performed 1000 times, wherein a different (~37%) subset of samples are eliminated at random in each penalized regression analysis iteration. A penalized linear regression analysis with bootstrapping delivered similar outcomes where *MALAT1* variants were identified as the top three candidates with a bootstrap frequency ranging from 57-78% (data not shown).

`

Isolation outcome: Transplantable Yes(1) (N=5) vs No(0) (N=13)

Transcript ID	Gene ID	Official Symbol	Gene type	Frequency (%)
ENST00000616691.1	ENSG00000251562.7	MALAT1	lincRNA	94.3
ENST00000619449.1	ENSG00000251562.7	MALAT1	lincRNA	46.1
ENST00000544034.1	ENSG00000256725.1	RP11-662M24.2	lincRNA	44.8
ENST00000620902.1	ENSG00000251562.7	MALAT1	lincRNA	28.4
ENST00000418403.1	ENSG00000237643.1	RP11-462G2.1	lincRNA	9.8

LncRNA: Post-Iso Purity (N=15)

Transcript ID	Gene ID	Official Symbol	Gene type	Frequency (%)
ENST00000616691.1	ENSG00000251562.7	MALAT1	lincRNA	68.8
ENST00000414790.5	ENSG00000130600.15	H19	processed_transcript	58.8
ENST00000620902.1	ENSG00000251562.7	MALAT1	lincRNA	49.3
ENST00000619449.1	ENSG00000251562.7	MALAT1	lincRNA	32.5
ENST00000602520.5	ENSG00000269893.6	SNHG8	lincRNA	28.6
ENST00000563192.1	ENSG00000260260.1	SNHG19	lincRNA	21.5
ENST00000544034.1	ENSG00000256725.1	RP11-662M24.2	lincRNA	13.8
ENST00000414002.5	ENSG00000203875.10	SNHG5	processed_transcript	12.8
ENST00000576784.1	ENSG00000262831.1	RP11-498C9.2	antisense	10.8
ENST00000450589.5	ENSG00000234741.7	GAS5	processed_transcript	10.5
ENST00000418403.1	ENSG00000237643.1	RP11-462G2.1	lincRNA	8.1
ENST00000430245.5	ENSG00000234741.7	GAS5	processed_transcript	4.1
ENST00000508832.2	ENSG00000251562.7	MALAT1	lincRNA	0.1

LncRNA: Post-Iso Viability (N=16)

Transcript ID	Gene ID	Official Symbol	Gene type	Frequency (%)
ENST00000541782.1	ENSG00000276232.1	SCARNA10	sense_intronic	65.1
ENST00000583934.1	ENSG00000265401.1	RP11-138I1.4	antisense	38.2
ENST00000616691.1	ENSG00000251562.7	MALAT1	lincRNA	37.2
ENST00000414002.5	ENSG00000203875.10	SNHG5	processed_transcript	35.3
ENST00000620902.1	ENSG00000251562.7	MALAT1	lincRNA	29.6
ENST00000619449.1	ENSG00000251562.7	MALAT1	lincRNA	24.5
ENST00000418403.1	ENSG00000237643.1	RP11-462G2.1	lincRNA	10.8
ENST00000530422.1	ENSG00000254911.3	SCARNA9	antisense	8.4
ENST00000602520.5	ENSG00000269893.6	SNHG8	lincRNA	7.7
ENST00000499732.2	ENSG00000245532.5	NEAT1	lincRNA	2.5
ENST00000601801.2	ENSG00000245532.5	NEAT1	lincRNA	1.9
ENST00000430245.5	ENSG00000234741.7	GAS5	processed_transcript	0.9
ENST00000491934.2	ENSG00000269926.1	RP11-442H21.2	antisense	0.1
ENST00000580180.5	ENSG00000175061.17	LRRC75A-AS1	processed_transcript	0.1
ENST00000548900.1	ENSG00000257764.2	RP11-1143G9.4	antisense	0.1
ENST00000508832.2	ENSG00000251562.7	MALAT1	lincRNA	0.1

