A. Bellacosa et al. Altered Gene Expression in Morphologically Normal Epithelial Cells from Heterozygous Carriers of BRCA1 or BRCA2 Mutations

SUPPLEMENTARY INFORMATION

Correlation between Affymetrix Data and Low-Density Arrays

In order to determine the correlation between fold changes from the original Affymetrix microarray data and the LDA real-time RT-PCR data, we measured Spearman's ρ for select candidate biomarkers, divided with respect to the genotypes and epithelial cultures in which the biomarkers were originally identified. The results, shown in Table S1, indicate that the highest correlations were found for breast and ovarian candidate biomarkers originally associated with the *BRCA1* genotype.

Exploratory Data Mining

For initial exploratory analyses, we considered the pre-processed data for breast and ovarian samples for each of the three genotypes (*BRCA1, BRCA2 and WT*). With 6 biological replicates in each condition, this dataset consisted of 36 samples and corresponding expression profiles for 54675 probe sets. In order to reduce the dimensionality, we applied several variation filters to this dataset. Specifically, we first removed probe sets whose maximum expression intensity was less than 200 across the 36 samples. We then removed probe sets whose coefficient of variation (computed as a percentage ratio of standard deviation to mean) was less than 75, i.e., we retained only the top 25% of the most variable probe sets. This resulted in 4635 most variable probe sets being retained in our dataset.

We applied the bi-clustering procedure based on non-negative matrix factorization (NMF) (Pascual-Montano et al., BMC Bioinformatics, 2006) to this reduced dataset consisting of breast and ovarian samples. This method potentially identifies sub-groups of genes that are strongly correlated with sub-groups of samples. The best model based on 200 random runs of the NMF algorithm was chosen. This model identified six clusters of probe sets, with clusters 1 and 5 preferentially correlating with ovarian epithelial cultures, and clusters 2, 3, 4 and 6 preferentially correlating with breast epithelial cultures. However, no difference related to the genotype was identified in this approach.

We then considered the breast and ovarian datasets separately and further filtered the datasets as above. These resulted in 1832 and 2553 probe sets, respectively, for breast and ovarian samples. We applied standard NMF in conjunction with consensus clustering (1, 2) based on 200 runs of the algorithm to the dataset for each target organ. In addition, we applied hierarchical clustering using average linkage with correlation as the metric. Neither approach was able to identify a clear separation of the genotypes (BRCA1, BRCA2 and WT) within each target organ (Figure S1).

Pathway and ontology analyses – Methods and discussion

An association analysis was conducted to identify the association between BRCA1 and BRCA2 profiles. We identified 4 datasets *viz.* i,ii,iii and iv described in main methods section. The datasets except Hedenfalk *et al.* study (iv), data were normalized using RMA and lists of differentially expressed genes were obtained by applying Linear Models for Microarray Data (LIMMA) (3) using a p-value cutoff of 0.001. These analyses involved a much larger number of microarrays relative to our study for the various comparisons of interest. LIMMA is suitable for

analyzing microarray data involving factorial designs (multiple conditions) and enables to extract relevant contrasts (treatment combinations) of interest for further analysis. For the Hedenfalk *et al.* study (4), pre-normalized data were obtained from supplemental information (http://www.nejm.org/general/content/supplemental/hedenfalk/index.html) and LIMMA was applied to obtain lists of differentially expressed genes between BRCA1 vs. MCF10-A, BRCA2 vs. MCF10-A, sporadic vs. MCF10-A, BRCA1 vs. BRCA2, BRCA2 vs. sporadic and BRCA1 vs. sporadic. We also generated a manually curated list of over 180 genes involved in DNA repair by parsing functional information from GeneRIF in Entrez Gene database.

In addition to pathway and association analyses, gene ontology analysis was performed to identify overrepresented biological processes on upregulated and downregulated genes, separately for all four comparisons. This approach allowed the identification of overrepresented categories for up- and down-regulated genes (see Table S2). This analysis revealed that a significant number of down-regulated genes in breast *BRCA1* heterozygous cells are involved in major cellular processes such as differentiation, development, proliferation, adhesion and apoptosis. On the other hand, significant numbers of up-regulated genes are involved in biosynthetic metabolic processes, including transcription, splicing, DNA replication and repair (Fig. 2, Table S2). Likewise, for down-regulated genes in breast *BRCA2* heterozygous cells, processes such as small GTPase-mediated signal transduction and cell cycle progression are enriched. For up-regulated genes in breast *BRCA2* heterozygous cells, biological processes such as immunologic and inflammatory processes, adhesion, oocyte differentiation and ovulation are over-represented (Fig. 2, Table S2).

