

Contents

1	Introduction	1
1.1	Cancer Research in the Post-Genomic Era	1
1.1.1	Cancer as a Global Health Concern	2
1.1.1.1	Genetics and Molecular Biology in Cancers	3
1.1.2	The Human Genome Revolution	5
1.1.2.1	The First Human Genome Sequence	5
1.1.2.2	Impact of Genomics	6
1.1.3	Technologies to Enable Genetics Research	7
1.1.3.1	DNA Sequencing and Genotyping Technologies	7
1.1.3.2	Microarrays and Quantitative Technologies	8
1.1.3.3	Massively Parallel “Next Generation” Sequencing	9
1.1.3.3.1	Molecular Profiling with Genomics Technology .	10
1.1.3.3.2	Established Sequencing Technologies	11
1.1.3.3.3	Emerging Sequencing Technologies	12
1.1.3.4	Bioinformatics as Interdisciplinary Genomic Analysis .	14
1.1.4	Follow-up Large-Scale Genomics Projects	14
1.1.5	Cancer Genomes	15
1.1.5.1	The Cancer Genome Atlas Project	16
1.1.5.2	The International Cancer Genome Consortium	17
1.1.5.2.1	Findings from Cancer Genomes	17
1.1.5.2.2	Genomic Comparisons Across Cancer Tissues .	19
1.1.5.2.3	Cancer Genomic Data Resources	20
1.1.6	Genomic Cancer Medicine	20
1.1.6.1	Cancer Genes and Driver Mutations	21
1.1.6.2	Personalised or Precision Cancer Medicine	22
1.1.6.2.1	Molecular Diagnostics and Pan-Cancer Medicine	22
1.1.6.3	Targeted Therapeutics and Pharmacogenomics	23
1.1.6.3.1	Targeting Oncogenic Driver Mutations	23
1.1.6.4	Systems and Network Biology	24
1.1.6.4.1	Network Medicine, and Polypharmacology	27
1.2	A Synthetic Lethal Approach to Cancer Medicine	28
1.2.1	Synthetic Lethal Genetic Interactions	28
1.2.2	Synthetic Lethal Concepts in Genetics	29
1.2.3	Studies of Synthetic Lethality	30
1.2.3.1	Synthetic Lethal Pathways and Networks	30
1.2.3.1.1	Evolution of Synthetic Lethality	31
1.2.4	Synthetic Lethal Concepts in Cancer	32
1.2.5	Clinical Impact of Synthetic Lethality in Cancer	33
1.2.6	High-throughput Screening for Synthetic Lethality	35
1.2.6.1	Synthetic Lethal Screens	37

1.2.7	Computational Prediction of Synthetic Lethality	40
1.2.7.1	Bioinformatics Approaches to Genetic Interactions	40
1.2.7.2	Comparative Genomics	41
1.2.7.3	Analysis and Modelling of Protein Data	44
1.2.7.4	Differential Gene Expression	46
1.2.7.5	Data Mining and Machine Learning	47
1.2.7.6	Bimodality	50
1.2.7.7	Rationale for Further Development	50
1.3	E-cadherin as a Synthetic Lethal Target	51
1.3.1	The <i>CDH1</i> gene and it's Biological Functions	51
1.3.1.1	Cytoskeleton	51
1.3.1.2	Extracellular and Tumour Micro-Environment	52
1.3.1.3	Cell-Cell Adhesion and Signalling	52
1.3.2	<i>CDH1</i> as a Tumour (and Invasion) Suppressor	52
1.3.2.1	Breast Cancers and Invasion	53
1.3.3	Hereditary Diffuse Gastric Cancer and Lobular Breast Cancer .	53
1.3.4	Somatic Mutations	54
1.3.4.1	Mutation Rate	54
1.3.4.2	Co-occurring Mutations	55
1.3.5	Models of <i>CDH1</i> loss in cell lines	56
1.4	Summary and Research Direction of Thesis	56
2	Methods and Resources	61
2.1	Bioinformatics Resources for Genomics Research	61
2.1.1	Public Data and Software Packages	61
2.1.1.1	Cancer Genome Atlas Data	62
2.1.1.2	Reactome and Annotation Data	63
2.2	Data Handling	63
2.2.1	Normalisation	63
2.2.2	Sample Triage	65
2.2.3	Metagenes and the Singular Value Decomposition	65
2.2.3.1	Candidate Triage and Integration with Screen Data .	66
2.3	Techniques	67
2.3.1	Statistical Procedures and Tests	67
2.3.2	Gene Set Over-representation Analysis	68
2.3.3	Clustering	68
2.3.4	Heatmap	68
2.3.5	Modeling and Simulations	69
2.3.5.1	Receiver Operating Characteristic (Performance) .	70
2.3.6	Resampling Analysis	70
2.4	Pathway Structure Methods	71
2.4.1	Network and Graph Analysis	71
2.4.2	Sourcing Graph Structure Data	72
2.4.3	Constructing Pathway Subgraphs	72
2.4.4	Network Analysis Metrics	72
2.5	Implementation	74

2.5.1	Computational Resources and Linux Utilities	74
2.5.2	R Language and Packages	74
2.5.3	High Performance and Parallel Computing	77
3	Methods Developed During Thesis	79
3.1	A Synthetic Lethal Detection Methodology	79
3.2	Synthetic Lethal Simulation and Modelling	81
3.2.1	A Model of Synthetic Lethality in Expression Data	82
3.2.2	Simulation Procedure	86
3.3	Detecting Simulated Synthetic Lethal Partners	87
3.3.1	Binomial Simulation of Synthetic lethality	89
3.3.2	Multivariate Normal Simulation of Synthetic lethality	91
3.3.2.1	Multivariate Normal Simulation with Correlated Genes	93
3.3.2.2	Specificity with Query-Correlated Pathways	99
3.3.2.2.1	Importance of Directional Testing	101
3.4	Graph Structure Methods	103
3.4.1	Upstream and Downstream Gene Detection	103
3.4.1.1	Permutation Analysis for Statistical Significance	104
3.4.1.2	Ranking Based on Biological Context	104
3.4.2	Simulating Gene Expression from Graph Structures	105
3.5	Customised Functions and Packages Developed	110
3.5.1	Synthetic Lethal Interaction Prediction Tool	110
3.5.2	Data Visualisation	111
3.5.3	Extensions to the iGraph Package	111
3.5.3.1	Sampling Simulated Data from Graph Structures	113
3.5.3.2	Plotting Directed Graph Structures	113
3.5.3.3	Computing Information Centrality	113
3.5.3.4	Testing Pathway Structure with Permutation Testing	114
3.5.3.5	Metapackage to Install iGraph Functions	114
4	Synthetic Lethal Analysis of Gene Expression Data	6
4.1	Synthetic lethal genes in breast cancer	8
4.1.1	Synthetic lethal pathways in breast cancer	10
4.1.2	Expression profiles of synthetic lethal partners	10
4.1.2.1	Subgroup pathway analysis	11
4.2	Comparison of synthetic lethal gene candidates	13
4.2.1	Comparison with differential expression	13
4.2.2	Comparison with correlation	13
4.2.3	Comparison with primary siRNA screen candidates	13
4.2.3.1	Comparison of screen at pathway level	16
4.2.3.1.1	Resampling of genes for pathway enrichment	16
4.2.4	Comparison with secondary screen siRNA screen candidates	20
4.2.4.1	Comparison of candidate SL Pathways	20
4.3	Mutation, Copy Number, and Methylation	20
4.3.1	Synthetic lethality by DNA copy number	22
4.3.2	Synthetic lethality by somatic mutation	22

4.3.2.1	Mutation analysis	22
4.3.3	ANOVA of Expression Predictors	22
4.4	Global Synthetic Lethality	23
4.4.1	Hub Genes	23
4.5	Metagene Analysis	23
4.5.1	Pathway expression	23
4.5.2	Somatic mutation	23
4.5.3	Synthetic lethal metagenes	23
4.6	Replication in stomach cancer	23
4.7	Replication in cell line encyclopaedia	24
4.8	Summary	26
5	Synthetic Lethal Pathway Structure	142
5.1	Reactome Network structure and Information Centrality as a measure of gene essentiality	143
5.2	Synthetic lethal genes in synthetic lethal pathways	143
5.3	Centrality and connectivity of synthetic lethal genes	143
5.4	Upstream or downstream synthetic lethal candidates	143
5.5	Hierachical approach	143
5.6	Discussion	143
5.7	Conclusion	143
6	Simulation and Modeling of Synthetic Lethal Pathways	144
6.1	Simulations and Modelling Synthetic Lethality in Expression Data . . .	146
6.2	Simulations over simple graph structures	147
6.2.1	Performance	147
6.2.2	Synthetic lethality across graph stuctures	147
6.2.3	Performance with inhibition links	147
6.2.4	Performance with 20,000 genes	147
6.3	Simulations over pathway-based graphs	147
6.4	Comparing methods	147
6.4.1	SLIPT and Chi-Squared	147
6.4.1.1	Correlated query genes	147
6.4.2	Correlation	147
6.4.3	Bimodality with BiSEp	147
7	Discussion	148
7.1	Significance	150
7.2	Future Directions	151
7.3	Conclusion	152
8	Conclusion	153
A	Sample Quality	154
A.1	Sample Correlation	154
A.2	Replicate Samples in TCGA Breast	156

B Software Used for Thesis	160
C Secondary Screen Data	169
D Mutation Analysis in Breast Cancer	171
D.1 Synthetic Lethal Genes and Pathways	171
D.2 Comparison to Primary Screen	175
D.3 Resampling Analysis	177
D.4 Metagene Analysis	179
D.5 Mutation Variation	180
D.5.1 Mutation Frequency	180
D.5.2 PI3K Mutation Expression	181
D.6 Compare SLIPT genes	186
E Expression Analysis in Stomach Cancer	187
F Mutation Analysis in Stomach Cancer	188

List of Figures

1.1	Synthetic genetic interactions	29
1.2	Synthetic lethality in cancer	33
2.1	Read count density	64
2.2	Read count sample mean	64
3.1	Framework for synthetic lethal prediction	80
3.2	Synthetic lethal prediction adapted for mutation	81
3.3	A model of synthetic lethal gene expression	83
3.4	Modeling synthetic lethal gene expression	84
3.5	Synthetic lethality with multiple genes	85
3.6	Simulating gene function	87
3.7	Simulating synthetic lethal gene function	88
3.8	Simulating synthetic lethal gene expression	88
3.9	Performance of binomial simulations	90
3.10	Comparison of statistical performance	90
3.11	Performance of multivariate normal simulations	92
3.12	Simulating expression with correlated gene blocks	95
3.13	Simulating expression with correlated gene blocks	96
3.14	Synthetic lethal prediction across simulations	97
3.15	Performance with correlations	98
3.16	Comparison of statistical performance with correlation structure	99
3.17	Performance with query correlations	100
3.18	Statistical evaluation of directional criteria	101
3.19	Performance of directional criteria	102
3.20	Simulated graph structures	106
3.21	Simulating expression from a graph structure	107
3.22	Simulating expression from graph structure with inhibitions	109
3.23	Demonstration of violin plots with custom features	112
3.24	Demonstration of annotated heatmap	112
3.25	Simulating graph structures	114
4.1	Synthetic lethal expression profiles of analysed samples	12
4.2	Comparison of SLIPT to siRNA	15
4.3	Resampled intersection of SLIPT and siRNA candidates	18
4.4	Synthetic lethal expression profiles of stomach samples	31
4.5	Comparison of SLIPT in stomach to siRNA	32
A.1	Correlation profiles of removed samples	154
A.2	Correlation analysis and sample removal	155
A.3	Replicate excluded samples	156

A.4	Replicate samples with all remaining	157
A.5	Replicate samples with some excluded	158
A.5	Replicate samples with some excluded	159
D.1	Synthetic lethal expression profiles of analysed samples	173
D.2	Comparison of mtSLIPT to siRNA	175
D.3	Somatic mutation locus	180
D.4	Somatic mutation against expression	181
D.5	Somatic mutation against PI3K protein	182
D.6	Somatic mutation against AKT protein	183
D.7	Somatic mutation against PI3K metagene	184
D.8	Somatic mutation against PIK3CA metagene	185
D.9	Compare mtSLIPT and siRNA genes with correlation	186
D.10	Compare mtSLIPT and siRNA genes with siRNA viability	186
F.1	Synthetic lethal expression profiles of stomach samples	190
F.2	Comparison of mtSLIPT in stomach to siRNA	192

List of Tables

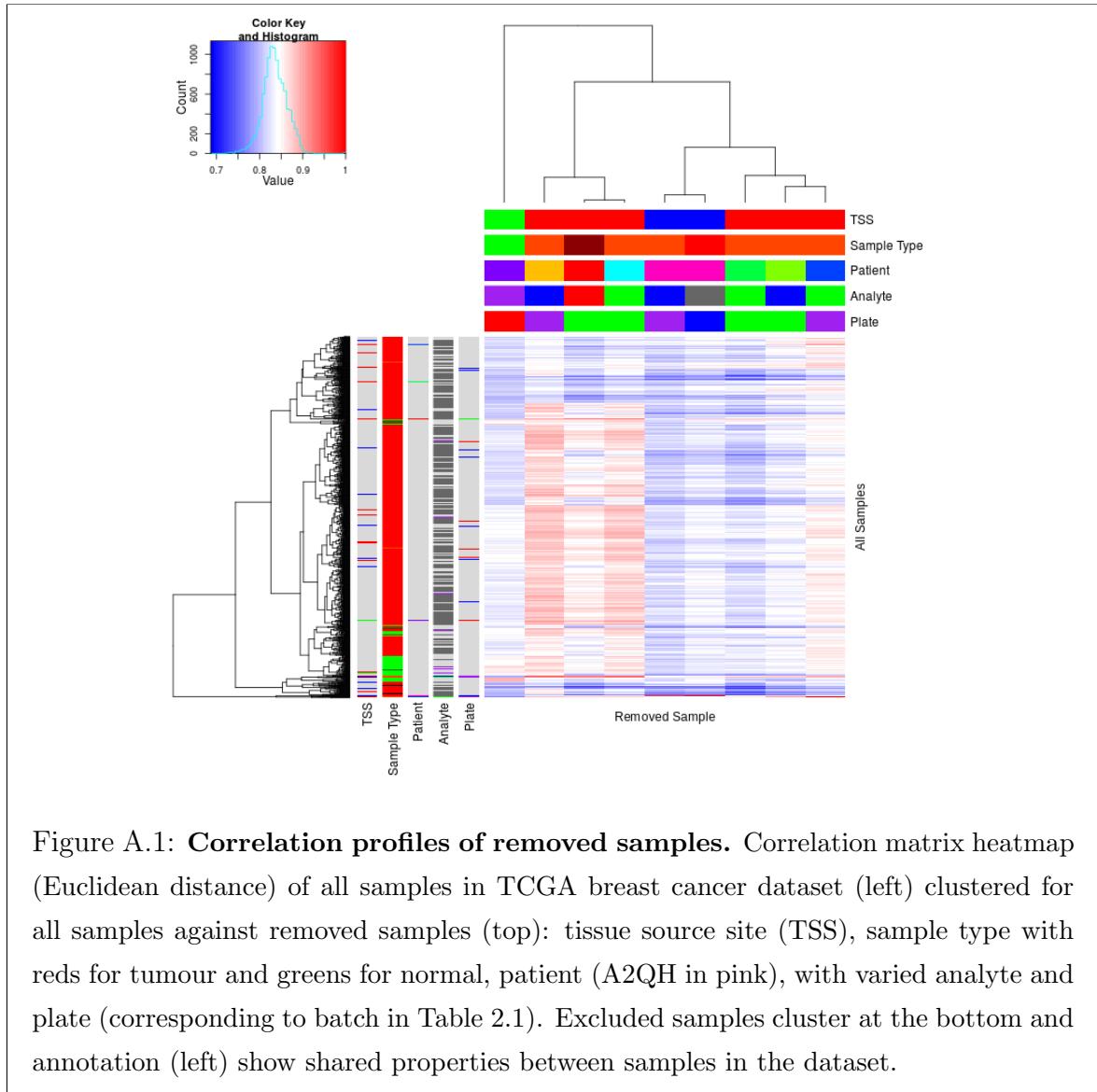
1.1	Methods for Predicting Genetic Interactions	40
1.2	Methods for Predicting Synthetic Lethality in Cancer	41
1.3	Methods used by ?	43
2.1	Excluded Samples by Batch and Clinical Characteristics.	65
2.2	Computers used during Thesis	74
2.3	Linux Utilities and Applications used during Thesis	75
2.4	R Installations used during Thesis	75
2.5	R Packages Developed during Thesis	76
2.6	R Packages used during Thesis	76
4.1	Candidate synthetic lethal genes against E-cadherin from SLIPT	9
4.2	Pathways for <i>CDH1</i> partners from SLIPT	10
4.3	Pathway composition for clusters of <i>CDH1</i> partners from SLIPT	14
4.4	Pathway composition for <i>CDH1</i> partners from SLIPT and siRNA screening	17
4.5	Pathways for <i>CDH1</i> partners from SLIPT	19
4.6	Pathways for <i>CDH1</i> partners from SLIPT and siRNA primary screen	21
4.7	Candidate synthetic lethal metagenes against <i>CDH1</i> from SLIPT	24
4.8	Candidate synthetic lethal genes against E-cadherin from SLIPT in stomach cancer	25
4.9	Pathways for <i>CDH1</i> partners from SLIPT in stomach cancer	26
4.10	Pathway composition for clusters of <i>CDH1</i> partners in stomach SLIPT	27
4.11	Pathway composition for <i>CDH1</i> partners from SLIPT and siRNA screening	28
4.12	Pathways for <i>CDH1</i> partners from SLIPT in stomach cancer	29
4.13	Pathways for <i>CDH1</i> partners from SLIPT in stomach and siRNA screen	30
4.14	Candidate synthetic lethal metagenes against <i>CDH1</i> from SLIPT in stomach cancer	33
4.15	Candidate synthetic lethal genes against E-cadherin from SLIPT in CCLE	34
4.16	Pathways for <i>CDH1</i> partners from SLIPT in CCLE	35
4.17	Candidate synthetic lethal genes against E-cadherin from SLIPT in breast CCLE	36
4.18	Pathways for <i>CDH1</i> partners from SLIPT in breast CCLE	37
4.19	Candidate synthetic lethal genes against E-cadherin from SLIPT in stomach CCLE	38
4.20	Pathways for <i>CDH1</i> partners from SLIPT in stomach CCLE	39
B.1	R Packages used during Thesis	160
C.1	Comparing SLIPT genes against Secondary siRNA Screen in breast cancer	169