LncRNA: Beta cell Viability Index (N=14)

Transcript ID	Gene ID	Official Symbol	Gene type	Frequency (%)
ENST00000616691.1	ENSG00000251562.7	MALAT1	lincRNA	79.9
ENST00000541782.1	ENSG00000276232.1	SCARNA10	sense_intronic	64.5
ENST00000419809.1	ENSG00000235495.1	AC010987.5	lincRNA	44.9
ENST00000619449.1	ENSG00000251562.7	MALAT1	lincRNA	34.7
ENST00000414790.5	ENSG00000130600.15	H19	processed_transcript	22.4
ENST00000544034.1	ENSG00000256725.1	RP11-662M24.2	lincRNA	21.3
ENST00000602520.5	ENSG00000269893.6	SNHG8	lincRNA	19.5
ENST00000418403.1	ENSG00000237643.1	RP11-462G2.1	lincRNA	8.9
ENST00000530422.1	ENSG00000254911.3	SCARNA9	antisense	7.1
ENST00000499732.2	ENSG00000245532.5	NEAT1	lincRNA	5.9
ENST00000602946.1	ENSG00000269968.1	RP5-940J5.9	antisense	2.9
ENST00000601801.2	ENSG00000245532.5	NEAT1	lincRNA	0.6
ENST00000602607.1	ENSG00000269971.1	RP3-426I6.5	sense_intronic	0.1

Supplementary Table 5: Bootstrap frequency tables of the random sampling revalidation performed using penalized logistic/linear regression analysis. Bootstrap revalidation was performed 1000 times, wherein a different (~37%) subset of samples are eliminated a random in each penalized regression analysis. This was performed separately on the annotated lncRNAs for each criterion (islet purity, islet viability, beta-cell viability and transplantable outcome).

			Spearman R			P-value			
Transcript ID	**Gene ID**	**Official Symbol**	**Post-Iso Purity**	**Post-Iso Viability**	**Beta-cell Viability Index**	**Post-Iso Purity**	**Post-Iso Viability**	**Beta-cell Viability Index**	**Transplantable (yes/no)**
ENST00000619449.1	ENSG00000251562.7	MALAT1	0.35	0.31	0.59	0.20	0.25	0.03	0.01
ENST00000620902.1	ENSG00000251562.7	MALAT1	0.52	0.51	0.33	0.05	0.04	0.25	0.35
ENST00000508832.2	ENSG00000251562.7	MALAT1	0.28	0.37	0.55	0.31	0.15	0.05	0.05
ENST00000616691.1	ENSG00000251562.7	MALAT1	0.47	0.37	0.63	0.08	0.16	0.02	0.01

Supplementary Table 6: Univariate analysis on key MALAT1 splice variants to the individual key parameters (used to categorize the islet samples into three groups). Results presented in Spearman correlation r and (two-tailed) P-value. Results of transplantable (yes/no) are presented in T-test (two-tail, unequal variance) P-value.