In the case of ovary, in *BRCA1* heterozygous cells, up-regulated genes show significant enrichment for processes involved in development, differentiation and cell morphogenesis, and

down-regulated genes are associated with DNA repair, replication and cell cycle (Fig. 2, Table S2). *BRCA2* heterozygous cells show an enrichment of genes involved in catabolic processes , antigen processing, DNA fragmentation and G2/M transition (Fig. 2, Table S2).

Table S1 – Correlation between fold changes from microarray and low density array (LDA) data for candidate biomarkers in breast and ovarian cultures for different genotypes (Spearman's ρ are shown)

Candidate markers originally identified in (culture, genotype):	Correlation in culture of *BRCA1* genotype	Correlation in culture of *BRCA2* genotype
Breast, *BRCA1*	0.94	0.85
Breast, *BRCA2*	0.79	0.09
Ovarian, *BRCA1*	0.94	0.38
Ovarian, *BRCA2*	0.6	0.2

Table S2. Over-represented gene ontology categories for clusters of up- and down-regulated genes in breast and ovary BRCA1 and BRCA21 mutant cells vs. WT

Cluster 1 Pvalue	OddsRatio	ExpCount	Count	Size	Term
0.001	2.069	21	35	2984	multicellular organismal process
0	2.268	19	34	2795	developmental process
0.002	2.45	7	16	1033	organ development
0	2.517	10	22	1506	cell differentiation
0.008	2.522	5	11	667	cell proliferation
0.007	2.534	5	11	664	cell adhesion
0.001	2.684	7	17	1018	negative regulation of biological process
0.004	2.742	4	11	617	proteolysis
0.003	2.751	5	12	676	apoptosis
0.002	2.809	5	13	724	cell death
0.003	3.022	4	10	508	regulation of cell cycle
0.007	3.055	3	8	396	cell morphogenesis
0.007	3.134	3	8	437	system development
0.004	3.167	3	9	434	cytoskeleton organization and biogenesis
0.008	3.676	2	6	245	enzyme linked receptor protein signaling pathway
0.006	3.892	2	6	232	growth
0	4.54	2	10	346	cell motility
0.002	4.744	1	6	192	negative regulation of progression through cell cycle
0.002	5.017	1	6	182	actin cytoskeleton organization and biogenesis
0.003	11.976	0	3	39	icosanoid metabolic process
0	12.11	1	9	123	ectoderm development
0.002	13.477	0	3	35	epidermis morphogenesis
0.01	14.976	0	2	21	response to UV
0.006	18.976	0	2	17	prostaglandin metabolic process
0.006	18.976	0	2	17	antigen processing and presentation of peptide or polysaccharide antigen via MHC class II
0.001	21.584	0	3	23	keratinocyte differentiation
0	25.4	0	3	20	keratinization
0.004	25.884	0	2	13	regulation of striated muscle development
0	30.85	0	3	17	peptide cross-linking
0.001	47.473	0	2	8	protein homotetramerization
0.001	47.473	0	2	8	regulation of nucleotide metabolic process
0	72.029	0	3	9	endothelial cell migration
0.007	Inf	0	1	1	pyrimidine ribonucleoside catabolic process
0.007	Inf	0	1	1	hydrogen peroxide biosynthetic process
0.007	Inf	0	1	1	regulatory T cell differentiation
0.007	Inf	0	1	1	isopeptide cross-linking via N6-(L-isoglutamyl)-L-lysine
0.007	Inf	0	1	1	pyrimidine salvage
0.007	Inf	0	1	1	cytidine deamination
0.007	Inf	0	1	1	skeletal muscle regeneration
0.007	Inf	0	1	1	negative regulation of low-density lipoprotein receptor catabolic process
0.007	Inf	0	1	1	extracellular transport
0.007	Inf	0	1	1	cytidine metabolic process