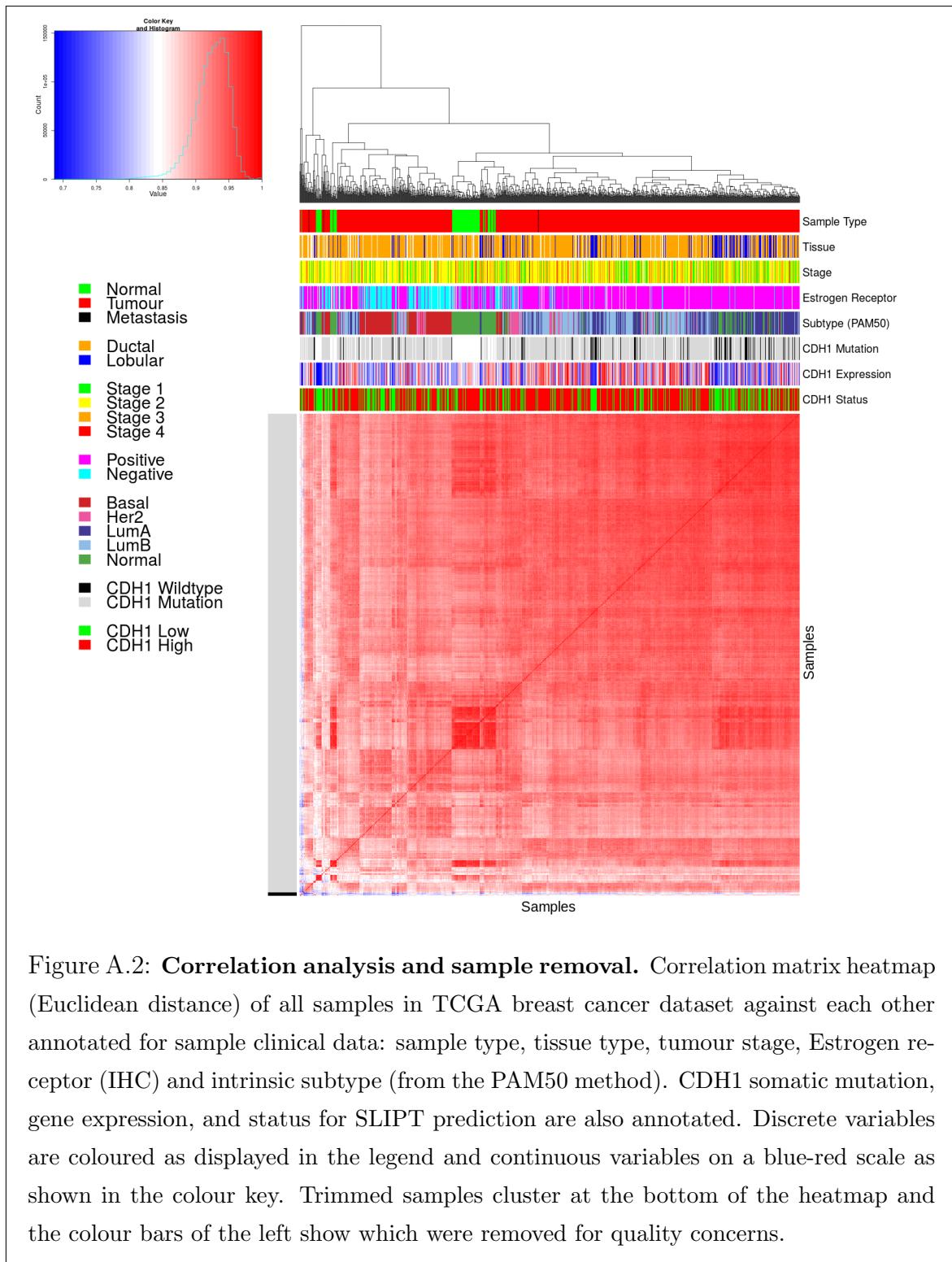
C.2	Comparing mtSLIPT genes against Secondary siRNA Screen in breast cancer	170
C.3	Comparing SLIPT genes against Secondary siRNA Screen in stomach cancer	170
D.1	Candidate synthetic lethal genes against E-cadherin from mtSLIPT	171
D.2	Pathways for <i>CDH1</i> partners from mtSLIPT	172
D.3	Pathway composition for clusters of <i>CDH1</i> partners from mtSLIPT	174
D.4	Pathway composition for <i>CDH1</i> partners from mtSLIPT and siRNA	176
D.5	Pathways for <i>CDH1</i> partners from mtSLIPT	177
D.6	Pathways for <i>CDH1</i> partners from mtSLIPT and siRNA primary screen	178
D.7	Candidate synthetic lethal metagenes against <i>CDH1</i> from mtSLIPT	179
F.1	Candidate synthetic lethal genes against E-cadherin from mtSLIPT in stomach cancer	188
F.2	Pathways for <i>CDH1</i> partners from mtSLIPT in stomach cancer	189
F.3	Pathway composition for clusters of <i>CDH1</i> partners in stomach mtSLIPT	191
F.4	Pathway composition for <i>CDH1</i> partners from mtSLIPT and siRNA	193
F.5	Pathways for <i>CDH1</i> partners from mtSLIPT in stomach cancer	194
F.6	Pathways for <i>CDH1</i> partners from mtSLIPT in stomach and siRNA screen	195
F.7	Candidate synthetic lethal metagenes against <i>CDH1</i> from mtSLIPT in stomach cancer	196

Appendix A

Sample Quality

A.1 Sample Correlation





A.2 Replicate Samples in TCGA Breast

Replicate samples were picked where possible from the TCGA breast cancer gene expression data to examine for sample quality. Independent samples of the same tumour are expected to have very high Pearson's correlation between their expression profiles unless there were issues with sample collection or preparation and are thus an indicator of sample quality. The log-transformed raw read counts for replicate samples were examined in Figures A.3–A.5. These were examined before normalisation which would be expected to increase sample concordance.

Another consideration are the samples which were removed for quality concerns (in section 2.2.2). While these were selected by unbiased hierarchical clustering (See Figure A.2), it is notable that many of the excluded (tumour) samples were performed in replicate despite relatively few replicate samples in the overall dataset. These samples correlate poorly with the rest of the dataset, in addition to with replicate samples.

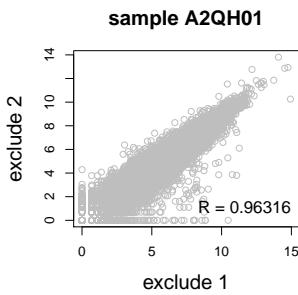


Figure A.3: **Replicate excluded samples.** Both tumour samples of patient A2QH were excluded as they were poorly correlated with other samples, although they are highly similar to each other as shown by Pearson's correlation of log-raw counts.

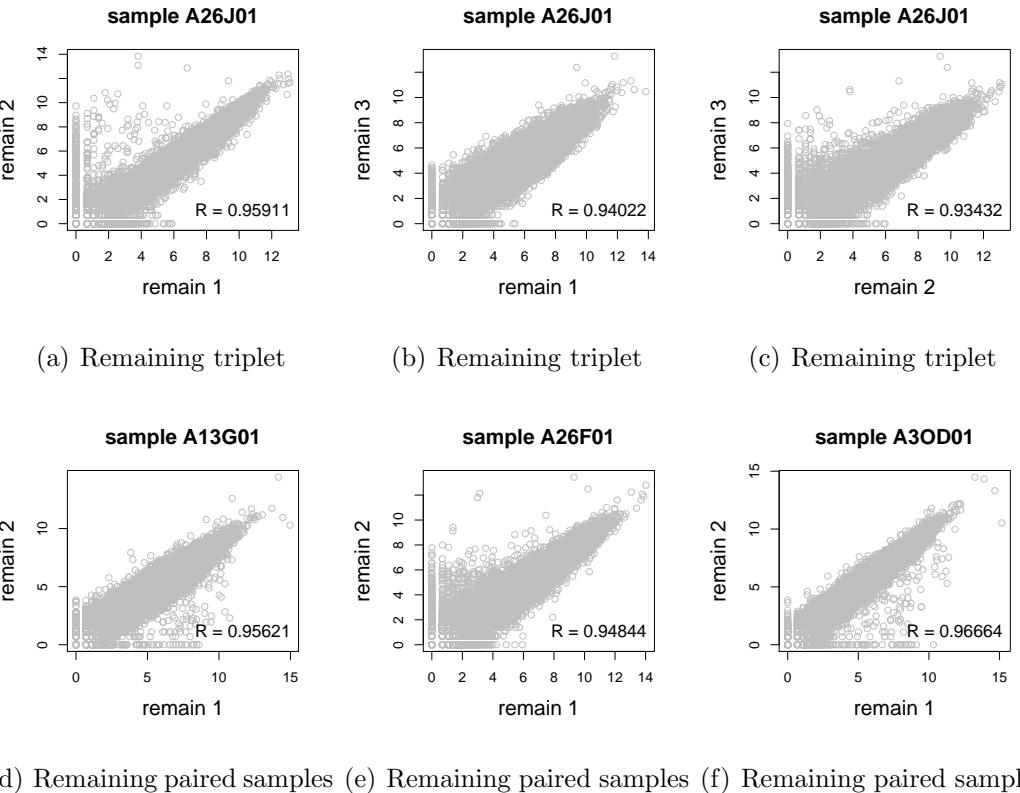


Figure A.4: **Replicate samples with all remaining.** Patient A26J was sampled 3 times and compared pairwise. Pairs of samples were also compared for other patients with replicate samples. In all cases, replicate samples remaining in the dataset were highly concordant as shown by Pearson's correlation of log-raw counts.

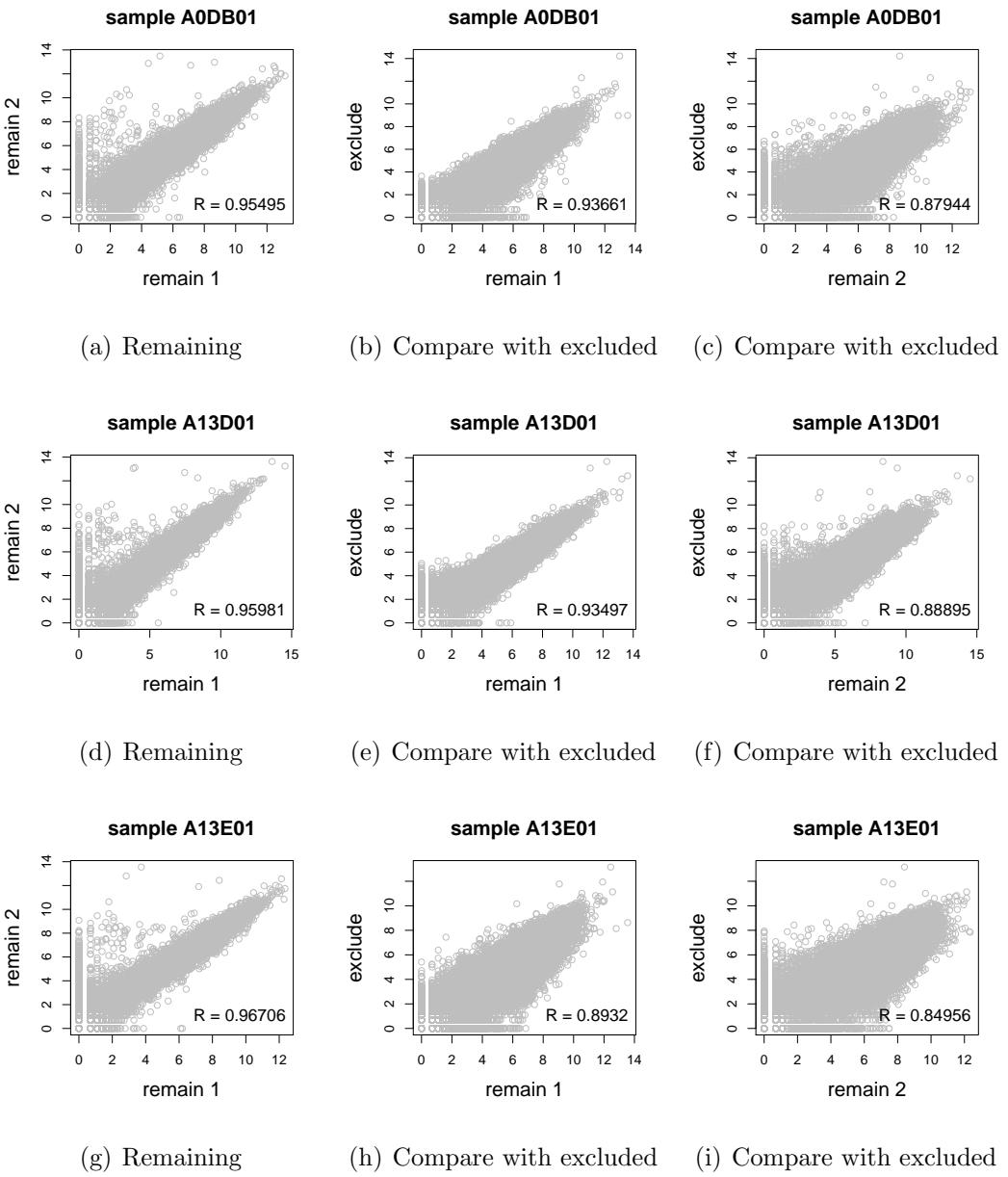


Figure A.5: Replicate samples with some excluded. Patients A0DB, A13D, A13E, and A26E were each sampled 3 times and compared pairwise. Pairs of samples were also compared for other patients with replicate samples. In all cases, the replicate samples remaining in the dataset more were highly concordant (as shown by Pearson's correlation of log-log counts) than those excluded from the analysis.

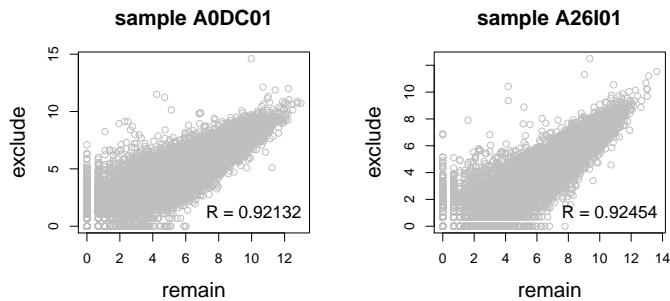
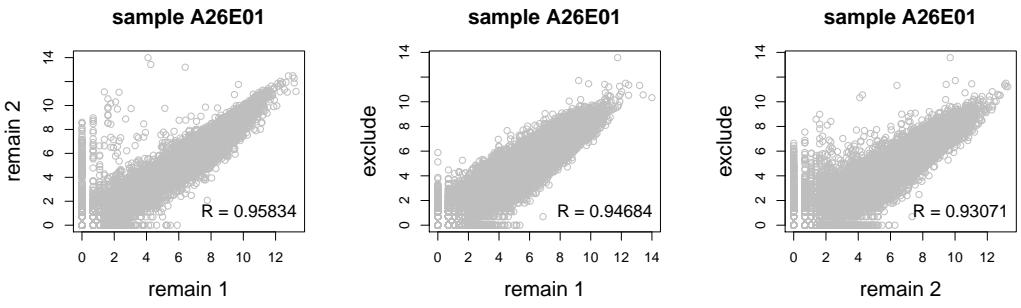


Figure A.5: Replicate samples with some excluded. Patients A0DB, A13D, A13E, and A26E were each sampled 3 times and compared pairwise. Pairs of samples were also compared for other patients with replicate samples. In all cases, the replicate samples remaining in the dataset more were highly concordant (as shown by Pearson's correlation of log-raw counts) than those excluded from the analysis.