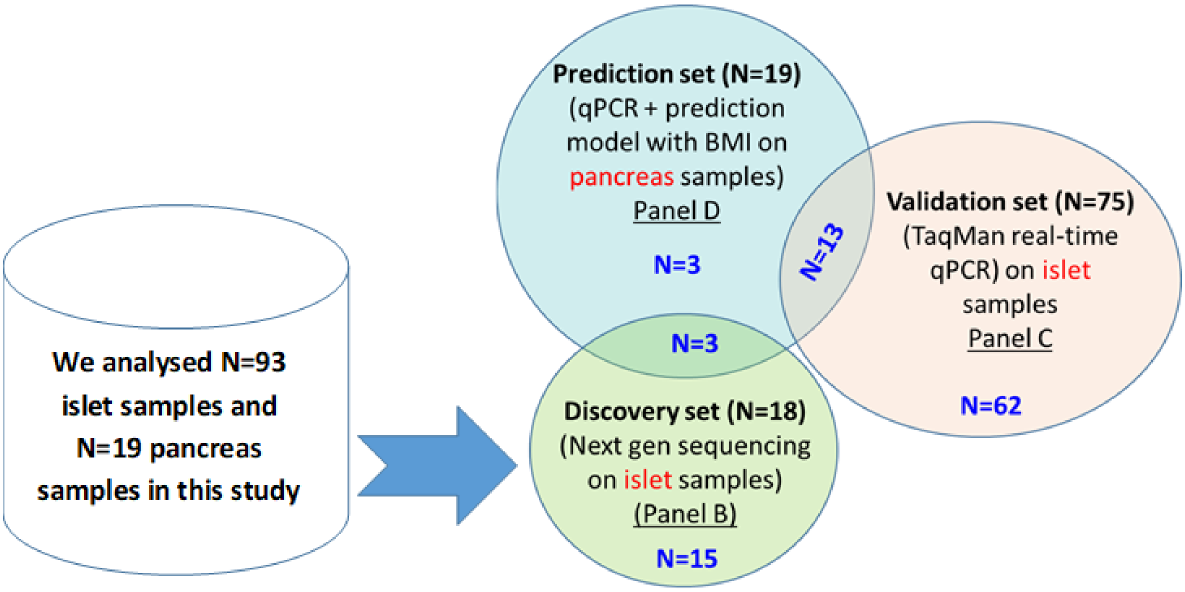
#	Sample name	Group	RNA integrity number (RIN)	Raw reads	Clean reads	Q20 (%)	Q30 (%)	GC content (%)
1	S164	1	8.4	78,802,166	77,339,700	97.44	93.53	55.41
2	S177	1	7.6	66,511,572	64,966,358	97.73	94.19	52.97
3	S321	1	8	72,222,866	70,996,486	97.49	93.65	52.22
4	S396	1	7.4	69,876,510	65,439,016	97.82	94.42	53.37
5	S712	1	7.5	82,376,566	79,133,946	97.42	93.67	47.03
6	S234	2	8.4	80,493,292	79,142,328	97.61	93.87	56.38
7	S263	2	7.9	80,505,416	79,142,530	97.56	93.79	53.56
8	S462	2	8.5	79,920,228	76,864,340	97.49	93.77	51.13
9	S574	2	8.1	87,752,734	82,251,092	97.5	93.76	52.81
10	S726	2	8.8	102,342,918	98,561,112	97.7	94.19	54.14
11	S75	2	7.7	79,893,996	77,589,020	96.37	91.09	52.89
12	S118	3	8.6	73,812,942	72,542,198	97.45	93.55	53.97
13	S178	3	8.7	76,429,010	75,167,708	97.54	93.74	52.98
14	S298	3	8.3	68,592,214	67,305,720	97.27	93.23	52.9
15	S458	3	7.9	76,175,086	71,963,744	97.53	93.94	53
16	S745	3	8.6	104,472,950	97,755,578	97.65	94.05	57.53
17	S975	3	8	80,372,466	74,648,616	97.54	93.94	52.01
18	S976	3	7.9	80,541,404	77,492,548	97.46	93.68	54.13

Supplementary Table 7. RNA integrity number, RNA-seq read counts, Quality (Q) score (Q20 and Q30) values and GC content of the human islet samples from discovery set (*n*=18).