Cluster

2					
Pvalue	OddsRatio	ExpCount	Count	Size	Term
0.007	1.645	60	73	6735	cellular metabolic process
0.008	1.722	22	33	2477	regulation of metabolic process
0.003	1.753	60	75	6788	primary metabolic process
0.003	1.869	20	32	2239	transcription
0.002	1.901	20	32	2208	regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process
0.001	1.918	52	70	5890	macromolecule metabolic process
0.001	2.019	18	31	2026	regulation of transcription, DNA-dependent
0.001	2.041	18	32	2083	RNA biosynthetic process
0.008	2.97	3	8	324	response to endogenous stimulus
0.003	3.066	4	10	476	RNA metabolic process
0.005	3.53	2	7	239	mRNA processing
0.006	3.969	2	6	182	induction of apoptosis
0.005	3.992	2	6	181	DNA replication
0.001	4.128	2	8	237	DNA repair
0	4.836	2	8	204	RNA splicing
0.002	45.498	0	2	7	ATP-dependent proteolysis
0.009	Inf	0	1	1	ethanolamine metabolic process
0.009	Inf	0	1	1	phosphatidylethanolamine biosynthetic process
0.009	Inf	0	1	1	deoxyribonucleoside catabolic process
0.009	Inf	0	1	1	nucleobase catabolic process
0.009	Inf	0	1	1	G0 to G1 transition
0.009	Inf	0	1	1	deoxyribonucleoside monophosphate biosynthetic process
0.009	Inf	0	1	1	pyrimidine deoxyribonucleoside monophosphate metabolic process
0.009	Inf	0	1	1	otic vesicle formation
0.009	Inf	0	1	1	glucose 1-phosphate metabolic process
0.009	Inf	0	1	1	endosomal lumen acidification
0.009	Inf	0	1	1	natural killer cell degranulation
0.009	Inf	0	1	1	cytotoxic T cell degranulation
0.009	Inf	0	1	1	pronephros development
0.009	Inf	0	1	1	negative regulation of astrocyte differentiation
0.009	Inf	0	1	1	adult somatic muscle development
0.009	Inf	0	1	1	uracil catabolic process
0.009	Inf	0	1	1	thymidine catabolic process
0.009	Inf	0	1	1	dTMP biosynthetic process

Cluster 3					
Pvalue	OddsRatio	ExpCount	Count	Size	Term
0.009	1.9	21	31	4365	biological regulation
0.006	1.971	18	28	3731	regulation of cellular process
0.004	3.764	2	7	418	small GTPase mediated signal transduction
0	3.788	3	11	688	cell cycle process
0.01	4.12	1	5	267	mitotic cell cycle
0.001	4.141	2	9	504	regulation of progression through cell cycle
0.008	16.723	0	2	27	DNA replication initiation
0.007	17.421	0	2	26	monosaccharide biosynthetic process
0.003	27.893	0	2	17	antigen processing and presentation of peptide or polysaccharide antigen via MHC class II
0.001	59.81	0	2	9	endothelial cell migration

0.01	206	0	1	2	isotype switching to IgA isotypes
0.01	206	0	1	2	positive regulation of isotype switching to IgA isotypes
0.005	Inf	0	1	1	succinate transport
0.005	Inf	0	1	1	inositol biosynthetic process
0.005	Inf	0	1	1	skeletal muscle regeneration

Cluster 4					
Pvalue	OddsRatio	ExpCount	Count	Size	Term
0.005	6.572	1	4	272	behavior
0.009	15.202	0	2	57	activation of immune response
0.008	16.4	0	2	53	immunoglobulin mediated immune response
0.001	20.258	0	3	67	humoral immune response
0.003	26.177	0	2	34	activation of plasma proteins during acute inflammatory response
0.002	33.525	0	2	27	complement activation, classical pathway
0.008	203.145	0	1	3	L-serine biosynthetic process
0.008	203.145	0	1	3	negative regulation of gliogenesis
0.008	203.145	0	1	3	Golgi to plasma membrane protein transport
0.003	Inf	0	1	1	negative regulation of Golgi to plasma membrane protein transport
0.003	Inf	0	1	1	regulation of Golgi to plasma membrane CFTR protein transport
0.003	Inf	0	1	1	cytoplasmic sequestering of CFTR protein
0.003	Inf	0	1	1	negative regulation of astrocyte differentiation