Appendix B

Software Used for Thesis

Table B.1: R Packages used during Thesis

Package	Repository	Laptop	Lab	Server	NeSI
base	base	3.3.2	3.3.2	3.3.1	3.3.0
abind	CRAN		1.4-5		1.4-3
acepack	CRAN		1.4.1		1.3-3.3
ade4	CRAN		1.7-5		
annaffy	Bioconductor		1.46.0		
AnnotationDbi	Bioconductor		1.36.0	1.36.0	1.34.4
apComplex	CRAN		2.40.0		
ape	CRAN		4		3.4
arm	CRAN		1.9-3		
assertthat	CRAN	0.1	0.1	0.1	0.1
backports	CRAN	1.0.5	1.0.4	1.0.5	1.0.2
base64	CRAN			2	2
base64enc	CRAN		0.1-3		0.1-3
beanplot	CRAN		1.2	1.2	1.2
BH	CRAN	1.60.0-2	1.62.0-1	1.62.0-1	1.60.0-2
Biobase	Bioconductor		2.34.0	2.34.0	2.32.0
BiocGenerics	Bioconductor		0.20.0	0.20.0	0.18.0
BiocInstaller	Bioconductor		1.24.0	1.20.3	1.22.3
BiocParallel	Bioconductor		1.8.1	1.8.1	
Biostings	Bioconductor		2.42.1	2.42.0	
BiSEp	Bioconductor		2.0.1	2.0.1	2.0.1
bitops	CRAN	1.0-6	1.0-6	1.0-6	1.0-6
boot	base	1.3-18	1.3-18	1.3-18	1.3-18
brew	CRAN	1.0-6	1.0-6	1.0-6	1.0-6
broom	CRAN	0.4.1			

caTools	CRAN	1.17.1	1.17.1	1.17.1	1.17.1
cgdsr	CRAN		1.2.5		
checkmate	CRAN		1.8.2		1.7.4
chron	CRAN	2.3-47	2.3-48	2.3-50	2.3-47
class	base	7.3-14	7.3-14	7.3-14	7.3-14
cluster	base	2.0.5	2.0.5	2.0.5	2.0.4
coda	CRAN		0.19-1		0.18-1
codetools	base	0.2-15	0.2-15	0.2-15	0.2-14
colorRamps	CRAN		2.3		
colorspace	CRAN	1.2-6	1.3-2	1.3-2	1.2-6
commonmark	CRAN	1.1		1.2	
compiler	base	3.3.2	3.3.2	3.3.1	3.3.0
corpcor	CRAN		1.6.8	1.6.8	1.6.8
Cprob	CRAN		1.2.4		
crayon	CRAN	1.3.2	1.3.2	1.3.2	1.3.2
crop	CRAN		0.0-2	0.0-2	
curl	CRAN	1.2	2.3	2.3	0.9.7
d3Network	CRAN		0.5.2.1		
data.table	CRAN	1.9.6	1.10.0	1.10.1	1.9.6
data.tree	CRAN		0.7.0	0.7.0	
datasets	base	3.3.2	3.3.2	3.3.1	3.3.0
DBI	CRAN	0.5-1	0.5-1	0.5-1	0.5-1
dendextend	CRAN	1.4.0	1.4.0	1.4.0	
DEoptimR	CRAN	1.0-8	1.0-8	1.0-8	1.0-4
desc	CRAN	1.1.0		1.1.0	
devtools	CRAN	1.12.0	1.12.0	1.12.0	1.12.0
DiagrammeR	CRAN		0.9.0	0.9.0	
dichromat	CRAN	2.0-0	2.0-0	2.0-0	2.0-0
digest	CRAN	0.6.10	0.6.11	0.6.12	0.6.9
diptest	CRAN	0.75-7	0.75-7	0.75-7	
doParallel	CRAN	1.0.10	1.0.10	1.0.10	1.0.10
dplyr	CRAN	0.5.0	0.5.0	0.5.0	0.5.0
ellipse	CRAN		0.3-8	0.3-8	0.3-8
evaluate	CRAN		0.1	0.1	0.9
fdrtool	CRAN		1.2.15		

fields	CRAN		8.1		
flexmix	CRAN	2.3-13	2.3-13	2.3-13	
forcats	CRAN	0.2.0			
foreach	CRAN	1.4.3	1.4.3	1.4.3	1.4.3
foreign	base	0.8-67	0.8-67	0.8-67	0.8-66
formatR	CRAN		1.4	1.4	1.4
Formula	CRAN		1.2-1		1.2-1
fpc	CRAN	2.1-10	2.1-10	2.1-10	
futile.logger	CRAN		1.4.3	1.4.3	1.4.1
futile.options	CRAN		1.0.0	1.0.0	1.0.0
gdata	CRAN	2.17.0	2.17.0	2.17.0	2.17.0
geepack	CRAN		1.2-1		
GenomeInfoDb	Bioconductor		1.10.2	1.10.1	
GenomicAlignments	Bioconductor		1.10.0	1.10.0	
GenomicRanges	Bioconductor		1.26.2	1.26.1	
ggm	CRAN		2.3		
ggplot2	CRAN	2.1.0	2.2.1	2.2.1	2.1.0
git2r	CRAN	0.15.0	0.18.0	0.16.0	0.15.0
glasso	CRAN		1.8		
GO.db	Bioconductor		3.4.0	3.2.2	3.3.0
GOSemSim	Bioconductor		2.0.3	1.28.2	1.30.3
gplots	CRAN	3.0.1	3.0.1	3.0.1	3.0.1
graph	Bioconductor		1.52.0		
graphics	base	3.3.2	3.3.2	3.3.1	3.3.0
graphsim	GitHub TomKellyGenetics	0.1.0	0.1.0	0.1.0	0.1.0
grDevices	base	3.3.2	3.3.2	3.3.1	3.3.0
grid	base	3.3.2	3.3.2	3.3.1	3.3.0
gridBase	CRAN	0.4-7	0.4-7	0.4-7	0.4-7
gridExtra	CRAN	2.2.1	2.2.1	2.2.1	2.2.1
gridGraphics	CRAN		0.1-5		
gtable	CRAN	0.2.0	0.2.0	0.2.0	0.2.0
gtools	CRAN	3.5.0	3.5.0	3.5.0	3.5.0
haven	CRAN	1.0.0			

heatmap.2x	GitHub TomKellyGenetics	0.0.0.9000	0.0.0.9000	0.0.0.9000	0.0.0.9000
hgu133plus2.db	Bioconductor	3.2.3			
highr	CRAN	0.6	0.6	0.6	
Hmisc	CRAN		4.0-2	4.0-2	3.17-4
hms	CRAN	0.2	0.3		
htmlTable	CRAN		1.8	1.9	
htmltools	CRAN	0.3.5	0.3.5	0.3.5	0.3.5
htmlwidgets	CRAN		0.8	0.8	
httpuv	CRAN	1.3.3		1.3.3	
httr	CRAN	1.2.1	1.2.1	1.2.1	1.1.0
huge	CRAN		1.2.7		
hunspell	CRAN		2.3		2
hypergraph	CRAN		1.46.0		
igraph	CRAN	1.0.1	1.0.1	1.0.1	1.0.1
igraph.extensions	GitHub TomKellyGenetics	0.1.0.9001	0.1.0.9001	0.1.0.9001	0.1.0.9001
influenceR	CRAN		0.1.0	0.1.0	
info.centrality	GitHub TomKellyGenetics	0.1.0	0.1.0	0.1.0	0.1.0
IRanges	Bioconductor		2.8.1	2.8.1	2.6.1
irlba	CRAN	2.1.1	2.1.2	2.1.2	2.0.0
iterators	CRAN	1.0.8	1.0.8	1.0.8	1.0.8
jpeg	CRAN		0.1-8		
jsonlite	CRAN	1.1	1.2	1.3	0.9.20
KEGG.db	Bioconductor		3.2.3		
kernlab	CRAN	0.9-25	0.9-25	0.9-25	
KernSmooth	base	2.23-15	2.23-15	2.23-15	2.23-15
knitr	CRAN		1.15.1	1.15.1	1.14
labeling	CRAN	0.3	0.3	0.3	0.3
lambda.r	CRAN		1.1.9	1.1.9	1.1.7
lattice	base	0.20-34	0.20-34	0.20-34	0.20-33
latticeExtra	CRAN		0.6-28		0.6-28
lava	CRAN		1.4.6		
lavaan	CRAN		0.5-22		

lazyeval	CRAN	0.2.0	0.2.0	0.2.0	0.2.0
les	CRAN		1.24.0		
lgtdl	CRAN		1.1.3		
limma	Bioconductor		3.30.7	3.30.3	
lme4	CRAN		1.1-12		1.1-12
lubridate	CRAN	1.6.0			
magrittr	CRAN	1.5	1.5	1.5	1.5
maps	CRAN		3.1.1		
markdown	CRAN		0.7.7	0.7.7	0.7.7
MASS	base	7.3-45	7.3-45	7.3-45	7.3-45
Matrix	base	1.2-7.1	1.2-7.1	1.2-8	1.2-6
matrixcalc	CRAN	1.0-3	1.0-3	1.0-3	1.0-3
mclust	CRAN	5.2	5.2.1	5.2.2	5.2
memoise	CRAN	1.0.0	1.0.0	1.0.0	1.0.0
methods	base	3.3.2	3.3.2	3.3.1	3.3.0
mgcv	base	1.8-16	1.8-16	1.8-17	1.8-12
mi	CRAN		1		
mime	CRAN	0.5	0.5	0.5	0.4
minqa	CRAN		1.2.4		1.2.4
mnormt	CRAN	1.5-5	1.5-5		1.5-4
modelr	CRAN	0.1.0			
modeltools	CRAN	0.2-21	0.2-21	0.2-21	
multtest	Bioconductor		2.30.0	2.30.0	
munsell	CRAN	0.4.3	0.4.3	0.4.3	0.4.3
mvtnorm	CRAN	1.0-5	1.0-5	1.0-6	1.0-5
network	CRAN		1.13.0		
nlme	base	3.1-128	3.1-128	3.1-131	3.1-128
nloptr	CRAN		1.0.4		1.0.4
NMF	CRAN	0.20.6	0.20.6	0.20.6	0.20.6
nnet	base	7.3-12	7.3-12	7.3-12	7.3-12
numDeriv	CRAN		2016.8-1		2014.2-1
openssl	CRAN	0.9.4	0.9.6	0.9.6	0.9.4
org.Hs.eg.db	Bioconductor		3.1.2		3.3.0
org.Sc.sgd.db	Bioconductor		3.4.0		
parallel	base	3.3.2	3.3.2	3.3.1	3.3.0

pathway.structure	GitHub		0.1.0	0.1.0	0.1.0	0.1.0
.permutation	TomKellyGenetics					
pbivnorm	CRAN		0.6.0			
PGSEA	Bioconductor		1.48.0			
pkgmaker	CRAN	0.22	0.22	0.22	0.22	
PKI	CRAN		0.1-3			
plogr	CRAN		0.1-1	0.1-1		
plot.igraph	GitHub		0.0.0.9001	0.0.0.9001	0.0.0.9001	0.0.0.9001
	TomKellyGenetics					
plotrix	CRAN		3.6-4			
plyr	CRAN	1.8.4	1.8.4	1.8.4	1.8.4	1.8.3
png	CRAN		0.1-7			0.1-7
prabclus	CRAN	2.2-6	2.2-6	2.2-6		
praise	CRAN	1.0.0	1.0.0			1.0.0
pROC	CRAN		1.8	1.9.1		
prodlim	CRAN		1.5.7			
prof.tree	CRAN		0.1.0			
protoools	CRAN		0.99-2			
progress	CRAN			1.1.2		
psych	CRAN	1.6.12	1.6.12			
purrr	CRAN	0.2.2	0.2.2	0.2.2	0.2.2	
qgraph	CRAN		1.4.1			
quadprog	CRAN		1.5-5	1.5-5	1.5-5	
R.methodsS3	CRAN		1.7.1			1.7.1
R.oo	CRAN		1.21.0			1.20.0
R.utils	CRAN		2.5.0			
R6	CRAN	2.1.3	2.2.0	2.2.0	2.1.3	
RBGL	CRAN		1.50.0			
RColorBrewer	CRAN	1.1-2	1.1-2	1.1-2	1.1-2	
Rcpp	CRAN	0.12.7	0.12.9	0.12.9	0.12.7	
RcppArmadillo	CRAN			0.7.700.0.0	0.6.700.6.0	
RcppEigen	CRAN		0.3.2.9.0			0.3.2.8.1
RCurl	CRAN		1.95-4.8	1.95-4.8	1.95-4.8	
reactome.db	Bioconductor		1.52.1	1.52.1		

		GitHub			
		TomKellyGenetics	0.1		
reactometree					
readr	CRAN	1.0.0	1.0.0		
readxl	CRAN	0.1.1			
registry	CRAN	0.3	0.3	0.3	0.3
reshape2	CRAN	1.4.1	1.4.2	1.4.2	1.4.1
rgefx	CRAN		0.15.3	0.15.3	
rgl	CRAN			0.97.0	0.95.1441
Rgraphviz	CRAN		2.18.0		
rjson	CRAN		0.2.15		
RJSONIO	CRAN		1.3-0		
rmarkdown	CRAN		1.3	1.3	1
Rmpi	CRAN		0.6-6		0.6-5
rngtools	CRAN	1.2.4	1.2.4	1.2.4	1.2.4
robustbase	CRAN	0.92-7	0.92-7	0.92-7	0.92-5
ROCR	CRAN	1.0-7	1.0-7	1.0-7	1.0-7
Rook	CRAN		1.1-1	1.1-1	
roxygen2	CRAN	6.0.1	5.0.1	6.0.1	5.0.1
rpart	base	4.1-10	4.1-10	4.1-10	4.1-10
rprojroot	CRAN	1.2	1.1	1.2	
Rsamtools	Bioconductor		1.26.1	1.26.1	
rsconnect	CRAN		0.7		
RSQLite	CRAN		1.1-2	1.1-2	1.0.0
rstudioapi	CRAN	0.6	0.6	0.6	0.6
rvest	CRAN	0.3.2			
S4Vectors	Bioconductor		0.12.1	0.12.0	0.10.3
safe	Bioconductor		3.14.0	3.10.0	
scales	CRAN	0.4.0	0.4.1	0.4.1	0.4.0
selectr	CRAN	0.3-1			
sem	CRAN		3.1-8		
shiny	CRAN	0.14		1.0.0	
slpt	GitHub TomKellyGenetics	0.1.0	0.1.0	0.1.0	0.1.0
sm	CRAN	2.2-5.4	2.2-5.4		
sna	CRAN		2.4		

snow	CRAN	0.4-1	0.4-2	0.4-2	0.3-13
sourcetools	CRAN	0.1.5		0.1.5	
SparseM	CRAN		1.74		1.7
spatial	base	7.3-11	7.3-11	7.3-11	7.3-11
splines	base	3.3.2	3.3.2	3.3.1	3.3.0
statnet.common	CRAN		3.3.0		
stats	base	3.3.2	3.3.2	3.3.1	3.3.0
stats4	base	3.3.2	3.3.2	3.3.1	3.3.0
stringi	CRAN	1.1.1	1.1.2	1.1.2	1.0-1
stringr	CRAN	1.1.0	1.1.0	1.2.0	1.0.0
Summarized Experiment	Bioconductor		1.4.0	1.4.0	
survival	base	2.39-4	2.40-1	2.40-1	2.39-4
tcltk	base	3.3.2	3.3.2	3.3.1	3.3.0
testthat	CRAN	1.0.2	1.0.2		1.0.2
tibble	CRAN	1.2	1.2	1.2	1.2
tidyverse	GitHub hadley	1.1.1			
timeline	CRAN		0.9		
tools	base	3.3.2	3.3.2	3.3.1	3.3.0
tpr	CRAN		0.3-1		
trimcluster	CRAN	0.1-2	0.1-2	0.1-2	
Unicode	CRAN	9.0.0-1	9.0.0-1	9.0.0-1	
utils	base	3.3.2	3.3.2	3.3.1	3.3.0
vioplot	CRAN		0.2		
vioplotx	GitHub TomKellyGenetics	0.0.0.9000	0.0.0.9000		
viridis	CRAN	0.3.4	0.3.4	0.3.4	
visNetwork	CRAN		1.0.3	1.0.3	
whisker	CRAN	0.3-2	0.3-2	0.3-2	0.3-2
withr	CRAN	1.0.2	1.0.2	1.0.2	1.0.2
XML	base	3.98-1.3	3.98-1.1	3.98-1.5	3.98-1.4
xml2	CRAN	1.1.1		1.1.1	1.0.0
xtable	CRAN	1.8-2	1.8-2	1.8-2	1.8-2

XVector	Bioconductor	0.14.0	0.14.0	
yaml	CRAN	2.1.14	2.1.14	2.1.13
zlibbioc	CRAN	1.20.0	1.20.0	
zoo	CRAN	1.7-13	1.7-14	1.7-13

Appendix C

Secondary Screen Data

A series of experimental genome-wide siRNA screens have been performed on synthetic lethal partners of *CDH1* (?). The strongest candidates from a primary screen were subject to a further secondary screen for validation by independent replication with 4 gene knockdowns with different targeting siRNA. As shown in Table C.1, there is significant ($p = 7.49 \times 10^{-3}$ by Fisher's exact test) association between SLIPT candidates and stronger validations of siRNA candidates. Since there were more SLIPT⁻ genes among those not validated and more SLIPT⁺ genes among those validated with several siRNAs, this supports the use of SLIPT as a synthetic lethal discovery procedure which may augment such screening experiments.

Table C.1: Comparing SLIPT genes against Secondary siRNA Screen in breast cancer

		Secondary Screen					Total
		0/4	1/4	2/4	3/4	4/4	
SLIPT⁺	Observed	70	46	31	8	2	157
	Expected	85	44	10	4	2	
SLIPT⁻	Observed	190	90	31	10	4	325
	Expected	175	91	42	12	4	
Total		280	136	52	18	6	482

Similar analysis with mtSLIPT, comparing SLIPT against *CDH1* somatic mutation with siRNA validation results was not significant ($p = 7.02 \times 10^{-1}$ by Fisher's exact test). However, as shown in Table C.2, the observed and expected values were in a direction consistent with that observed above for SLIPT against low *CDH1* expression. It is not unexpected that this result does not have comparable statistical support due to the lower sample size for mutation data.

This analysis was replicated on a (smaller) stomach cancer dataset but it was less conclusive ($p = 2.36 \times 10^{-1}$ by Fisher's exact test). As shown in Table C.3, fewer

Table C.2: Comparing mtSLIPT genes against Secondary siRNA Screen in breast cancer

		Secondary Screen					Total
		0/4	1/4	2/4	3/4	4/4	
mtSLIPT+	Observed	54	35	17	4	6	111
	Expected	60	31	14	4	1	
mtSLIPT-	Observed	206	101	45	14	5	371
	Expected	200	105	48	14	4	
Total		269	143	63	19	6	482

SLIPT candidates were validated than expected statistically. However, these results in stomach cancer may not be directly comparable to experiments in a breast cell line. Genes validated by 0 or 1 siRNA behave consistently with the results above.