`

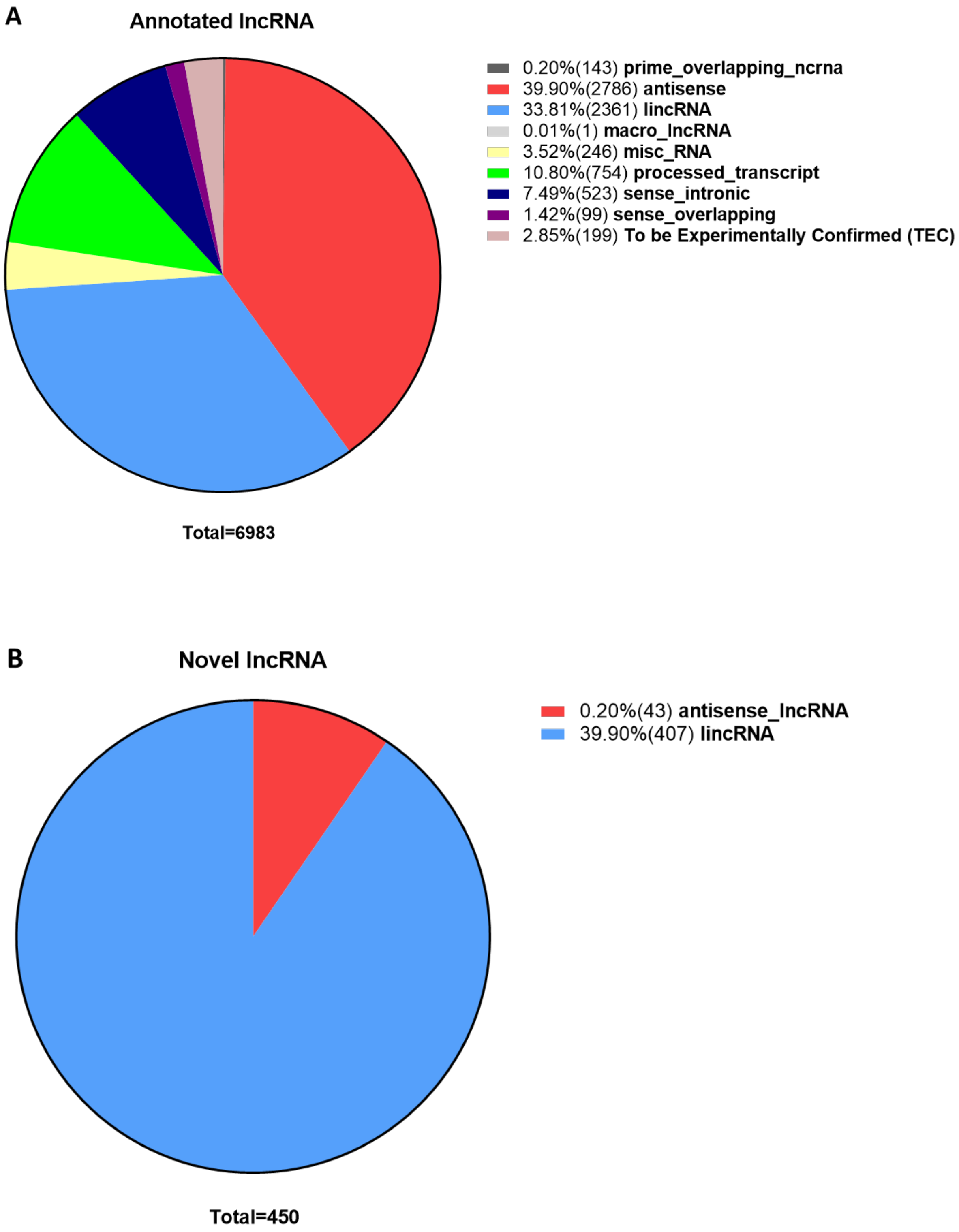
#	Assay ID	Gene symbol	Dye	Stock
1	Hs03003631_g1	18S	VIC-MGB	60x*
2	Hs01910177_s1 (“177”)	MALAT1	FAM-MGB	20x
3	Hs 00273907_s1 (“907”)	MALAT1	FAM-MGB	20x
4	Hs02741908_m1	INS	FAM-MGB	20x

Supplementary Table 8: List of the TaqMan® primer/probe gene expression assays (Thermo Fisher Scientific, Waltham, MA) used for real-time qPCR on the ViiA7 96-well platform. *A one in three dilution was performed to the stock 18s (60x) to match other assays concentration.