Cluster 5					
Pvalue	OddsRatio	ExpCount	Count	Size	Term
0.005	22.077	0	2	73	cell-matrix adhesion
0.01	123.578	0	1	7	positive regulation of tumor necrosis factor biosynthetic process
0.01	123.578	0	1	7	oocyte differentiation
0.01	123.578	0	1	7	pH reduction
0.01	123.578	0	1	7	regulation of cellular pH
0.009	148.306	0	1	6	ovulation (sensu Mammalia)
0.009	148.306	0	1	6	positive regulation of interleukin-6 biosynthetic process
0.007	185.397	0	1	5	phagocytosis, engulfment
0.007	185.397	0	1	5	positive regulation of erythrocyte differentiation
0.007	185.397	0	1	5	stress fiber formation
0.006	247.216	0	1	4	negative regulation of bone mineralization
0.004	370.853	0	1	3	oocyte maturation
0.004	370.853	0	1	3	antral ovarian follicle growth
0.001	Inf	0	1	1	initiation of primordial ovarian follicle growth
0.001	Inf	0	1	1	endosomal lumen acidification

Cluster 6					
Pvalue	OddsRatio	ExpCount	Count	Size	Term
0.005	2.136	10	19	1529	cellular developmental process
0.001	2.29	13	25	1984	multicellular organismal development
0	2.586	12	25	1791	anatomical structure development
0.001	3.809	3	9	417	nervous system development
0.01	4.094	1	5	196	neuron differentiation

0.004	4.231	2	6	230	neurogenesis
0	4.32	4	16	664	cell adhesion
0.007	4.474	1	5	180	cell projection organization and biogenesis
0.007	4.474	1	5	180	cell part morphogenesis
0.006	5.981	1	4	108	wound healing
0.001	6.841	1	5	120	homophilic cell adhesion
0.001	10.765	0	4	62	morphogenesis of an epithelium
0.006	19.099	0	2	18	neural plate development
0.006	20.374	0	2	17	neural tube formation
0.004	23.512	0	2	15	negative regulation of angiogenesis
0.003	27.792	0	2	13	neural tube closure
0.001	61.171	0	2	7	apical protein localization
0	101.967	0	2	5	establishment of planar polarity
0.007	Inf	0	1	1	PML body organization and biogenesis
0.007	Inf	0	1	1	intermediate mesoderm development
0.007	Inf	0	1	1	synaptic vesicle priming
0.007	Inf	0	1	1	androgen catabolic process

Cluster 7					
Pvalue	OddsRatio	ExpCount	Count	Size	Term
0.005	2.094	13	22	2317	cellular component organization and biogenesis
0.008	2.542	5	11	883	response to stress
0.003	3.642	2	8	465	regulation of progression through cell cycle
0	5.082	2	8	324	response to endogenous stimulus
0.002	5.083	1	6	237	DNA repair
0	5.343	2	9	352	chromosome organization and biogenesis
0.003	5.497	1	5	181	DNA replication
0	6.547	3	16	648	cell cycle process
0.009	7.474	0	3	113	cell cycle
0	9.616	1	8	177	microtubule-based process
0.001	9.848	0	4	82	phosphoinositide-mediated signaling
0.001	10.526	0	4	77	interphase
0.002	12.93	0	3	47	viral reproduction
0.009	15.592	0	2	26	ear morphogenesis
0.009	15.592	0	2	26	regulation of progression through mitotic cell cycle
0.008	16.271	0	2	25	establishment of organelle localization
0.001	17.255	0	3	36	viral infectious cycle
0	17.45	1	15	231	mitotic cell cycle
0	18.921	1	18	248	M phase
0	19.855	1	14	185	cell division
0.001	21.913	0	3	29	sister chromatid segregation
0.001	22.058	0	3	30	chromosome segregation
0	24.275	1	18	198	mitosis
0.004	24.965	0	2	17	positive regulation of progression through cell cycle
0.003	26.751	0	2	16	mitotic chromosome condensation
0.002	34.054	0	2	13	G2/M transition of mitotic cell cycle
0	36.744	0	4	25	cytokinesis
0.002	41.629	0	2	11	regulation of exit from mitosis
0.001	46.836	0	2	10	deoxyribonucleotide metabolic process
0.001	49	0	2	10	spindle organization and biogenesis
0.001	53.531	0	2	9	regulation of viral genome replication
0	71.318	0	3	11	mitotic spindle organization and biogenesis