Table C.3: Comparing SLIPT genes against Secondary siRNA Screen in stomach cancer

		Secondary Screen					Total
		0/4	1/4	2/4	3/4	4/4	
SLIPT+	Observed	67	47	13	4	1	132
	Expected	71	37	17	5	2	
SLIPT-	Observed	195	90	50	14	5	354
	Expected	190	100	46	13	4	
Total		262	137	63	19	6	486

Appendix D

Mutation Analysis in Breast Cancer

D.1 Synthetic Lethal Genes and Pathways

Table D.1: Candidate synthetic lethal genes against E-cadherin from mtSLIPT

Gene	Observed	Expected	χ^2 value	p-value	p-value (FDR)
<i>TFAP2B</i>	8	36.7	89.5	3.60×10^{-20}	8.37×10^{-17}
<i>ZNF423</i>	15	36.7	78.8	7.89×10^{-18}	1.22×10^{-14}
<i>CALCOCO1</i>	11	36.7	76.8	2.09×10^{-17}	2.59×10^{-14}
<i>RBM5</i>	13	36.7	75.7	3.65×10^{-17}	4.00×10^{-14}
<i>BTG2</i>	7	36.7	71.7	2.72×10^{-16}	1.81×10^{-13}
<i>RXRA</i>	6	36.7	70.5	5.00×10^{-16}	2.97×10^{-13}
<i>SLC27A1</i>	11	36.7	70.3	5.42×10^{-16}	2.97×10^{-13}
<i>MEF2D</i>	12	36.7	69.6	7.86×10^{-16}	3.95×10^{-13}
<i>NISCH</i>	12	36.7	69.6	7.86×10^{-16}	3.95×10^{-13}
<i>AVPR2</i>	9	36.7	69.2	9.36×10^{-16}	4.58×10^{-13}
<i>CRY2</i>	13	36.7	68.9	1.07×10^{-15}	4.98×10^{-13}
<i>RAPGEF3</i>	13	36.7	68.9	1.07×10^{-15}	4.98×10^{-13}
<i>NRIP2</i>	10	36.7	68.2	1.58×10^{-15}	7.18×10^{-13}
<i>DARC</i>	12	36.7	66.4	3.76×10^{-15}	1.54×10^{-12}
<i>SFRS5</i>	12	36.7	66.4	3.76×10^{-15}	1.54×10^{-12}
<i>NOSTRIN</i>	5	36.7	65.1	7.40×10^{-15}	2.70×10^{-12}
<i>KIF13B</i>	12	36.7	63.4	1.69×10^{-14}	5.16×10^{-12}
<i>TENC1</i>	10	36.7	62.5	2.67×10^{-14}	7.40×10^{-12}
<i>MFAP4</i>	12	36.7	60.5	7.17×10^{-14}	1.67×10^{-11}
<i>ELN</i>	13	36.7	59.7	1.07×10^{-13}	2.32×10^{-11}
<i>SGK223</i>	14	36.7	59	1.51×10^{-13}	3.05×10^{-11}
<i>KIF12</i>	11	36.7	58.8	1.74×10^{-13}	3.34×10^{-11}
<i>SELP</i>	11	36.7	58.8	1.74×10^{-13}	3.34×10^{-11}
<i>CIRBP</i>	9	36.7	58.7	1.83×10^{-13}	3.41×10^{-11}
<i>CTDSP1</i>	9	36.7	58.7	1.83×10^{-13}	3.41×10^{-11}

Strongest candidate SL partners for *CDH1* by mtSLIPT with observed and expected mutant samples with low expression of partner genes

Table D.2: Pathways for *CDH1* partners from mtSLIPT

Pathways Over-represented	Pathway Size	SL Genes	p-value (FDR)
Eukaryotic Translation Elongation	86	60	2.0×10^{-128}
Peptide chain elongation	83	59	2.0×10^{-128}
Eukaryotic Translation Termination	83	58	2.3×10^{-125}
Viral mRNA Translation	81	57	2.5×10^{-124}
Nonsense Mediated Decay independent of the Exon Junction Complex	88	59	8.6×10^{-124}
Nonsense-Mediated Decay	103	61	5.2×10^{-117}
Nonsense Mediated Decay enhanced by the Exon Junction Complex	103	61	5.2×10^{-117}
Formation of a pool of free 40S subunits	93	58	1.6×10^{-116}
L13a-mediated translational silencing of Ceruloplasmin expression	103	59	1.3×10^{-111}
3' -UTR-mediated translational regulation	103	59	1.3×10^{-111}
GTP hydrolysis and joining of the 60S ribosomal subunit	104	59	6.2×10^{-111}
SRP-dependent cotranslational protein targeting to membrane	104	58	2.9×10^{-108}
Eukaryotic Translation Initiation	111	59	3.0×10^{-106}
Cap-dependent Translation Initiation	111	59	3.0×10^{-106}
Influenza Viral RNA Transcription and Replication	108	57	5.1×10^{-103}
Influenza Infection	117	59	1.5×10^{-102}
Translation	141	64	3.7×10^{-101}
Influenza Life Cycle	112	57	1.4×10^{-100}
GPCR downstream signaling	472	116	1.0×10^{-80}
Hemostasis	422	105	1.4×10^{-78}

Gene set over-representation analysis (hypergeometric test) for Reactome pathways in mtSLIPT partners for *CDH1*

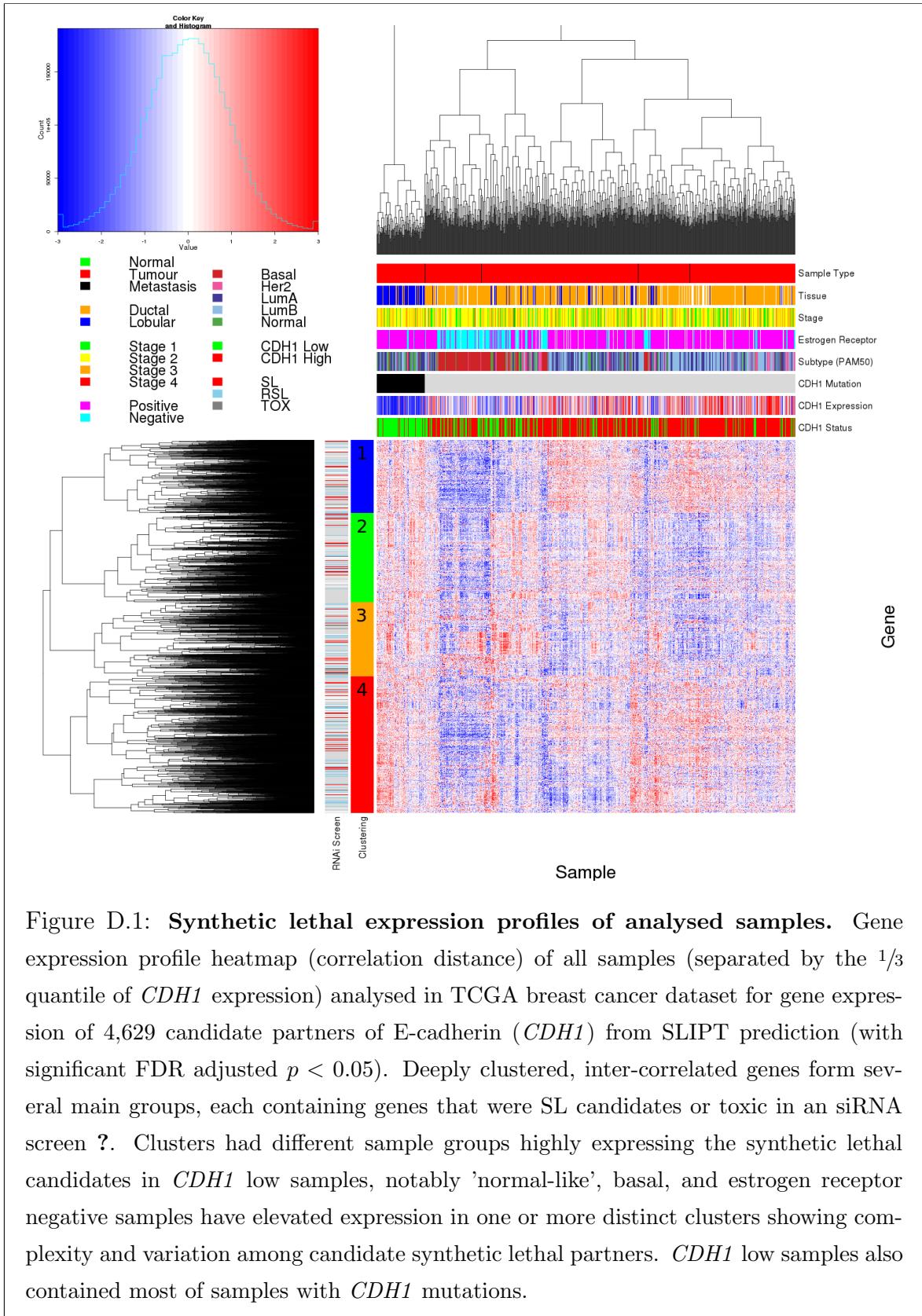


Table D.3: Pathway composition for clusters of *CDH1* partners from mtSLIPT

Pathways Over-represented in Cluster 1	Pathway Size	Cluster Genes	p-value (FDR)
Olfactory Signaling Pathway	57	8	7.1×10^{-9}
Assembly of the primary cilium	149	14	8×10^{-9}
Sphingolipid metabolism	62	8	9.6×10^{-9}
Signaling by ERBB4	133	12	5.1×10^{-8}
PI3K Cascade	65	7	4.9×10^{-7}
Circadian Clock	33	5	4.9×10^{-7}
Nuclear signaling by ERBB4	34	5	4.9×10^{-7}
Intraflagellar transport	35	5	4.9×10^{-7}
PI3K events in ERBB4 signaling	87	8	4.9×10^{-7}
PIP3 activates AKT signaling	87	8	4.9×10^{-7}
PI3K events in ERBB2 signaling	87	8	4.9×10^{-7}
PI-3K cascade:FGFR1	87	8	4.9×10^{-7}
PI-3K cascade:FGFR2	87	8	4.9×10^{-7}
PI-3K cascade:FGFR3	87	8	4.9×10^{-7}
PI-3K cascade:FGFR4	87	8	4.9×10^{-7}
Deadenylation of mRNA	22	4	5.6×10^{-7}
PI3K/AKT activation	90	8	5.6×10^{-7}
Cargo trafficking to the periciliary membrane	38	5	5.6×10^{-7}
Signaling by Hedgehog	108	9	5.6×10^{-7}
Downstream signal transduction	143	11	5.6×10^{-7}

Pathways Over-represented in Cluster 2	Pathway Size	Cluster Genes	p-value (FDR)
G _{αs} signalling events	83	19	5.1×10^{-25}
Extracellular matrix organization	238	30	1.4×10^{-18}
Hemostasis	422	46	2.7×10^{-16}
Aquaporin-mediated transport	32	9	2.7×10^{-16}
Transcriptional regulation of white adipocyte differentiation	56	11	1.7×10^{-15}
Degradation of the extracellular matrix	102	15	1.7×10^{-15}
Integration of energy metabolism	84	13	8.8×10^{-15}
GPCR downstream signaling	472	48	2.8×10^{-14}
G _{αs} signalling events	15	6	5×10^{-14}
Molecules associated with elastic fibres	33	8	5.4×10^{-14}
Phase 1 - Functionalization of compounds	67	11	5.6×10^{-14}
Platelet activation, signaling and aggregation	179	20	5.6×10^{-14}
Vasopressin regulates renal water homeostasis via Aquaporins	24	7	6.1×10^{-14}
Elastic fibre formation	37	8	3×10^{-13}
Calmodulin induced events	27	7	3.3×10^{-13}
CaM pathway	27	7	3.3×10^{-13}
cGMP effects	18	6	3.6×10^{-13}
G _{αs} signalling events	167	18	6.3×10^{-13}
Ca-dependent events	29	7	8.2×10^{-13}
Binding and Uptake of Ligands by Scavenger Receptors	40	8	8.2×10^{-13}

Pathways Over-represented in Cluster 3	Pathway Size	Cluster Genes	p-value (FDR)
Eukaryotic Translation Elongation	86	55	1.1×10^{-112}
Peptide chain elongation	83	54	1.3×10^{-112}
Viral mRNA Translation	81	53	1.6×10^{-111}
Eukaryotic Translation Termination	83	53	7.1×10^{-110}
Nonsense Mediated Decay independent of the Exon Junction Complex	88	54	1×10^{-108}
Formation of a pool of free 40S subunits	93	53	4.1×10^{-102}
Nonsense-Mediated Decay	103	54	3.9×10^{-98}
Nonsense-Mediated Decay enhanced by the Exon Junction Complex	103	54	3.9×10^{-98}
L13a-mediated translational silencing of Ceruloplasmin expression	103	53	1.2×10^{-95}
3' -UTR-mediated translational regulation	103	53	1.2×10^{-95}
SRP-dependent cotranslational protein targeting to membrane	104	53	4.3×10^{-95}
GTP hydrolysis and joining of the 60S ribosomal subunit	104	53	4.3×10^{-95}
Influenza Viral RNA Transcription and Replication	108	53	9.6×10^{-93}
Eukaryotic Translation Initiation	111	53	4.2×10^{-91}
Cap-dependent Translation Initiation	111	53	4.2×10^{-91}
Influenza Life Cycle	112	53	1.4×10^{-90}
Influenza Infection	117	53	6.2×10^{-88}
Translation	141	55	3×10^{-81}
Formation of the ternary complex, and subsequently, the 43S complex	47	23	2.3×10^{-48}
Translation initiation complex formation	54	23	9.1×10^{-45}

Pathways Over-represented in Cluster 4	Pathway Size	Cluster Genes	p-value (FDR)
ECM proteoglycans	66	10	2.9×10^{-11}
deactivation of the beta-catenin transactivating complex	38	7	5.1×10^{-10}
Arachidonic acid metabolism	41	7	1.1×10^{-9}
Gαq signalling events	149	14	4×10^{-9}
HS-GAG degradation	21	5	4.5×10^{-9}
Uptake and actions of bacterial toxins	22	5	6.1×10^{-9}
Gastrin-CREB signalling pathway via PKC and MAPK	170	15	6.1×10^{-9}
RNA Polymerase I, RNA Polymerase III, and Mitochondrial Transcription	64	8	6.1×10^{-9}
Non-integrin membrane-ECM interactions	53	7	1.5×10^{-8}
Syndecan interactions	25	5	1.5×10^{-8}
NOTCH1 Intracellular Domain Regulates Transcription	40	6	2.3×10^{-8}
Synthesis of Leukotrienes and Exoxins	15	4	3.2×10^{-8}
Signaling by NOTCH1	59	7	5.3×10^{-8}
Regulation of insulin secretion	44	6	6×10^{-8}
Metabolism of lipids and lipoproteins	471	37	8.2×10^{-8}
Signaling by NOTCH	80	8	1.2×10^{-7}
Platelet activation, signaling and aggregation	179	14	1.2×10^{-7}
Recruitment of mitotic centrosome proteins and complexes	64	7	1.2×10^{-7}
Centrosome maturation	64	7	1.2×10^{-7}
Biological oxidations	133	11	1.5×10^{-7}

D.2 Comparison to Primary Screen

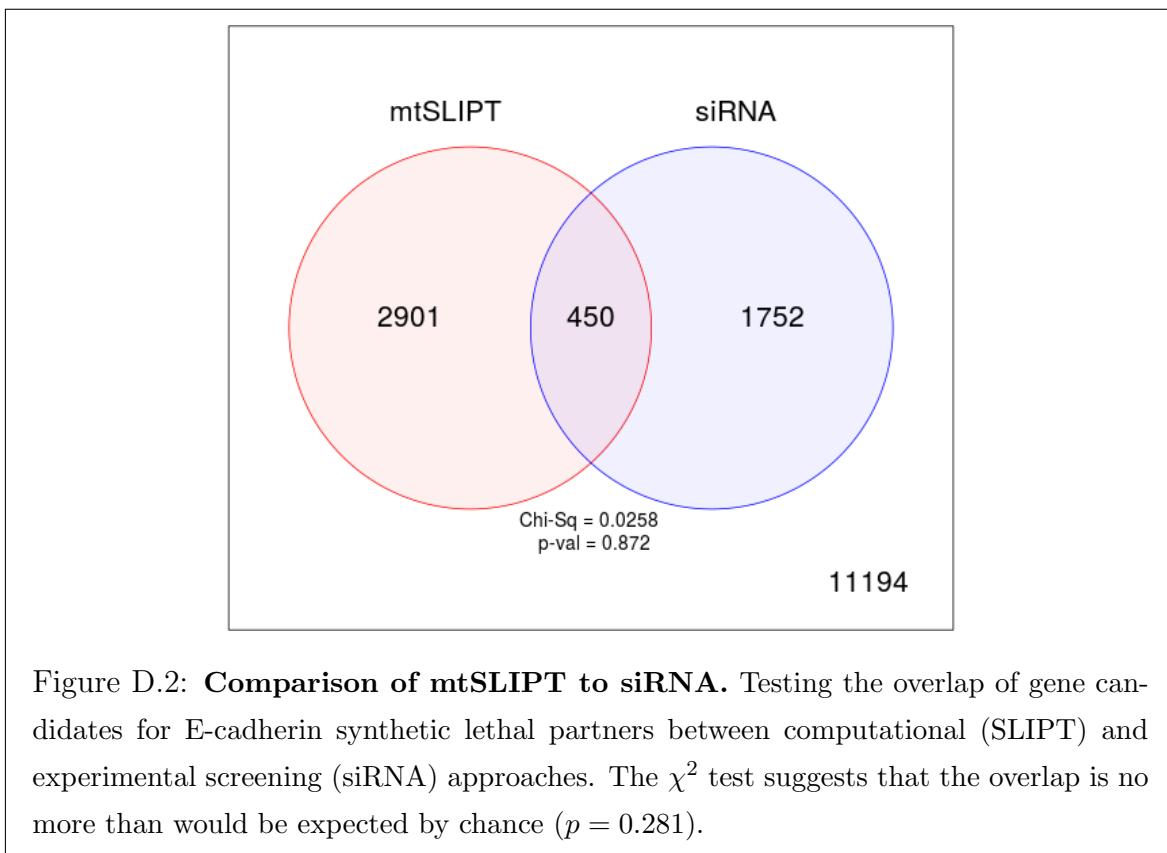


Table D.4: Pathway composition for *CDH1* partners from mtSLIPT and siRNA

Predicted only by SLIPT (2901 genes)	Pathway Size	Genes Identified	p-value (FDR)
Eukaryotic Translation Elongation	87	57	2.8×10^{-120}
Peptide chain elongation	84	56	3.1×10^{-120}
Eukaryotic Translation Termination	84	55	2.8×10^{-117}
Viral mRNA Translation	82	54	4.1×10^{-116}
Nonsense Mediated Decay independent of the Exon Junction Complex	89	55	3.7×10^{-113}
Formation of a pool of free 40S subunits	94	55	2.8×10^{-109}
Nonsense-Mediated Decay	104	57	8.4×10^{-108}
Nonsense Mediated Decay enhanced by the Exon Junction Complex	104	57	8.4×10^{-108}
L13a-mediated translational silencing of Ceruloplasmin expression	104	56	3.4×10^{-105}
3' -UTR-mediated translational regulation	104	56	3.4×10^{-105}
GTP hydrolysis and joining of the 60S ribosomal subunit	105	56	1.4×10^{-104}
Eukaryotic Translation Initiation	112	56	2.8×10^{-100}
Cap-dependent Translation Initiation	112	56	2.8×10^{-100}
SRP-dependent cotranslational protein targeting to membrane	105	54	2.2×10^{-99}
Influenza Viral RNA Transcription and Replication	109	54	5.3×10^{-97}
Influenza Life Cycle	113	54	9.6×10^{-95}
Influenza Infection	118	55	1.7×10^{-94}
Translation	142	60	3.5×10^{-94}
Infectious disease	349	77	5.9×10^{-62}
Extracellular matrix organization	241	54	3.0×10^{-52}