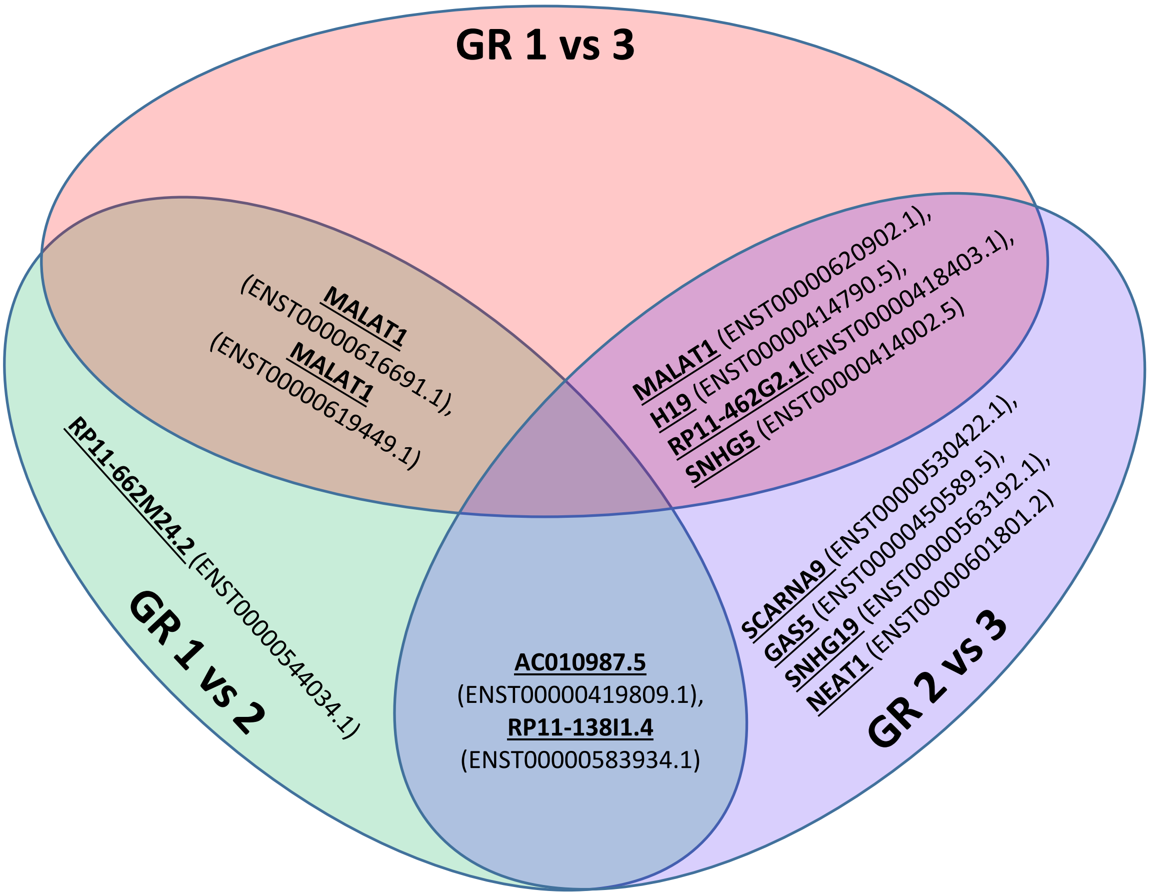
Supplementary Figure 1: Schematic diagram of workflow on the cadaveric donor pancreas/islets used in this study.
Through the biobank at Westmead Islet Transplant Center, we had access to *n*=93 human pancreatic islet samples and *n*=19 human pancreas samples. The numbers in blue color represent the number of cadaveric donors from whom islets or pancreas were available. None of the samples from the "Discovery set" (*n*=18) were used for "Validation set" (*n*=75). However, pancreas from 13 cadaveric donors who contributed islets to Validation set were used in the "Prediction set". Similarly, pancreas from *n*=3 discovery set donors were used in the prediction set. Discovery analysis (*n*=18 islet samples, **Supplementary Table 1**) was based on Next generation sequencing whilst the validation (*n*=75 islet samples, **Supplementary Table 2**) and prediction (*n*=19 pancreas samples, **Supplementary Table 3**) studies were based on TaqMan qPCR measurements of the two MALAT1 gene transcripts.

