A Bayesian Integrative Model for Genetical Genomics with Spatially Informed Variable Selection

Alberto Cassese[*][†], Michele Guindani[†] and Marina Vannucci[*]

April 10, 2014

Gene Ontology Analyses

We report full tables of the Gene Ontology analyses performed on the lists of target genes and CGH probes selected by our model. More specifically, Table 1 shows the results from the enrichment analysis of the selected target genes, and Table 2 reports the results from the enrichment analysis of the selected CGH probes. We report, for each molecular function, the list of genes identified by our model. Moreover, we show the counts of the genes identified by our model and the percentages of the full list of genes belonging to that molecular function. The rightmost two columns show p-values of enrichment, before and after the Benjamini correction.

---

[*] Department of Statistics, Rice University, Houston, TX, USA

[†] Department of Biostatistics, The University of Texas MD Anderson Cancer Center, Houston, TX, USA

Gene Ontology Analysis of Target Genes

Table 1: Case study: Gene Ontology analysis of the target genes with expressions that were found to be significantly associated with the CNVs.

Molecular Function Term	Genes	Count	Percent	P-Value	Benjamini Correction
Structural molecule activity	NELL1, CLDN8, FBN2, KRT13, KRT23, SPRR1A, SPRR3, STATH, TFPI2, UPK1B	10	16,7%	0.00019	0.031
Eukaryotic cell surface binding	FGA, FGB, FGG	3	5%	0.00140	0.11
Protein binding, bridging	FGA, FGB, FGG, SPRR1A	4	6,7%	0.0037	0.18
Cell surface binding	FGA, FGB, FGG	3	5%	0.0041	0.16
Calcium ion binding	NELL1, CALB1, DLK1, FBN2, FGG, MMP10, NRXN1, TKTL1	8	13,3%	0.03	0.64

Continued on next page

Table 1 – continued from previous page

Molecular Function Term	**Genes**	**Count**	**Percent**	**P-Value**	**Benjamini Correction**
Extracellular matrix structural constituent	FBN2, STATH, TFPI2	3	5%	0.033	0.6
Serine-type endopeptidase inhibitor activity	SERPINB13, SERPINB4, TFPI2	3	5%	0.037	0.59
Heme binding	CYP26A1, CYP4F11, CYP4F3	3	5%	0.061	0.73
Retinoid binding	ADH7, CYP26A1	2	3.3%	0.067	0.72
Tetrapyrrole binding	CYP26A1, CYP4F11, CYP4F3	3	5%	0.068	0.69
Isoprenoid binding	ADH7, CYP26A1	2	3.3%	0.074	0.68
Iron ion binding	CYP26A1, CYP4F11, CYP4F3, PAH	4	6.7%	0.081	0.69
Carboxylic acid binding	AZGP1, CYP26A1, PAH	3	5%	0.082	0.67
Endopeptidase inhibitor activity	SERPINB13, SERPINB4, TFPI2	3	5%	0.083	0.64
Peptidase inhibitor activity	SERPINB13, SERPINB4, TFPI2	3	5%	0.091	0.65

Gene Ontology Analysis of CGH probes

Table 2: Case study: Gene Ontology analysis of the CGH genes that were found to be significantly associated with the target genes.

Molecular Function Term	**Genes**	**Count**	**Percent**	**P-Value**	**Benjamini Correction**
Ephrin receptor activity	EPHA3, EPHA6, EPHB1	3	0.6%	0.0012	0.16
Protein kinase activity	EPHA3, EPHA6, EPHB1, NEK11, IRAK2, KALRN, PRKCI	7	1.4%	0.0088	0.48
Phosphatidylinositol-4,5-bisphosphate 3-kinase activity	TRAT1, PIK3CB	2	0.4%	0.015	0.52
Transmembrane receptor protein tyrosine kinase activity	EPHA3, EPHA6, EPHB1	3	0.6%	0.017	0.47
Rho guanylnucleotide exchange factor activity	MCF2L2, ECT2, KALRN	3	0.6%	0.021	0.46
Phospholipase activity	CASR, PLD1, PLCXD2, PHLDB2	3	0.6%	0.0260	0.4700

Continued on next page

Table 2 – continued from previous page

Molecular Function Term	**Genes**	**Count**	**Percent**	**P-Value**	**Benjamini Correction**
Phosphoric diester hydrolase activity	CASR, PLD1, PLCXD2, PHLDB2	3	0.6%	0.026	0.43
ATP binding	EPHA3, EPHA6, EPHB1, NEK11, IRAK2, KALRN, LARS2, PIK3CB, PRKCI, RFC4	10	2%	0.028	0.41
Ras guanyl-nucleotide exchange factor activity	MCF2L2, ECT2, KALRN	3	0.6%	0.029	0.38
Adenyl ribonucleotide binding	EPHA3, EPHA6, EPHB1, NEK11, IRAK2, KALRN, LARS2, PIK3CB, PRKCI, RFC4	10	2%	0.03	0.37
Phospho-inositide 3-kinase activity	TRAT1, PIK3CB	2	0.4%	0.035	0.38

Continued on next page

Table 2 – continued from previous page

Molecular Function Term	**Genes**	**Count**	**Percent**	**P-Value**	**Benjamini Correction**
Lipase activity	CASR, PLD1, PLCXD2, PHLDB2	3	0.6%	0.036	0.37
Adenyl nucleotide binding	EPHA3, EPHA6, EPHB1, NEK11, IRAK2, KALRN, LARS2, PIK3CB, PRKCI, RFC4	10	2%	0.041	0.38
Purine nucleoside binding	EPHA3, EPHA6, EPHB1, NEK11, IRAK2, KALRN, LARS2, PIK3CB, PRKCI, RFC4	10	2%	0.044	0.38

Continued on next page

Table 2 – continued from previous page

Molecular Function Term	**Genes**	**Count**	**Percent**	**P-Value**	**Benjamini Correction**
Nucleoside binding	EPHA3, EPHA6, EPHB1, NEK11, IRAK2, KALRN, LARS2, PIK3CB, PRKCI, RFC4	10	2%	0.046	0.37
Guanyl-nucleotide exchange factor activity	MCF2L2, ECT2, KALRN	3	0.6%	0.076	0.52
Lipid kinase activity	TRAT1, PIK3CB	2	0.4%	0.084	0.53
Protein tyrosine kinase activity	EPHA3, EPHA6, EPHB1	3	0.6%	0.089	0.53
Phospholipase C activity	CASR, PLCXD2, PHLDB2	2	0.4%	0.089	0.52
Purine ribonucleotide binding	EPHA3, EPHA6, EPHB1, NEK11, IRAK2, KALRN, LARS2, PIK3CB, PRKCI, RFC4	10	2%	0.09	0.5

Continued on next page

Table 2 – continued from previous page

Molecular Function Term	**Genes**	**Count**	**Percent**	**P-Value**	**Benjamini Correction**
Ribonucleotide binding	EPHA3, EPHA6, EPHB1, NEK11, IRAK2, KALRN, LARS2, PIK3CB, PRKCI, RFC4	10	2%	0.09	0.5
Phospholipid binding	ITPR1, PLD1, PRKCI	3	0.6%	0.099	0.52