0.001	74.955	0	2	7	embryonic skeletal morphogenesis
0	93.701	0	2	6	retroviral genome replication
0	124.945	0	2	5	sensory perception of temperature stimulus
0.005	Inf	0	1	1	cytokinesis after mitosis
0.005	Inf	0	1	1	centric heterochromatin formation
0.005	Inf	0	1	1	maintenance of DNA methylation
0.005	Inf	0	1	1	deoxyribonucleoside monophosphate biosynthetic process
0.005	Inf	0	1	1	pyrimidine deoxyribonucleoside monophosphate metabolic process
0.005	Inf	0	1	1	DNA mediated transformation
0.005	Inf	0	1	1	positive regulation of retroviral genome replication
0.005	Inf	0	1	1	dTMP biosynthetic process

Cluster 8					
Pvalue	OddsRatio	ExpCount	Count	Size	Term
0.006	3.931	4	9	2815	developmental process
0.01	4.695	1	5	1033	organ development
0.003	4.931	4	9	2941	multicellular organismal process
0.006	9.235	0	3	289	biopolymer catabolic process
0.004	10.953	0	3	245	enzyme linked receptor protein signaling pathway
0.001	40.881	0	2	43	blood pressure regulation
0.001	64.544	0	2	28	peptidoglycan metabolic process
0	88.372	0	2	21	multicellular organismal catabolic process
0	93.289	0	2	20	collagen catabolic process
0	93.289	0	2	20	multicellular organismal macromolecule catabolic process
0	93.289	0	2	20	multicellular organismal protein metabolic process
0	93.289	0	2	20	protein digestion
0.009	131.312	0	1	7	chitin catabolic process
0.008	157.588	0	1	6	endosome to lysosome transport
0.008	157.588	0	1	6	positive regulation of bone remodeling
0.008	157.588	0	1	6	negative regulation of G-protein coupled receptor protein signaling pathway
0.007	197	0	1	5	negative regulation of body size
0.007	197	0	1	5	brown fat cell differentiation
0.007	197	0	1	5	heat generation
0.007	197	0	1	5	imprinting
0.005	262.688	0	1	4	adaptation of signaling pathway
0.005	262.688	0	1	4	segment specification
0.005	262.688	0	1	4	negative regulation of erythrocyte differentiation
0.004	394.062	0	1	3	positive regulation of bone mineralization
0.004	394.062	0	1	3	diet induced thermogenesis
0.004	394.062	0	1	3	norepinephrine-epinephrine vasodilation during regulation of blood pressure
0.004	394.062	0	1	3	regulation of organismal metabolic process
0.004	394.062	0	1	3	transmembrane receptor protein tyrosine kinase activation (dimerization)
0.003	788.188	0	1	2	regulation of sodium ion transport
0.003	788.188	0	1	2	arrestin mediated desensitization of G-protein coupled receptor protein signaling pathway
0.003	788.188	0	1	2	negative regulation of smooth muscle contraction

Cluster 9					
Pvalue	OddsRatio	ExpCount	Count	Size	Term
0.009	131.312	0	1	17	antigen processing and presentation of peptide or polysaccharide antigen via MHC class II
0.008	161.654	0	1	14	DNA fragmentation during apoptosis
0.007	175.139	0	1	13	G2/M transition of mitotic cell cycle
0.006	210.2	0	1	11	positive regulation of JNK cascade
0.003	420.567	0	1	6	adult feeding behavior
0.003	525.75	0	1	5	regulation of response to food
0.003	525.75	0	1	5	regulation of response to extracellular stimulus
0.003	525.75	0	1	5	Golgi to plasma membrane transport
0.002	1051.667	0	1	3	positive regulation of appetite
0.002	1051.667	0	1	3	positive regulation of response to external stimulus
0.002	1051.667	0	1	3	positive regulation of response to nutrient levels
0.001	2103.5	0	1	2	Golgi to plasma membrane CFTR protein transport