`

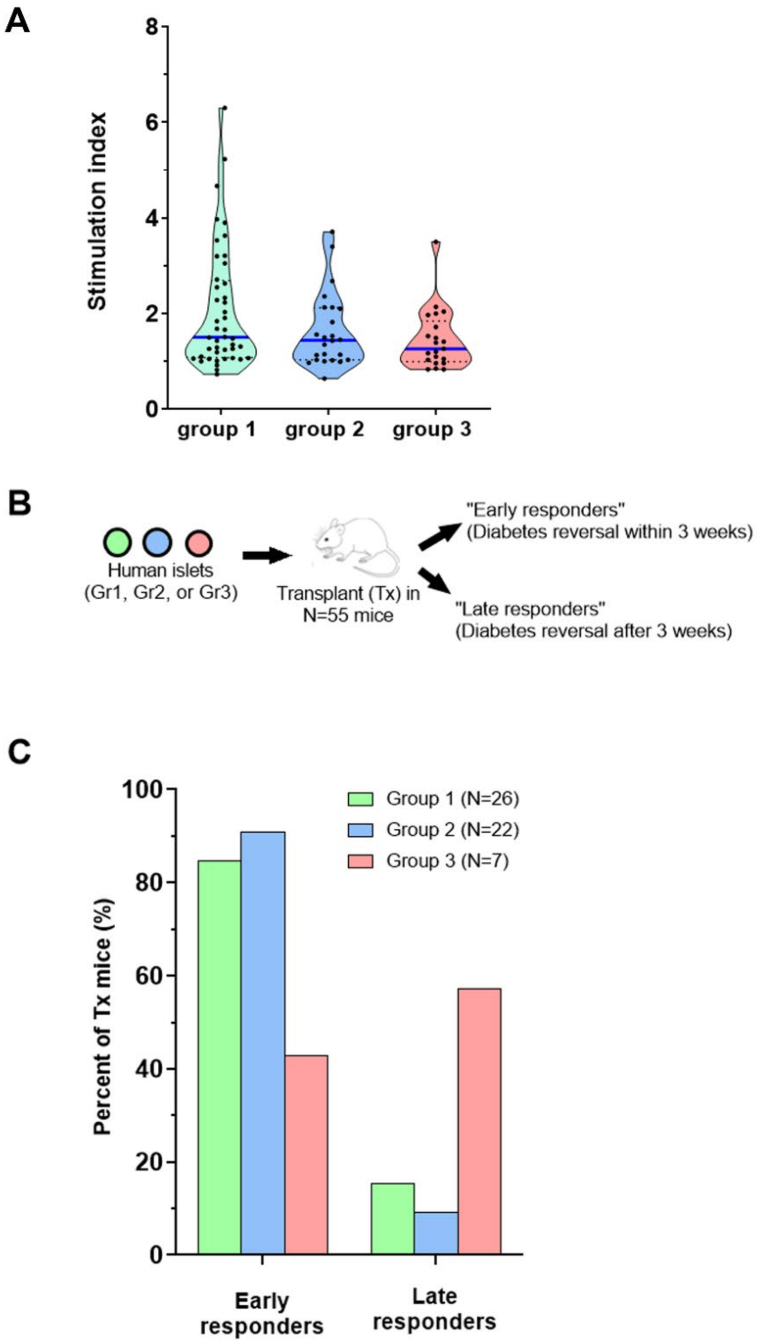


Supplementary Figure 2: A pie-chart detailing the types of annotated lncRNAs (**A**) and novel lncRNAs (**B**) identified through whole transcriptome sequencing of *n*=18 human islet preparations using the HiSeq4000, 150 Paired-end (PE) reads platform.