Detected only by siRNA screen (1752 genes)	Pathway Size	Genes Identified	p-value (FDR)
Class A/1 (Rhodopsin-like receptors)	282	69	1.9×10^{-59}
GPCR ligand binding	363	78	2.7×10^{-54}
Peptide ligand-binding receptors	175	41	1.5×10^{-42}
$G_{\alpha i}$ signalling events	184	41	1.1×10^{-40}
Gastrin-CREB signalling pathway via PKC and MAPK	180	37	1.5×10^{-35}
$G_{\alpha q}$ signalling events	159	34	3.7×10^{-35}
DAP12 interactions	159	27	1.1×10^{-24}
VEGFA-VEGFR2 Pathway	91	19	1.0×10^{-23}
Downstream signal transduction	146	24	1.9×10^{-22}
Signaling by VEGF	99	19	2.6×10^{-22}
DAP12 signaling	149	24	4.2×10^{-22}
Organelle biogenesis and maintenance	264	34	4.3×10^{-20}
Downstream signaling of activated FGFR1	134	21	4.3×10^{-20}
Downstream signaling of activated FGFR2	134	21	4.3×10^{-20}
Downstream signaling of activated FGFR3	134	21	4.3×10^{-20}
Downstream signaling of activated FGFR4	134	21	4.3×10^{-20}
Signaling by ERBB2	146	22	5.3×10^{-20}
Signaling by FGFR	146	22	5.3×10^{-20}
Signaling by FGFR1	146	22	5.3×10^{-20}
Signaling by FGFR2	146	22	5.3×10^{-20}

Intersection of SLIPT and siRNA screen (450 genes)	Pathway Size	Genes Identified	p-value (FDR)
HS-GAG degradation	21	4	4.9×10^{-6}
Retinoid metabolism and transport	39	5	4.9×10^{-6}
Platelet activation, signaling and aggregation	186	13	4.9×10^{-6}
Signaling by NOTCH4	11	3	4.9×10^{-6}
$G_{\alpha s}$ signalling events	100	8	5.0×10^{-6}
Defective EXT2 causes exostoses 2	12	3	5.0×10^{-6}
Defective EXT1 causes exostoses 1, TRPS2 and CHDS	12	3	5.0×10^{-6}
Class A/1 (Rhodopsin-like receptors)	289	18	2.2×10^{-5}
Signaling by PDGF	173	11	2.9×10^{-5}
Circadian Clock	34	4	2.9×10^{-5}
Signaling by ERBB4	139	9	4.3×10^{-5}
Role of LAT2/NTAL/LAB on calcium mobilization	99	7	4.4×10^{-5}
Peptide ligand-binding receptors	181	11	4.5×10^{-5}
Defective B4GALT7 causes EDS, progeroid type	19	3	4.5×10^{-5}
Defective B3GAT3 causes JDSSDH	19	3	4.5×10^{-5}
Signaling by NOTCH	80	6	4.5×10^{-5}
$G_{\alpha q}$ signalling events	164	10	5.1×10^{-5}
Response to elevated platelet cytosolic Ca^{2+}	84	6	7.1×10^{-5}
Signaling by ERBB2	148	9	7.1×10^{-5}
Signaling by SCF-KIT	129	8	8.3×10^{-5}

D.3 Resampling Analysis

Table D.5: Pathways for *CDH1* partners from mtSLIPT

Reactome Pathway	Over-representation	Permutation
Eukaryotic Translation Elongation	3.2×10^{-128}	$< 7.035 \times 10^{-4}$
Peptide chain elongation	3.2×10^{-128}	$< 7.035 \times 10^{-4}$
Eukaryotic Translation Termination	3.7×10^{-125}	$< 7.035 \times 10^{-4}$
Viral mRNA Translation	4.1×10^{-124}	$< 7.035 \times 10^{-4}$
Nonsense Mediated Decay independent of the Exon Junction Complex	1.4×10^{-123}	$< 7.035 \times 10^{-4}$
Nonsense-Mediated Decay	8.4×10^{-117}	$< 7.035 \times 10^{-4}$
Nonsense Mediated Decay enhanced by the Exon Junction Complex	8.4×10^{-117}	$< 7.035 \times 10^{-4}$
Formation of a pool of free 40S subunits	2.6×10^{-116}	$< 7.035 \times 10^{-4}$
L13a-mediated translational silencing of Ceruloplasmin expression	2.0×10^{-111}	$< 7.035 \times 10^{-4}$
3' -UTR-mediated translational regulation	2.0×10^{-111}	$< 7.035 \times 10^{-4}$
GTP hydrolysis and joining of the 60S ribosomal subunit	9.9×10^{-111}	$< 7.035 \times 10^{-4}$
SRP-dependent cotranslational protein targeting to membrane	4.7×10^{-108}	$< 7.035 \times 10^{-4}$
Eukaryotic Translation Initiation	4.8×10^{-106}	$< 7.035 \times 10^{-4}$
Cap-dependent Translation Initiation	4.8×10^{-106}	$< 7.035 \times 10^{-4}$
Influenza Viral RNA Transcription and Replication	8.1×10^{-103}	$< 7.035 \times 10^{-4}$
Influenza Infection	2.4×10^{-102}	$< 7.035 \times 10^{-4}$
Translation	6.0×10^{-101}	$< 7.035 \times 10^{-4}$
Influenza Life Cycle	2.2×10^{-100}	$< 7.035 \times 10^{-4}$
Disease	2.1×10^{-90}	0.013347
GPCR downstream signaling	1.6×10^{-80}	0.095478
Hemostasis	2.1×10^{-78}	0.2671
Signaling by GPCR	1.2×10^{-73}	0.44939
<i>Extracellular matrix organization</i>	2.2×10^{-67}	0.054008
Metabolism of proteins	1.4×10^{-66}	0.9607
Signal Transduction	2.1×10^{-66}	0.48184
Developmental Biology	2.5×10^{-66}	0.54075
Innate Immune System	5.3×10^{-66}	0.9589
Infectious disease	9.6×10^{-66}	0.21075
Signalling by NGF	1.1×10^{-62}	0.43356
Immune System	2.8×10^{-62}	0.23052

Over-representation (hypergeometric test) and Permutation p-values adjusted for multiple tests across pathways (FDR). Significant pathways are marked in bold (FDR < 0.05) and italics (FDR < 0.1).

Table D.6: Pathways for *CDH1* partners from mtSLIPT and siRNA primary screen

Reactome Pathway	Over-representation	Permutation
Visual phototransduction	1.2×10^{-9}	0.86279
G_{αs} signalling events	2.9×10^{-7}	0.023066
Retinoid metabolism and transport	2.9×10^{-7}	0.299
Acylic chain remodelling of PS	1.1×10^{-5}	0.42584
Transcriptional regulation of white adipocyte differentiation	1.1×10^{-5}	0.53928
Chemokine receptors bind chemokines	1.1×10^{-5}	0.95259
<i>Signaling by NOTCH4</i>	1.2×10^{-5}	0.079229
Defective EXT2 causes exostoses 2	1.2×10^{-5}	0.22292
Defective EXT1 causes exostoses 1, TRPS2 and CHDS	1.2×10^{-5}	0.22292
Platelet activation, signaling and aggregation	1.2×10^{-5}	0.48853
Serotonin receptors	1.4×10^{-5}	0.34596
Nicotinamide salvaging	1.4×10^{-5}	0.70881
Phase 1 - Functionalization of compounds	2×10^{-5}	0.31142
Amine ligand-binding receptors	2.5×10^{-5}	0.34934
Acylic chain remodelling of PE	3.8×10^{-5}	0.42615
Signaling by GPCR	3.8×10^{-5}	0.93888
Molecules associated with elastic fibres	3.9×10^{-5}	0.017982
DAP12 interactions	3.9×10^{-5}	0.71983
Beta defensins	3.9×10^{-5}	0.91458
Cytochrome P ₄₅₀ - arranged by substrate type	4.7×10^{-5}	0.83493
GPCR ligand binding	5.7×10^{-5}	0.95258
Acylic chain remodelling of PC	6.1×10^{-5}	0.42584
Response to elevated platelet cytosolic Ca ²⁺	6.4×10^{-5}	0.54046
Arachidonic acid metabolism	6.7×10^{-5}	0.026696
Defective B4GALT7 causes EDS, progeroid type	7.3×10^{-5}	0.24921
Defective B3GAT3 causes JDSSDHD	7.3×10^{-5}	0.24921
Hydrolysis of LPC	7.3×10^{-5}	0.80663
Elastic fibre formation	7.4×10^{-5}	0.0058768
HS-GAG degradation	9.4×10^{-5}	0.0083179
<i>Bile acid and bile salt metabolism</i>	9.4×10^{-5}	0.079905
Netrin-1 signaling	0.00011	0.92216
Integration of energy metabolism	0.00011	0.011152
Dectin-2 family	0.00012	0.10385
Platelet sensitization by LDL	0.00012	0.34596
DAP12 signaling	0.00012	0.62787
Defensins	0.00012	0.77542
GPCR downstream signaling	0.00012	0.79454
<i>Diseases associated with glycosaminoglycan metabolism</i>	0.00013	0.065927
<i>Diseases of glycosylation</i>	0.00013	0.065927
Signaling by Retinoic Acid	0.00013	0.22292
Signaling by Leptin	0.00013	0.34596
Signaling by SCF-KIT	0.00013	0.70881
Opioid Signalling	0.00013	0.96053
Signaling by NOTCH	0.00015	0.26884
Platelet homeostasis	0.00015	0.4878
Signaling by NOTCH1	0.00016	0.13043
Class B/2 (Secretin family receptors)	0.00016	0.13994
<i>Diseases of Immune System</i>	0.0002	0.0795
<i>Diseases associated with the TLR signaling cascade</i>	0.0002	0.0795
A tetrasaccharide linker sequence is required for GAG synthesis	0.0002	0.42615

Over-representation (hypergeometric test) and Permutation p-values adjusted for multiple tests across pathways (FDR). Significant pathways are marked in bold (FDR < 0.05) and italics (FDR < 0.1).

D.4 Metagene Analysis

Table D.7: Candidate synthetic lethal metagenes against *CDH1* from mtSLIPT

Pathway	ID	Observed	Expected	χ^2 value	p-value	p-value (FDR)
Linoleic acid (LA) metabolism	2046105	79	36.70	87.03	1.2637×10^{-19}	2.0839×10^{-160}
ATF6-alpha activates chaperone genes	381183	78	36.70	80.25	3.7449×10^{-18}	3.0877×10^{-150}
Neurotoxicity of clostridium toxins	168799	8	36.70	79.41	5.7092×10^{-18}	3.1382×10^{-150}
Aquaporin-mediated transport	445717	8	36.70	76.28	2.7327×10^{-17}	9.0124×10^{-150}
Toxicity of botulinum toxin type G (BoNTG)	5250989	8	36.70	76.278	2.7327×10^{-17}	9.0124×10^{-150}
Purine metabolism	73847	75	36.70	75.86	3.3623×10^{-17}	9.2407×10^{-150}
Chk1Chk2(Cds1) mediated inactivation of Cyclin B:Cdk1 complex	75035	74	36.70	71.68	2.7211×10^{-16}	6.41×10^{-140}
Scavenging by Class F Receptors	3000484	75	36.70	69.56	7.8573×10^{-16}	1.4396×10^{-130}
Cytosolic tRNA aminoacylation	379716	75	36.70	69.56	7.8573×10^{-16}	1.4396×10^{-130}
G1S Transition	69206	74	36.70	69.21	9.3593×10^{-16}	1.5433×10^{-130}
ABC-family proteins mediated transport	382556	10	36.70	68.16	1.5826×10^{-15}	1.8641×10^{-130}
MG1 Transition	68874	74	36.70	68.16	1.5826×10^{-15}	1.8641×10^{-130}
DNA Replication Pre-Initiation	69002	74	36.70	68.16	1.5826×10^{-15}	1.8641×10^{-130}
Cell Cycle Checkpoints	69620	74	36.70	68.16	1.5826×10^{-15}	1.8641×10^{-130}
Basigin interactions	210991	74	36.70	67.23	2.5162×10^{-15}	2.7661×10^{-130}
Mitotic G1-G1S phases	453279	72	36.70	64.98	7.7471×10^{-15}	7.9843×10^{-130}
Metabolism of folate and pterines	196757	73	36.70	63.42	1.6932×10^{-14}	1.6424×10^{-120}
Tetrahydrobiopterin (BH4) synthesis, recycling, salvage and regulation	1474151	73	36.70	62.68	2.4547×10^{-14}	2.0427×10^{-120}
DNA Replication	69306	72	36.70	62.51	2.6652×10^{-14}	2.0427×10^{-120}
Separation of Sister Chromatids	2467813	71	36.70	62.47	2.7252×10^{-14}	2.0427×10^{-120}
M Phase	68886	71	36.70	62.47	2.7252×10^{-14}	2.0427×10^{-120}
Cell Cycle, Mitotic	69278	71	36.70	62.47	2.7252×10^{-14}	2.0427×10^{-120}
G0 and Early G1	1538133	70	36.70	61.62	4.1658×10^{-14}	2.8623×10^{-120}
Regulation of PLK1 Activity at G2M Transition	2565942	70	36.70	61.62	4.1658×10^{-14}	2.8623×10^{-120}
alpha-linolenic (omega3) and linoleic (omega6) acid metabolism	2046104	70	36.70	60.07	9.0139×10^{-14}	5.1255×10^{-120}

Strongest candidate SL partners for *CDH1* by mtSLIPT with observed and expected mutant samples with low expression of partner metagenes

D.5 Mutation Variation

D.5.1 Mutation Frequency

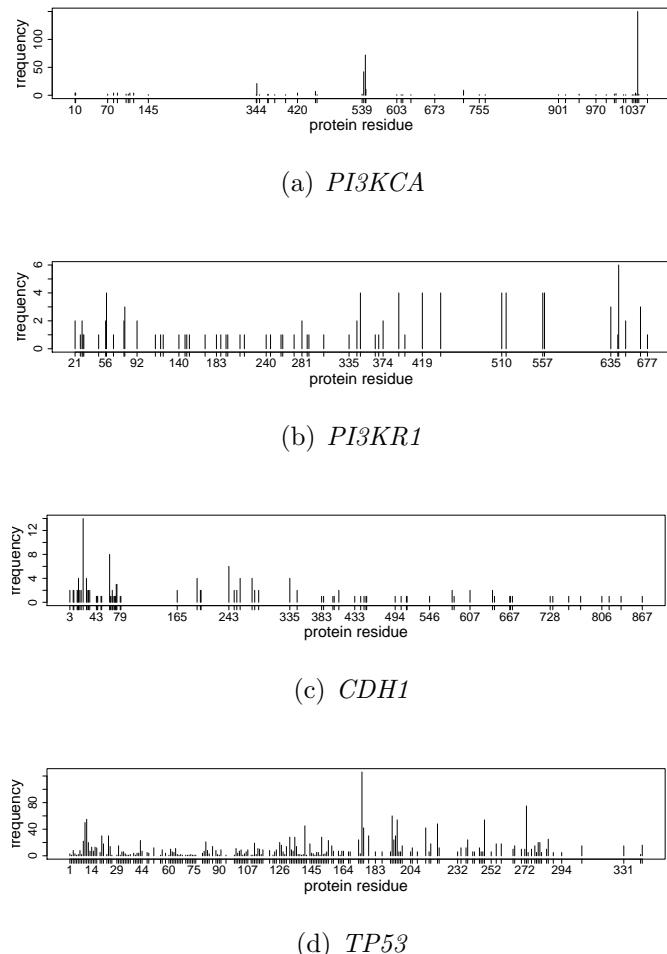


Figure D.3: **Somatic mutation locus.** Mutation frequency at each locus in TCGA breast cancer. *PIK3CA* shows clear recurrent E545K and H1047R oncogene mutations consistent with it being an oncogene. *PIK3R1* and *CDH1* are tumour suppressors with inactivating mutations distributed throughout the gene, whereas *TP53* exhibits both of these properties and a very high mutation frequency compared to other genes.