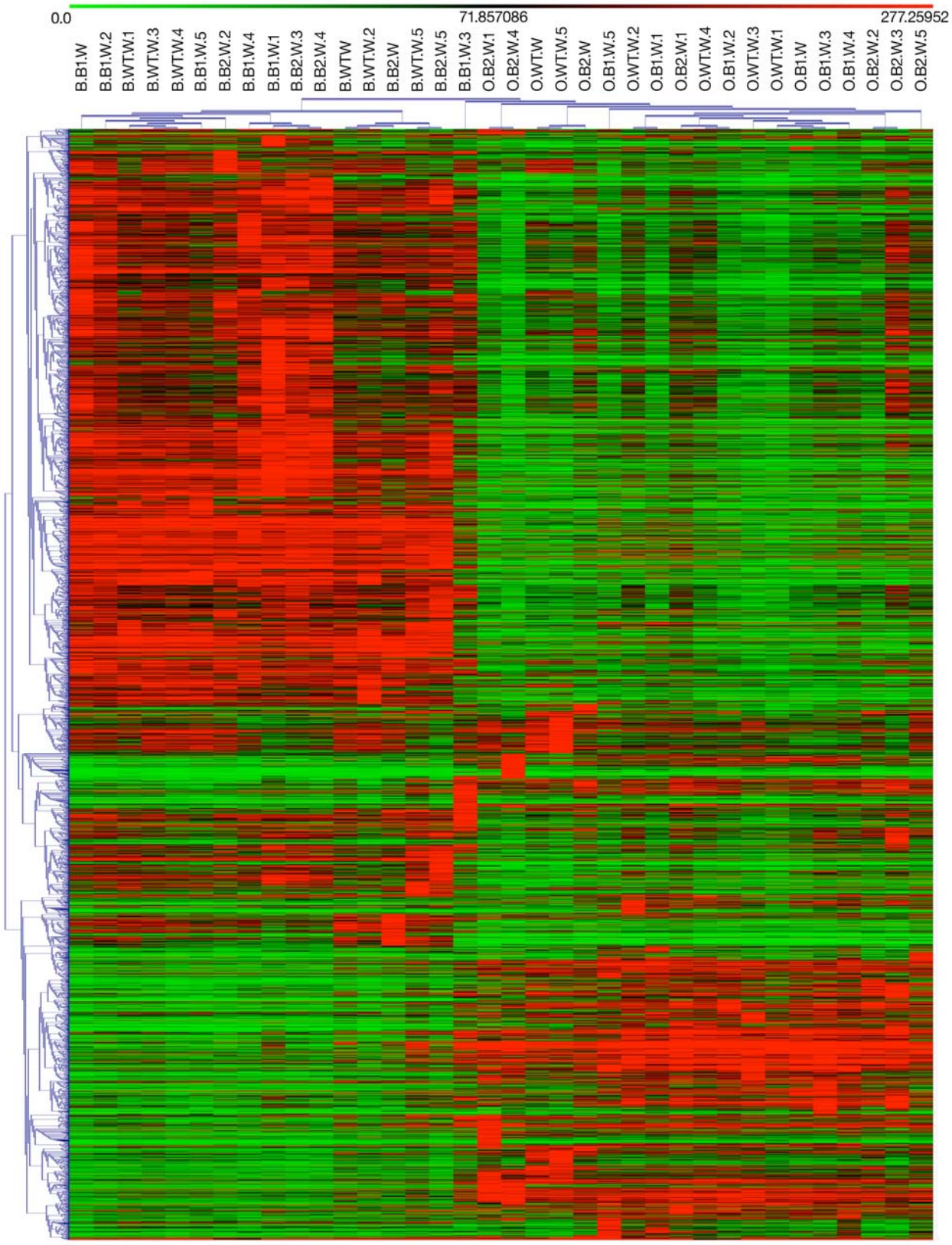
Table S3. Msig Db gene sets enriched in breast and ovarian BRCA1 and BRCA2 mutant cells