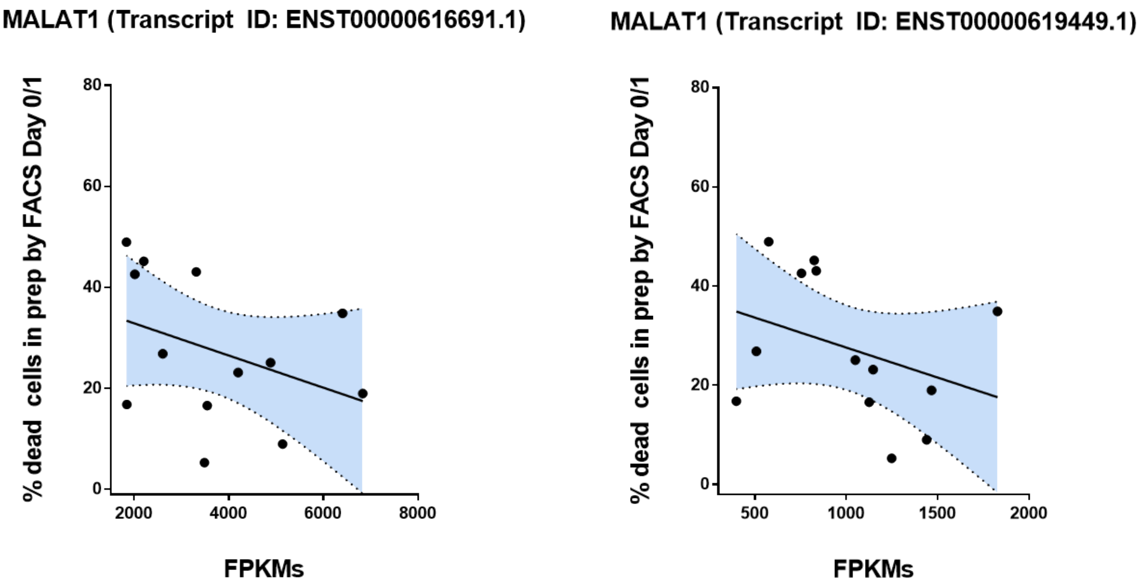
`



Supplementary Figure 3: Venn diagram of the set of lncRNAs from the bootstrap PLR group comparison analyses (**Supplementary Table 4).**



Supplementary Figure 4: The functionality of all human islet preparations was determined in *n*=89 of the available 93 samples using (**A**) the in vitro stimulation index obtained after GSIS and in vivo islet transplantation in diabetic animals. (**B**) Schematic diagram of the workflow for human islet transplantation in mice (**C**) Percent of diabetes reversal within three weeks (early responders) or after (late responders) in each group following transplantation in mice is calculated. Panel A: Each dot in the violin plot represents a different islet preparation. Horizontal solid blue line represents median for each group, the horizontal dotted line represents quartiles, while the polygons represent the density of individual data points and extend to min/max values.



Supplementary Figure 5: Flow cytometric assessment of isolated islet cells was carried out in 24 hours from pancreatic islet isolation and the proportion of dead cells assessed were plotted relative to the levels of MALAT1 variants (ENST00000616691.1 and ENST00000619449.1) measured in each of these islet preparations (r= -0.45, P-value= 0.13 and r= -0.42, P-value=0.16 respectively) . Islet cell death was negatively associated with MALAT1 transcript abundance for both the variants shown here. Please note that viability data was not available for *n*=5 samples and therefore data for remaining (*n*=13) samples are presented herein. Results are presented in Spearman correlation r and (two-tailed) P-value.