D.5.2 PI3K Mutation Expression

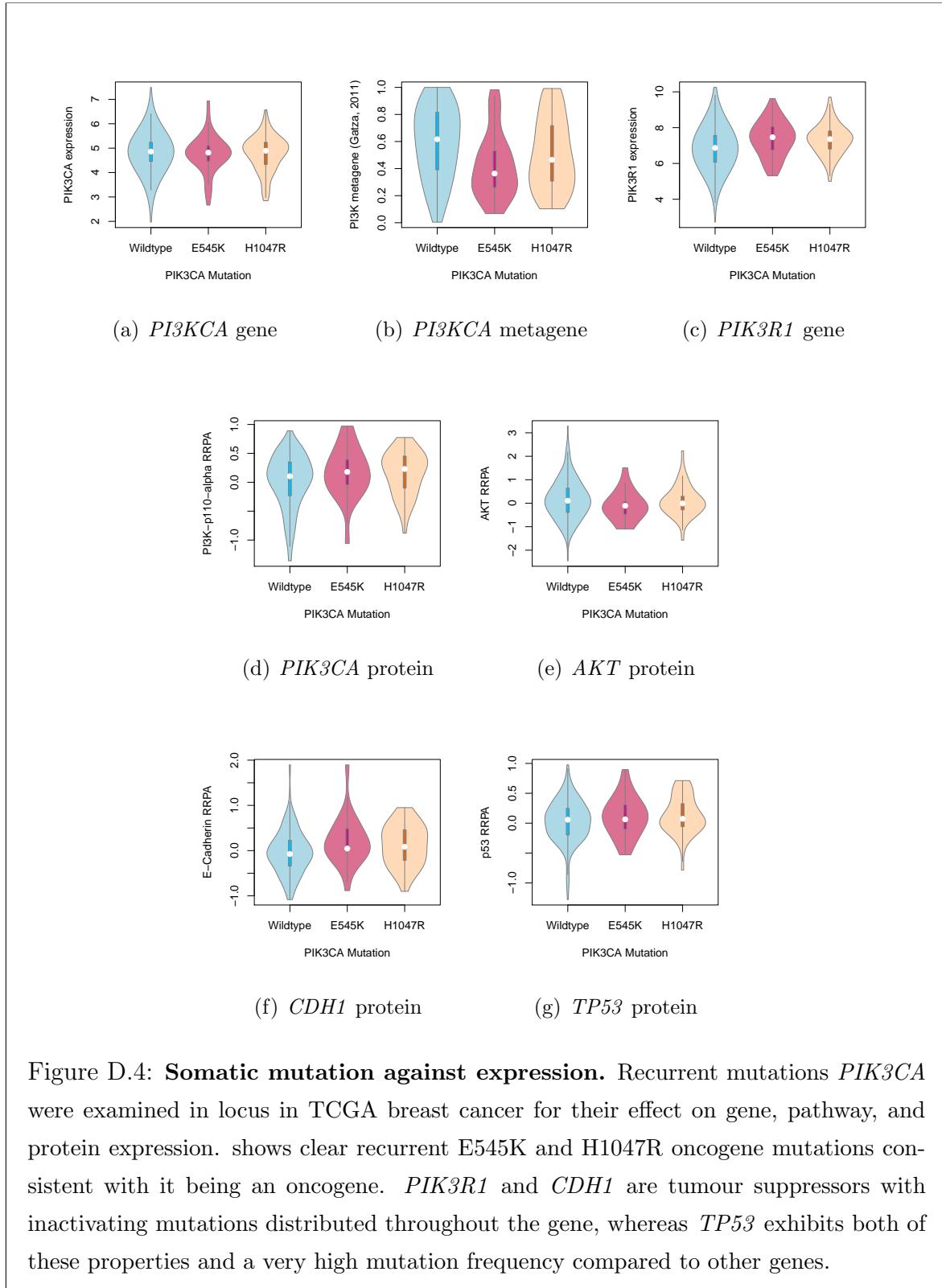


Figure D.4: Somatic mutation against expression. Recurrent mutations *PIK3CA* were examined in locus in TCGA breast cancer for their effect on gene, pathway, and protein expression. shows clear recurrent E545K and H1047R oncogene mutations consistent with it being an oncogene. *PIK3R1* and *CDH1* are tumour suppressors with inactivating mutations distributed throughout the gene, whereas *TP53* exhibits both of these properties and a very high mutation frequency compared to other genes.

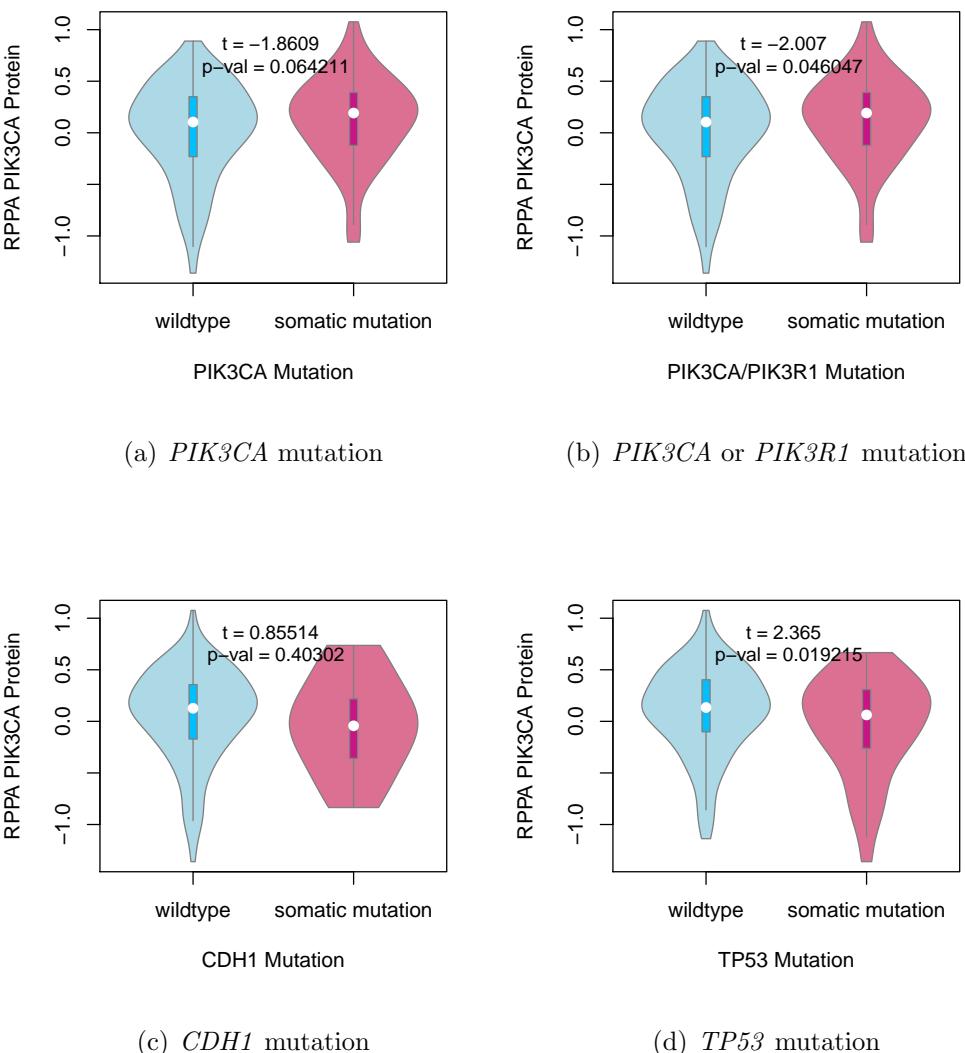


Figure D.5: Somatic mutation against PI3K protein. Recurrent mutations *PIK3CA* were examined in locus in TCGA breast cancer for their effect on gene, pathway, and protein expression (p110 α protein). shows clear recurrent E545K and H1047R oncogene mutations consistent with it being an oncogene. *PIK3R1* and *CDH1* are tumour suppressors with inactivating mutations distributed throughout the gene, whereas *TP53* exhibits both of these properties and a very high mutation frequency compared to other genes.

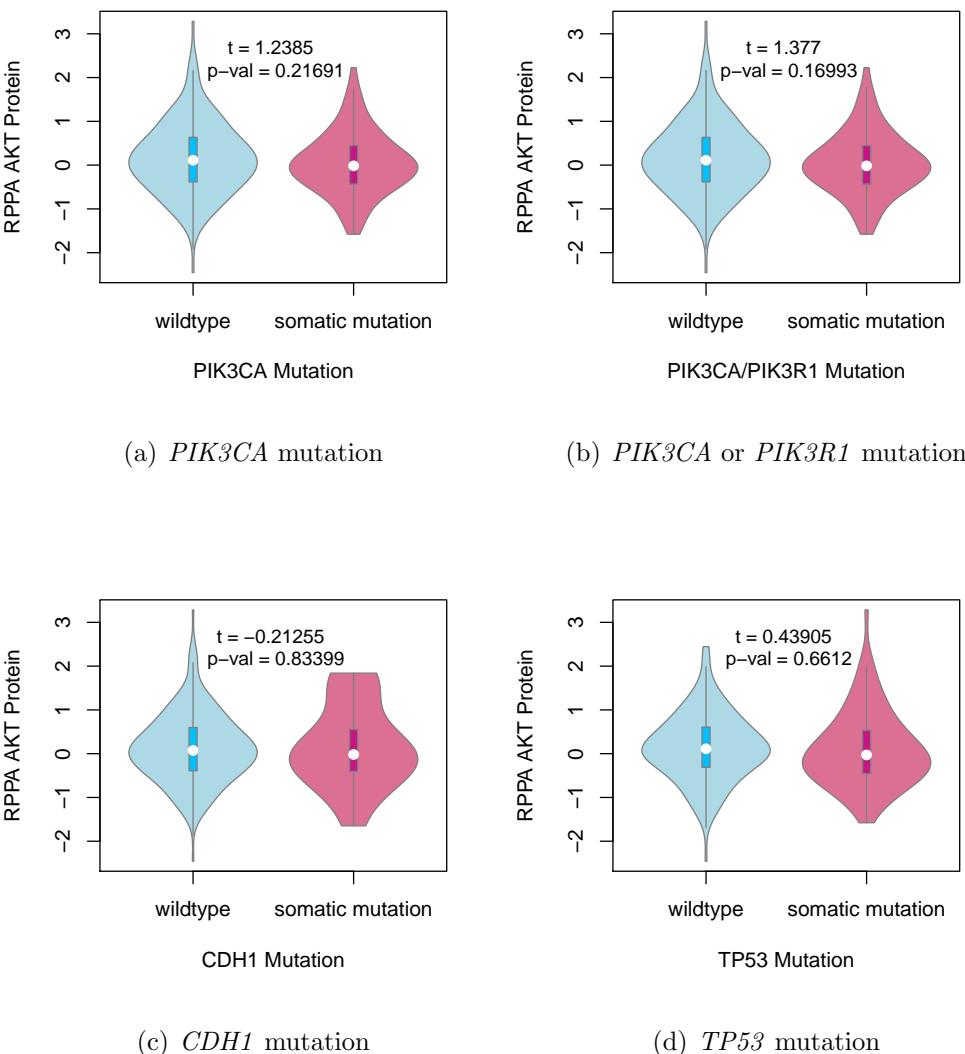


Figure D.6: Somatic mutation against AKT protein. Recurrent mutations *PIK3CA* were examined in locus in TCGA breast cancer for their effect on gene, pathway, and protein expression. shows clear recurrent E545K and H1047R oncogene mutations consistent with it being an oncogene. *PIK3R1* and *CDH1* are tumour suppressors with inactivating mutations distributed throughout the gene, whereas *TP53* exhibits both of these properties and a very high mutation frequency compared to other genes.

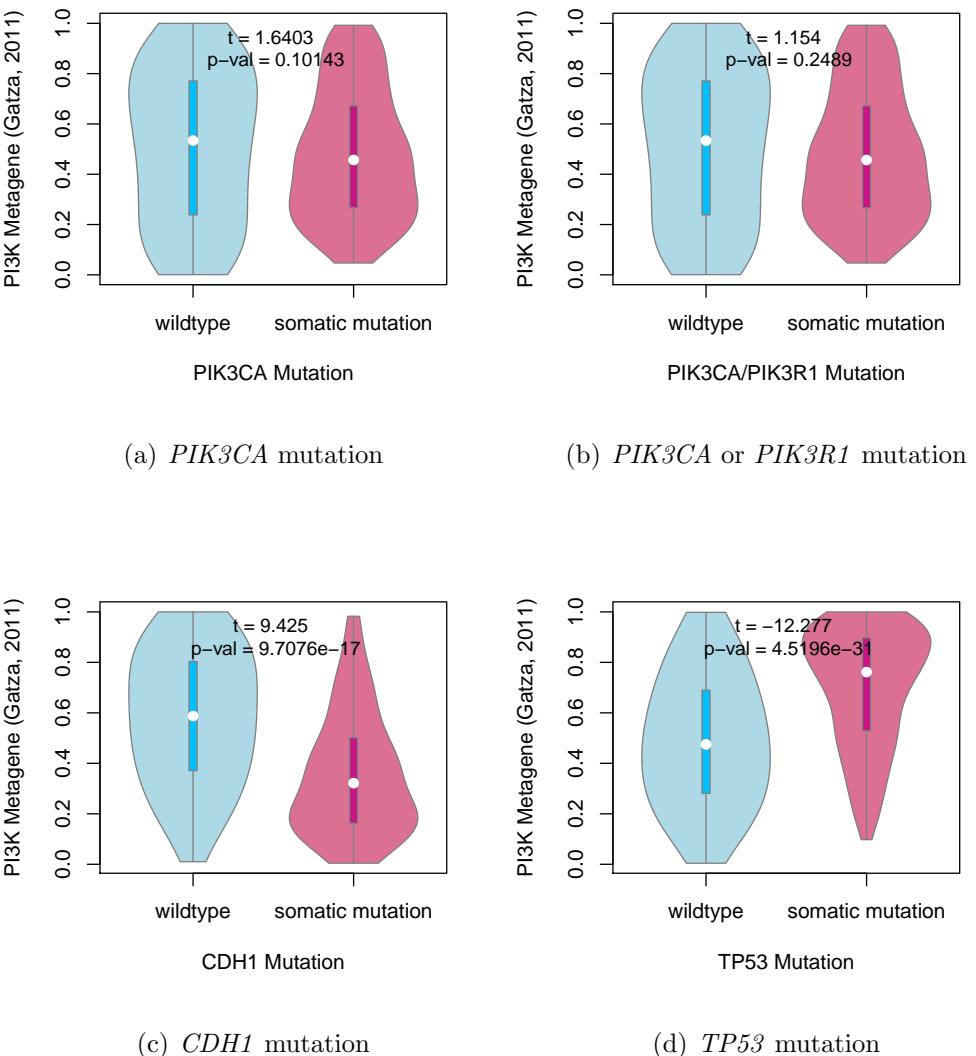


Figure D.7: Somatic mutation against PI3K metagene. Recurrent mutations *PIK3CA* were examined in locus in TCGA breast cancer for their effect on gene, pathway, and protein expression. shows clear recurrent E545K and H1047R oncogene mutations consistent with it being an oncogene. *PIK3R1* and *CDH1* are tumour suppressors with inactivating mutations distributed throughout the gene, whereas *TP53* exhibits both of these properties and a very high mutation frequency compared to other genes.

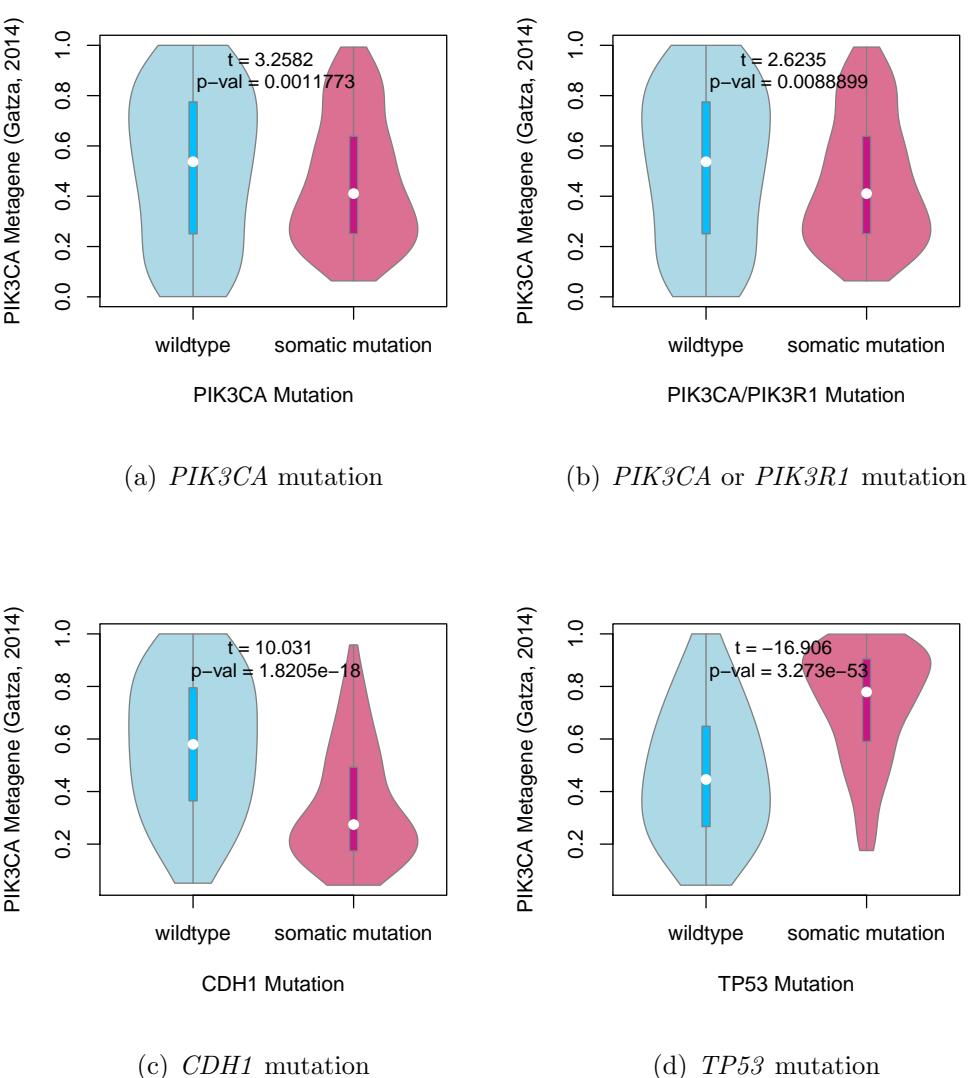


Figure D.8: Somatic mutation against PIK3CA metagene. Recurrent mutations *PIK3CA* were examined in locus in TCGA breast cancer for their effect on gene, pathway, and protein expression. shows clear recurrent E545K and H1047R oncogene mutations consistent with it being an oncogene. *PIK3R1* and *CDH1* are tumour suppressors with inactivating mutations distributed throughout the gene, whereas *TP53* exhibits both of these properties and a very high mutation frequency compared to other genes.