Gene Set	Description	Gene set size	PMID
ALCALAY_AML_NPMC_DN	Genes downregulated in NPMc+ leukemias	187	PMID:15831697
	Upregulated gene signature putatively involved in the maintenance of a stem-cell phenotype suggesting that NPMc+ AML may derive from a multipotent hematopoietic progenitor		
BRCA_ER_NEG	Down regulated genes whose expression is consistently negatively correlated with estrogen receptor status in breast cancer - higher expression is associated with ER-negative tumors	929	PMID:11823860
DOX_RESIST_GASTRIC_UP	Upregulated in gastric cancer cell lines reistant to doxorubicin, compared to parent chemosensitive lines	44	PMID:14734480
GAY_YY1_DN	List of YY1 target genes identified in MEFs expressing ~25% of YY1 Down	281	PMID:16611997
HOFFMANN_BIVSBII_BI_TABLE2	Genes with at least five fold change in expression between Pre-BI and Large Pre-BII cells (B-cell)	254	PMID:11779835
IDX_TSA_UP_CLUSTER3	Strongly up-regulated at 16-24 hours during differentiation of 3T3-L1 fibroblasts into adipocytes with IDX (insulin, dexamethasone and isobutylxanthine), vs. fibroblasts treated with IDX + TSA to prevent differentiation (cluster 3)	90	PMID:15033539
IRITANI_ADPROX_LYMPH	Lympocyte proliferation expression profile	127	PMID:12234922
LEE_TCELLS2_UP	Transcripts enriched in more mature cells (SP4, CB4, and AB4) more than 3-fold, with average signal value differences of at least 100 between less mature (ITTP, DP) and more mature (SP4, CB4, and AB4) cells	1141	PMID:15210650
LEE_TCELLS3_UP	Transcripts enriched in both ITTP and DP more than 3-fold, with average signal value differences of at least 100 between less mature (ITTP, DP) and more mature (SP4, CB4, and AB4) cells		PMID:15210650
POD1_KO_UP	Up-regulated in glomeruli isolated from Pod1 knockout mice, versus wild-type controls	416	PMID:16207825
SERUM_FIBROBLAST_CELLCYCL	Cell-cycle dependent genes regulated following exposure to serum in a variety of human fibroblast cell lines	138	PMID:14737219
SHEPARD_BMYB_MORPHOLINO_	Genes upregulated in control vs bmyb morpholino knockdown in zebra fish	208	PMID:16150706

STEMCELL_EMBRYONIC_UP	Enriched in mouse embryonic stem cells, compared to differentiated brain and bone marrow cells	1344	PMID:12228720
STEMCELL_NEURAL_UP	Enriched in mouse neural stem cells, compared to differentiated brain and bone marrow cells	1838	PMID:12228720
TARTE_PLASMA_BLASTIC	Genes overexpressed in mature plasma cells isolated from tonsils (TPCs) and mature plasma cells isolated from bone marrow (BMPCs) as compared to polyclonal plasmablastic cells (PPCs)	310	PMID:12663452
ZHAN_MM_CD138_PR_VS_REST	Top 50 genes from various sub-groups of multiple myeloma molecular classification	49	PMID:16728703
LEI_MYB_REGULATED_GENES	Myb-regulated genes	325	PMID:15105423

NEXT PAGE:

Figure S1. Hierarchical clustering for comparisons between *BRCA1* and *WT*, and *BRCA2* and *WT*, for breast and ovarian epithelial cultures.


0.0
71.857086
277.25952
B.B1.W
B.B1.W.2
B.WT.W.1
B.WT.W.3
B.WT.W.4
B.B1.W.5
B.B2.W.2
B.B1.W.4
B.B1.W.1
B.B2.W.3
B.B2.W.4
B.WTW
B.WT.W.2
B.B2.W
B.WT.W.5
B.B2.W.5
B.B1.W.3
O.B2.W.1
O.B2.W.4
O.WT.W
O.WT.W.5
O.B2.W
O.B1.W.5
O.WT.W.2
O.B1.W.1
O.B2.W.1
O.WT.W.4
O.B1.W.2
O.WT.W.3
O.WT.W.1
O.B1.W
O.B1.W.3
O.B1.W.4
O.B2.W.2
O.B2.W.3
O.B2.W.5

References

1. Brunet A, Bonni A, Zigmond MJ, et al. Akt promotes cell survival by phosphorylating and inhibiting a Forkhead transcription factor. Cell 1999;96:857-68.
2. Devarajan K. Nonnegative matrix factorization – A new paradigm in large-scale biological data analysis, Proceedings of the Joint Statistical Meetings, Seattle, Washington (CD-ROM). 2006.
3. Smyth GK. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. Stat Appl Genet Mol Biol 2004;3:Article3.
4. Hedenfalk I, Duggan D, Chen Y, et al. Gene-expression profiles in hereditary breast cancer. N Engl J Med 2001;344:539-48.