D.6 Compare SLIPT genes

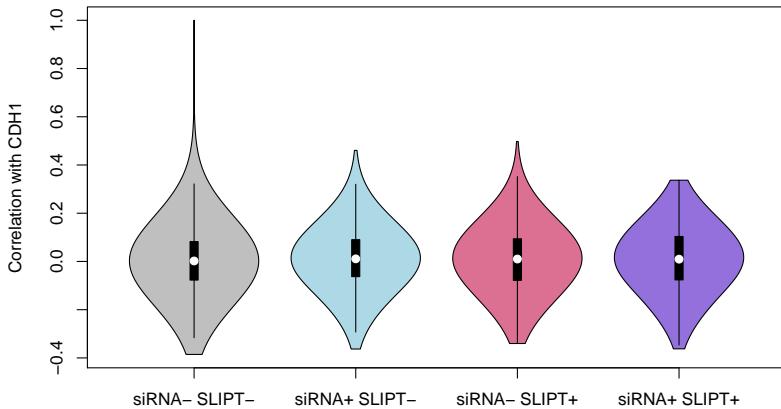


Figure D.9: **Compare mtSLIPT and siRNA genes with correlation.** Caption, Caption, Caption...

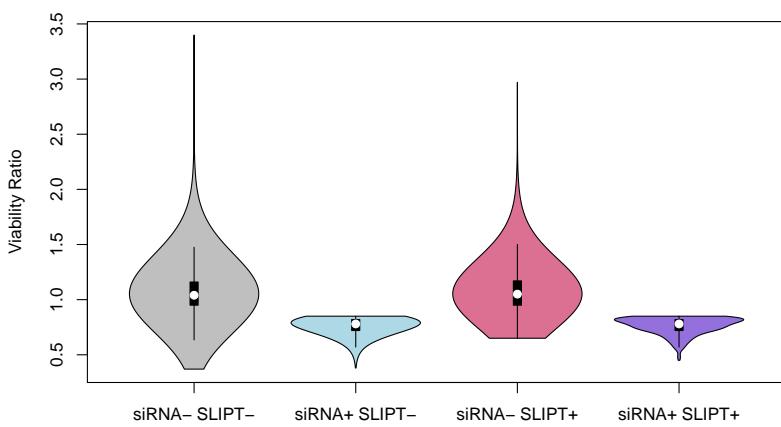


Figure D.10: **Compare mtSLIPT and siRNA genes with siRNA viability.** Caption, Caption, Caption...

Appendix E

Expression Analysis in Stomach Cancer

Appendix F

Mutation Analysis in Stomach Cancer

Table F.1: Candidate synthetic lethal genes against E-cadherin from mtSLIPT in stomach cancer

Gene	Observed	Expected	χ^2 value	p-value	p-value (FDR)
<i>OLFML1</i>	5	10.1	29.2	4.53×10^{-7}	0.0031
<i>NRIP2</i>	6	10.1	25.4	3.11×10^{-6}	0.00706
<i>VIM</i>	3	10.1	24.7	4.29×10^{-6}	0.00706
<i>TCF4</i>	5	10.1	24.7	4.33×10^{-6}	0.00706
<i>ZEB2</i>	5	10.1	24.7	4.33×10^{-6}	0.00706
<i>BCL2</i>	2	10.1	22	1.66×10^{-5}	0.0155
<i>SMARCA2</i>	2	10.1	22	1.66×10^{-5}	0.0155
<i>CCND2</i>	3	10.1	21.1	2.61×10^{-5}	0.0155
<i>MMP19</i>	3	10.1	21.1	2.61×10^{-5}	0.0155
<i>NEURL1B</i>	3	10.1	21.1	2.61×10^{-5}	0.0155
<i>IGFBP6</i>	6	10.1	21.1	2.65×10^{-5}	0.0155
<i>OGN</i>	6	10.1	21.1	2.65×10^{-5}	0.0155
<i>THY1</i>	6	10.2	21	2.7×10^{-5}	0.0155
<i>DZIP1</i>	4	10.1	20.6	3.29×10^{-5}	0.0155
<i>LOC650368</i>	4	10.1	20.6	3.29×10^{-5}	0.0155
<i>PCOLCE</i>	4	10.1	20.6	3.29×10^{-5}	0.0155
<i>PTGFR</i>	4	10.1	20.6	3.29×10^{-5}	0.0155
<i>RUNX1T1</i>	4	10.1	20.6	3.29×10^{-5}	0.0155
<i>CLEC2B</i>	5	10.1	20.6	3.3×10^{-5}	0.0155
<i>MSC</i>	5	10.1	20.6	3.3×10^{-5}	0.0155
<i>NISCH</i>	5	10.1	20.6	3.3×10^{-5}	0.0155
<i>TSPAN11</i>	5	10.1	20.6	3.3×10^{-5}	0.0155
<i>KCTD12</i>	2	10.1	19.1	7.19×10^{-5}	0.0246
<i>LRRC55</i>	2	10.1	19.1	7.19×10^{-5}	0.0246
<i>PCBP3</i>	2	10.1	19.1	7.19×10^{-5}	0.0246

Strongest candidate SL partners for *CDH1* by mtSLIPT with observed and expected mutant samples with low expression of partner genes

Table F.2: Pathways for *CDH1* partners from mtSLIPT in stomach cancer

Pathways Over-represented	Pathway Size	SL Genes	p-value (FDR)
Extracellular matrix organization	241	20	9.6×10^{-9}
Elastic fibre formation	38	6	3.7×10^{-8}
Diseases associated with glycosaminoglycan metabolism	26	5	3.7×10^{-8}
Diseases of glycosylation	26	5	3.7×10^{-8}
Nitric oxide stimulates guanylate cyclase	24	4	3.1×10^{-6}
Molecules associated with elastic fibres	34	4	3.7×10^{-5}
Platelet homeostasis	54	5	3.7×10^{-5}
Initial triggering of complement	17	3	3.7×10^{-5}
Regulation of IGF transport and uptake by IGFBPs	17	3	3.7×10^{-5}
Collagen degradation	58	5	5.6×10^{-5}
Defective B4GALT7 causes EDS, progeroid type	19	3	5.6×10^{-5}
Defective B3GAT3 causes JDSSDHD	19	3	5.6×10^{-5}
Degradation of the extracellular matrix	104	7	8.0×10^{-5}
ECM proteoglycans	66	5	0.00017
A tetrasaccharide linker sequence is required for GAG synthesis	25	3	0.00025
RHO GTPases Activate WASPs and WAVEs	29	3	0.00059
Non-integrin membrane-ECM interactions	53	4	0.00065
Creation of C4 and C2 activators	11	2	0.00079
Dermatan sulfate biosynthesis	11	2	0.00079
Integrin cell surface interactions	82	5	0.00098

Gene set over-representation analysis (hypergeometric test) for Reactome pathways in mtSLIPT partners for *CDH1*

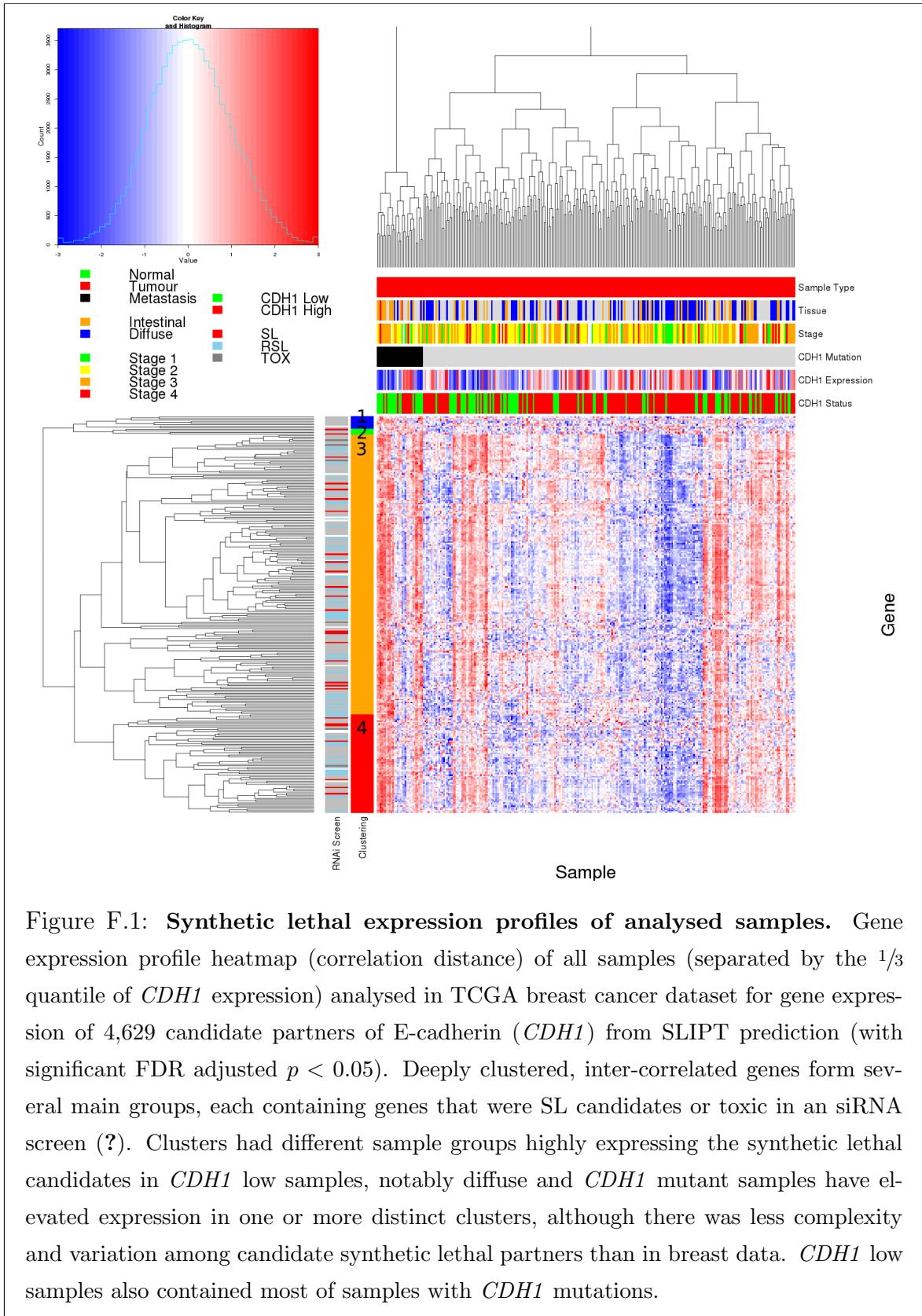


Table F.3: Pathway composition for clusters of *CDH1* partners in stomach mtSLIPT

Pathways Over-represented in Cluster 1	Pathway Size	Cluster Genes	p-value (FDR)
CD28 dependent PI3K/Akt signaling	15	1	1
Hormone-sensitive lipase (HSL)-mediated triacylglycerol hydrolysis	19	1	1
CD28 co-stimulation	26	1	1
Lipid digestion, mobilization, and transport	48	1	1
Costimulation by the CD28 family	51	1	1
Dectin-1 mediated noncanonical NF- κ B signaling	58	1	1
CLEC7A (Dectin-1) signaling	99	1	1
C-type lectin receptors (CLRs)	123	1	1
Adaptive Immune System	418	1	1
Metabolism of lipids and lipoproteins	494	1	1
Interleukin-6 signaling	10	0	1
Apoptosis	150	0	1
Hemostasis	445	0	1
Intrinsic Pathway for Apoptosis	36	0	1
Cleavage of Growing Transcript in the Termination Region	33	0	1
PKB-mediated events	28	0	1
PI3K Cascade	68	0	1
RAF/MAP kinase cascade	10	0	1
Global Genomic NER (GG-NER)	35	0	1
Repair synthesis for gap-filling by DNA polymerase in TC-NER	15	0	1

Pathways Over-represented in Cluster 2	Pathway Size	Cluster Genes	p-value (FDR)
Kinesins	22	1	1
O-linked glycosylation of mucins	49	1	1
O-linked glycosylation	59	1	1
MHC class II antigen presentation	85	1	1
Factors involved in megakaryocyte development and platelet production	120	1	1
Post-translational protein modification	303	1	1
Adaptive Immune System	418	1	1
Hemostasis	445	1	1
Interleukin-6 signaling	10	0	1
Apoptosis	150	0	1
Intrinsic Pathway for Apoptosis	36	0	1
Cleavage of Growing Transcript in the Termination Region	33	0	1
PKB-mediated events	28	0	1
PI3K Cascade	68	0	1
RAF/MAP kinase cascade	10	0	1
Global Genomic NER (GG-NER)	35	0	1
Repair synthesis for gap-filling by DNA polymerase in TC-NER	15	0	1
Gap-filling DNA repair synthesis and ligation in TC-NER	17	0	1
Formation of transcription-coupled NER (TC-NER) repair complex	29	0	1
Dual incision reaction in TC-NER	29	0	1

Pathways Over-represented in Cluster 3	Pathway Size	Cluster Genes	p-value (FDR)
Extracellular matrix organization	241	20	9.6×10^{-9}
Elastic fibre formation	38	6	3.7×10^{-8}
Diseases associated with glycosaminoglycan metabolism	26	5	3.7×10^{-8}
Diseases of glycosylation	26	5	3.7×10^{-8}
Molecules associated with elastic fibres	34	4	4.8×10^{-5}
Initial triggering of complement	17	3	4.8×10^{-5}
Regulation of IGF transport and uptake by IGFBPs	17	3	4.8×10^{-5}
Collagen degradation	58	5	6.7×10^{-5}
Defective B4GALT7 causes EDS, progeroid type	19	3	6.7×10^{-5}
Defective B3GAT3 causes JDSSDH	19	3	6.7×10^{-5}
Degradation of the extracellular matrix	104	7	9.5×10^{-5}
ECM proteoglycans	66	5	0.0002
A tetrasaccharide linker sequence is required for GAG synthesis	25	5	0.00029
Non-integrin membrane-ECM interactions	53	4	0.00079
Creation of C4 and C2 activators	11	2	0.00093
Dermatan sulfate biosynthesis	11	2	0.00093
Integrin cell surface interactions	82	5	0.0012
Keratan sulfate degradation	12	2	0.0012
Complement cascade	34	3	0.0013
CS/DS degradation	13	2	0.0015

Pathways Over-represented in Cluster 4	Pathway Size	Cluster Genes	p-value (FDR)
cGMP effects	18	2	0.11
Nitric oxide stimulates guanylate cyclase	24	2	0.19
Neurotoxicity of clostridium toxins	10	1	1
Platelet homeostasis	54	2	1
Eicosanoid ligand-binding receptors	14	1	1
Prolactin receptor signaling	15	1	1
Acyl chain remodelling of PI	15	1	1
Signaling by FGFR1 fusion mutants	15	1	1
PKA activation	16	1	1
PKA-mediated phosphorylation of CREB	17	1	1
Synthesis of glycosylphosphatidylinositol (GPI)	17	1	1
PKA activation in glucagon signalling	17	1	1
Butyrate Response Factor 1 (BRF1) destabilizes mRNA	17	1	1
Other semaphorin interactions	19	1	1
Acyl chain remodelling of PE	21	1	1
Signaling by Leptin	21	1	1
DARPP-32 events	22	1	1
Glucagon-like Peptide-1 (GLP1) regulates insulin secretion	22	1	1
Uptake and actions of bacterial toxins	22	1	1
Acyl chain remodelling of PC	23	1	1

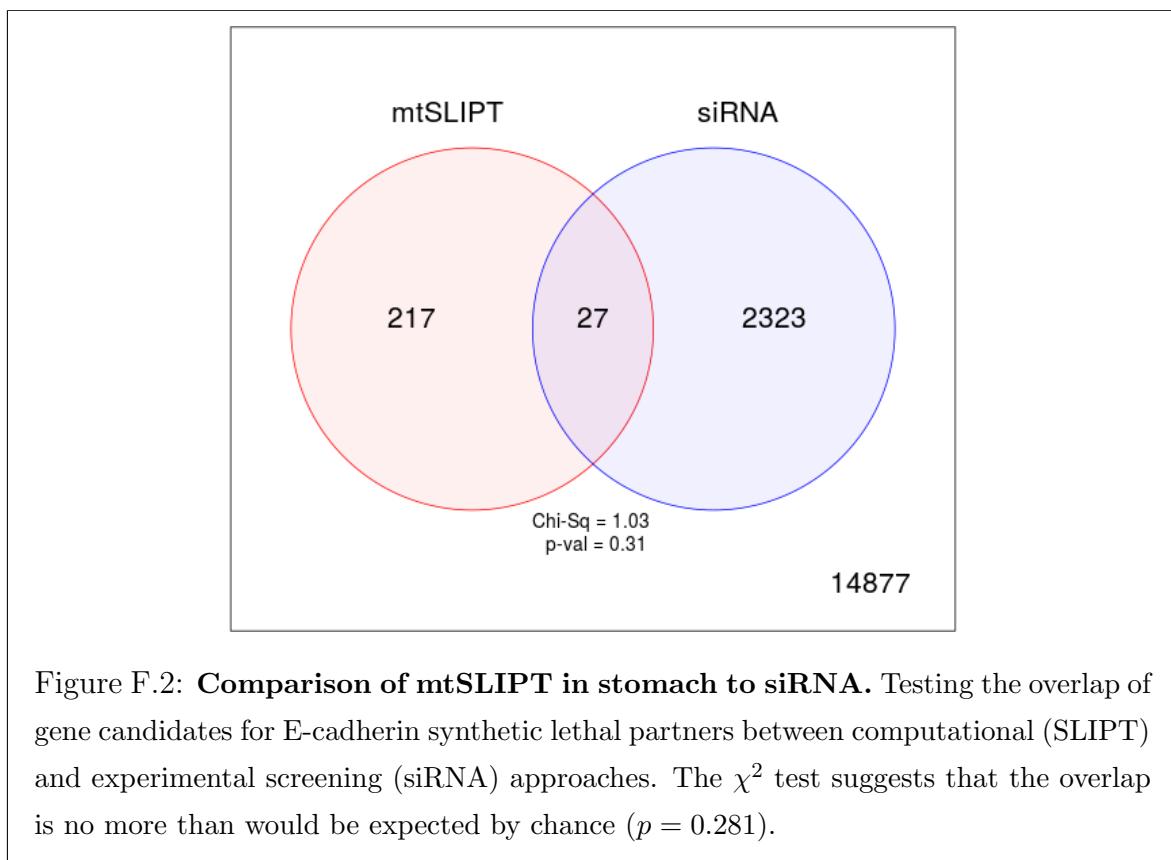


Table F.4: Pathway composition for *CDH1* partners from mtSLIPT and siRNA

Predicted only by SLIPT (217 genes)	Pathway Size	Genes Identified	p-value (FDR)
Eukaryotic Translation Elongation	87	57	2.8×10^{-120}
Peptide chain elongation	84	56	3.1×10^{-120}
Eukaryotic Translation Termination	84	55	2.8×10^{-117}
Viral mRNA Translation	82	54	4.1×10^{-116}
Nonsense Mediated Decay independent of the Exon Junction Complex	89	55	3.7×10^{-113}
Formation of a pool of free 40S subunits	94	55	2.8×10^{-109}
Nonsense-Mediated Decay	104	57	8.4×10^{-108}
Nonsense Mediated Decay enhanced by the Exon Junction Complex	104	57	8.4×10^{-108}
L13a-mediated translational silencing of Ceruloplasmin expression	104	56	3.4×10^{-105}
3' -UTR-mediated translational regulation	104	56	3.4×10^{-105}
GTP hydrolysis and joining of the 60S ribosomal subunit	105	56	1.4×10^{-104}
Eukaryotic Translation Initiation	112	56	2.8×10^{-100}
Cap-dependent Translation Initiation	112	56	2.8×10^{-100}
SRP-dependent cotranslational protein targeting to membrane	105	54	2.2×10^{-99}
Influenza Viral RNA Transcription and Replication	109	54	5.3×10^{-97}
Influenza Life Cycle	113	54	9.6×10^{-95}
Influenza Infection	118	55	1.7×10^{-94}
Translation	142	60	3.5×10^{-94}
Infectious disease	349	77	5.9×10^{-62}
Extracellular matrix organization	241	54	3×10^{-52}

Detected only by siRNA screen (2323 genes)	Pathway Size	Genes Identified	p-value (FDR)
Class A/1 (Rhodopsin-like receptors)	282	69	1.9×10^{-59}
GPCR ligand binding	363	78	2.7×10^{-54}
Peptide ligand-binding receptors	175	41	1.5×10^{-42}
G _{αi} signalling events	184	41	1.1×10^{-40}
Gastrin-CREB signalling pathway via PKC and MAPK	180	37	1.5×10^{-35}
G _{αq} signalling events	159	34	3.7×10^{-35}
DAP12 interactions	159	27	1.1×10^{-24}
VEGFA-VEGFR2 Pathway	91	19	1.0×10^{-23}
Downstream signal transduction	146	24	1.9×10^{-22}
Signaling by VEGF	99	19	2.6×10^{-22}
DAP12 signaling	149	24	4.2×10^{-22}
Organelle biogenesis and maintenance	264	34	4.3×10^{-20}
Downstream signaling of activated FGFR1	134	21	4.3×10^{-20}
Downstream signaling of activated FGFR2	134	21	4.3×10^{-20}
Downstream signaling of activated FGFR3	134	21	4.3×10^{-20}
Downstream signaling of activated FGFR4	134	21	4.3×10^{-20}
Signaling by ERBB2	146	22	5.3×10^{-20}
Signaling by FGFR	146	22	5.3×10^{-20}
Signaling by FGFR1	146	22	5.3×10^{-20}
Signaling by FGFR2	146	22	5.3×10^{-20}

Intersection of SLIPT and siRNA screen (23 genes)	Pathway Size	Genes Identified	p-value (FDR)
HS-GAG degradation	21	4	4.9×10^{-6}
Retinoid metabolism and transport	39	5	4.9×10^{-6}
Platelet activation, signaling and aggregation	186	13	4.9×10^{-6}
Signaling by NOTCH4	11	3	4.9×10^{-6}
G _{αs} signalling events	100	8	5×10^{-6}
Defective EXT2 causes exostoses 2	12	3	5×10^{-6}
Defective EXT1 causes exostoses 1, TRPS2 and CHDS	12	3	5×10^{-6}
Class A/1 (Rhodopsin-like receptors)	289	18	2.2×10^{-5}
Signaling by PDGF	173	11	2.9×10^{-5}
Circadian Clock	34	4	2.9×10^{-5}
Signaling by ERBB4	139	9	4.3×10^{-5}
Role of LAT2/NTAL/LAB on calcium mobilization	99	7	4.4×10^{-5}
Peptide ligand-binding receptors	181	11	4.5×10^{-5}
Defective B4GALT7 causes EDS, progeroid type	19	3	4.5×10^{-5}
Defective B3GAT3 causes JDSSDH	19	3	4.5×10^{-5}
Signaling by NOTCH	80	6	4.5×10^{-5}
G _{αq} signalling events	164	10	5.1×10^{-5}
Response to elevated platelet cytosolic Ca ²⁺	84	6	7.1×10^{-5}
Signaling by ERBB2	148	9	7.1×10^{-5}
Signaling by SCF-KIT	129	8	8.3×10^{-5}

Table F.5: Pathways for *CDH1* partners from mtSLIPT in stomach cancer

Reactome Pathway	Over-representation	Permutation
Eukaryotic Translation Elongation	2×10^{-128}	$< 8.802 \times 10^{-4}$
Peptide chain elongation	2×10^{-128}	$< 8.802 \times 10^{-4}$
Eukaryotic Translation Termination	2.3×10^{-125}	$< 8.802 \times 10^{-4}$
Viral mRNA Translation	2.5×10^{-124}	$< 8.802 \times 10^{-4}$
Nonsense Mediated Decay independent of the Exon Junction Complex	8.6×10^{-124}	$< 8.802 \times 10^{-4}$
Nonsense-Mediated Decay	5.2×10^{-117}	$< 8.802 \times 10^{-4}$
Nonsense Mediated Decay enhanced by the Exon Junction Complex	5.2×10^{-117}	$< 8.802 \times 10^{-4}$
Formation of a pool of free 40S subunits	1.6×10^{-116}	$< 8.802 \times 10^{-4}$
L13a-mediated translational silencing of Ceruloplasmin expression	1.3×10^{-111}	$< 8.802 \times 10^{-4}$
3' -UTR-mediated translational regulation	1.3×10^{-111}	$< 8.802 \times 10^{-4}$
GTP hydrolysis and joining of the 60S ribosomal subunit	6.2×10^{-111}	$< 8.802 \times 10^{-4}$
SRP-dependent cotranslational protein targeting to membrane	2.9×10^{-108}	$< 8.802 \times 10^{-4}$
Eukaryotic Translation Initiation	3×10^{-106}	$< 8.802 \times 10^{-4}$
Cap-dependent Translation Initiation	3×10^{-106}	$< 8.802 \times 10^{-4}$
Influenza Viral RNA Transcription and Replication	5.1×10^{-103}	$< 8.802 \times 10^{-4}$
Influenza Infection	1.5×10^{-102}	$< 8.802 \times 10^{-4}$
Translation	3.7×10^{-101}	$< 8.802 \times 10^{-4}$
Influenza Life Cycle	1.4×10^{-100}	$< 8.802 \times 10^{-4}$
GPCR downstream signaling	1×10^{-80}	0.034498
Hemostasis	1.4×10^{-78}	0.086519
Extracellular matrix organization	1.5×10^{-67}	0.040016
Developmental Biology	1.8×10^{-66}	0.18385
Infectious disease	7.3×10^{-66}	0.068426
Signalling by NGF	8.5×10^{-63}	0.16798
Metabolism of lipids and lipoproteins	4.9×10^{-58}	0.51411
Platelet activation, signaling and aggregation	2.7×10^{-55}	0.081717
GPCR ligand binding	7.3×10^{-55}	0.28898
Signaling by PDGF	8.4×10^{-55}	0.16025
Class A/1 (Rhodopsin-like receptors)	3.2×10^{-54}	0.22801
Fc epsilon receptor (FCER1) signaling	6.2×10^{-53}	0.15229
Adaptive Immune System	5.1×10^{-52}	0.037698
Signaling by ERBB4	5.9×10^{-52}	0.10088
Axon guidance	8.8×10^{-52}	0.40234
Formation of the ternary complex, and subsequently, the 43S complex	1.6×10^{-51}	0.00088017
Ribosomal scanning and start codon recognition	2.2×10^{-50}	0.00088017
Translation initiation complex formation	2.2×10^{-50}	0.0017305
NGF signalling via TRKA from the plasma membrane	6.7×10^{-50}	0.28811
Activation of the mRNA upon binding of the cap-binding complex and eIFs, and subsequent binding to 43S	7.1×10^{-50}	0.0017305
Transmembrane transport of small molecules	1.8×10^{-49}	0.081229
Signaling by ERBB2	5.9×10^{-49}	0.11896
Rho GTPase cycle	3.6×10^{-48}	0.035735
Gαs signalling events	1.1×10^{-47}	0.0088487
Downstream signal transduction	1.7×10^{-47}	0.11909
Signaling by FGFR	1.7×10^{-47}	0.11896
Signaling by FGFR1	1.7×10^{-47}	0.11896
Signaling by FGFR2	1.7×10^{-47}	0.11896
Signaling by FGFR3	1.7×10^{-47}	0.11896
Signaling by FGFR4	1.7×10^{-47}	0.11896
DAP12 interactions	1.9×10^{-47}	0.28811
DAP12 signaling	1×10^{-46}	0.12442

Over-representation (hypergeometric test) and Permutation p-values adjusted for multiple tests across pathways (FDR).

Significant pathways are marked in bold (FDR < 0.05) and italics (FDR < 0.1).

Table F.6: Pathways for *CDH1* partners from mtSLIPT in stomach and siRNA screen

Reactome Pathway	Over-representation	Permutation
Signaling by NOTCH4	4.9×10^{-6}	0.050121
HS-GAG degradation	4.9×10^{-6}	0.013193
Platelet activation, signaling and aggregation	4.9×10^{-6}	0.28053
Retinoid metabolism and transport	4.9×10^{-6}	0.0927
Defective EXT2 causes exostoses 2	5×10^{-6}	0.14898
Defective EXT1 causes exostoses 1, TRPS2 and CHDS	5×10^{-6}	0.14898
<i>Gαs</i> signalling events	5×10^{-6}	0.048426
Class A/1 (Rhodopsin-like receptors)	2.2×10^{-5}	0.60435
Signaling by PDGF	2.9×10^{-5}	0.43907
Circadian Clock	2.9×10^{-5}	0.012519
Signaling by ERBB4	4.3×10^{-5}	0.12835
Role of LAT2/NTAL/LAB on calcium mobilization	4.4×10^{-5}	0.27344
Defective B4GALT7 causes EDS, progeroid type	4.5×10^{-5}	0.23536
Defective B3GAT3 causes JDSSDHD	4.5×10^{-5}	0.23536
Peptide ligand-binding receptors	4.5×10^{-5}	0.41193
Signaling by NOTCH	4.5×10^{-5}	0.10912
<i>Gαq</i> signalling events	5.1×10^{-5}	0.28937
Signaling by ERBB2	7.1×10^{-5}	0.50797
Response to elevated platelet cytosolic Ca^{2+}	7.1×10^{-5}	0.38513
Signaling by SCF-KIT	8.3×10^{-5}	0.55412
PI3K events in ERBB4 signaling	0.0001	0.24486
PIP3 activates AKT signaling	0.0001	0.24486
Collagen formation	0.0001	0.15296
PI3K events in ERBB2 signaling	0.0001	0.24486
PI-3K cascade:FGFR1	0.0001	0.24486
PI-3K cascade:FGFR2	0.0001	0.24486
PI-3K cascade:FGFR3	0.0001	0.24486
PI-3K cascade:FGFR4	0.0001	0.24486
Growth hormone receptor signaling	0.0001	0.057494
PI3K Cascade	0.00011	0.20906
Effects of PIP2 hydrolysis	0.00012	0.14898
A tetrasaccharide linker sequence is required for GAG synthesis	0.00012	0.29766
PI3K/AKT activation	0.00013	0.24486
GAB1 signalosome	0.00013	0.4648
Diseases associated with glycosaminoglycan metabolism	0.00013	0.050121
Diseases of glycosylation	0.00013	0.050121
Heparan sulfate/heparin (HS-GAG) metabolism	0.00016	0.19
HS-GAG biosynthesis	0.00016	0.29681
Integrin alphaIIb beta3 signaling	0.00016	0.63007
Interferon gamma signaling	0.00018	0.43088
Gastrin-CREB signalling pathway via PKC and MAPK	0.00018	0.77958
Chemokine receptors bind chemokines	0.00023	0.62702
Downstream signal transduction	0.00027	0.54921
Platelet homeostasis	0.00029	0.24577
IRS-mediated signalling	0.00029	0.31766
<i>Gαi</i> signalling events	0.00029	$< 2.749 \times 10^{-4}$
Diseases of signal transduction	0.00029	0.65733
Signaling by activated point mutants of FGFR1	0.00029	0.24892
FGFR1c ligand binding and activation	0.00029	0.24892
Signaling by NOTCH3	0.00029	0.017419

Over-representation (hypergeometric test) and Permutation p-values adjusted for multiple tests across pathways (FDR). Significant pathways are marked in bold (FDR < 0.05) and italicics (FDR < 0.1).

Table F.7: Candidate synthetic lethal metagenes against *CDH1* from mtSLIPT in stomach cancer

Pathway	ID	Observed	Expected	χ^2 value	p-value	p-value (FDR)
Prostacyclin signalling through prostacyclin receptor	392851	1	10.07	26.53	1.7307×10^{-6}	0.0028590
Cell surface interactions at the vascular wall	202733	3	10.07	21.11	2.6107×10^{-5}	0.00642330
The NLRP1 inflammasome	844455	3	10.07	21.11	2.6107×10^{-5}	0.00642330
Innate Immune System	168249	6	10.07	21.07	2.6548×10^{-5}	0.00642330
Keratan sulfatekeratin metabolism	1638074	4	10.07	20.65	3.2861×10^{-5}	0.00642330
Keratan sulfate biosynthesis	2022854	4	10.07	20.65	3.2861×10^{-5}	0.00642330
Signaling by SCF-KIT	1433557	5	10.07	20.64	3.3045×10^{-5}	0.00642330
VEGFA-VEGFR2 Pathway	4420097	5	10.07	20.64	3.3045×10^{-5}	0.00642330
ERK1 activation	110056	21	10.07	20.12	4.277×10^{-5}	0.00642330
Cholesterol biosynthesis	191273	21	10.07	20.12	4.277×10^{-5}	0.00642330
G2 Phase	68911	21	10.07	20.12	4.277×10^{-5}	0.00642330
p130Cas linkage to MAPK signaling for integrins	372708	2	10.07	19.08	7.1872×10^{-5}	0.00651340
cGMP effects	418457	8	10.07	19.01	7.4597×10^{-5}	0.00651340
Regulation of cytoskeletal remodeling and cell spreading by IPP complex components	446388	8	10.07	19.01	7.4597×10^{-5}	0.00651340
Post-translational modification: synthesis of GPI-anchored proteins	163125	20	10.07	18.59	9.1878×10^{-5}	0.00651340
Fcgamma receptor (FCGR) dependent phagocytosis	2029480	3	10.07	17.95	0.00012676	0.00651340
A third proteolytic cleavage releases NICD	157212	7	10.07	17.90	0.00012995	0.00651340
Signalling by NGF	166520	7	10.07	17.90	0.00012995	0.00651340
Signaling by VEGF	194138	7	10.07	17.90	0.00012995	0.00651340
Regulation of thyroid hormone activity	350864	7	10.07	17.90	0.00012995	0.00651340
Nitric oxide stimulates guanylate cyclase	392154	7	10.07	17.90	0.00012995	0.00651340
Platelet homeostasis	418346	7	10.07	17.90	0.00012995	0.00651340
Termination of translesion DNA synthesis	5656169	20	10.07	17.46	0.00016155	0.00651340
PI3K events in ERBB4 signaling	1250342	4	10.07	17.26	0.00017862	0.00651340
PIP3 activates AKT signaling	1257604	4	10.07	17.26	0.00017862	0.00651340

Strongest candidate SL partners for *CDH1* by mtSLIPT with observed and expected mutant samples with low expression of partner metagenes