Integration of Promoter & Exon Arrays for colorectal carcinoma cell line under Oxaliplatin treatment indicates activation of oxidative phosphorylation pathways

RECOMB-DREAM
Columbia University
New York City
November 16-20, 2010

Yi-An Ko[1]*, Feng-Hsiang Chung[1], Qingdong Ling[1,2], and HC Lee[1,2]
[1]Institute of Systems Biology & Bioinformatics, National Central University, Zhongli, Taiwan
[2]Cathay General Hospital, Taipei, Taiwan
*Current address: Albert Einstein College of Medicine, New York City, USA

Abstract

A novel method, a genome-wide unrestrained functional elements assay (UFEA), is employed to identify changes in regulatory elements and to pursue a comprehensive characterization of pathway enrichment of HCT-116 colorectal cancer cell line under Oxaliplatin treatment. Data from high-resolution promoter tiling arrays and exon arrays are analyzed in an integrated fashion and then enriched with information from KEGG and GO for insight in pathway activity. The general trend of our analysis is that many cancer related pathways are down-regulated whereas three neurodegenerative diseases, Huntington's, Parkingson's and Alzheimers's, are conspicuous among the upregulated pathways.

Materials & Methods

Cell line treatment. HCT-116, an epithelial-like cell line that comes from human colon carcinoma, was obtained from the American Type Culture Collection as a gift from Dr. C.J. Huang. The p26 HCT-116 cells were grown in DMEM (Gibco; pH 7.4) supplemented with 10% FBS (Gibco), 100 units/mL Penicillin/Streptomycin (Gibco) cultures. For each experi-ment, cells (107/10 ml medium in replicate 100-mm dishes) were exposed to 100 μM Oxaliplatin. Nuclei were extracted and treated under seven different concentrations of DNase I (Promega), from 5, 10, 20, 40, 60, 80, to 100 U/mL. Protease K was added to a final concentration of 25μg/mL into the DNase I treated nuclei and incubated overnight at 55°C. DNA is then purified with Puregene system (Gentra Systems, Minneapolis, MN) according to the manufacturer s protocol and resuspended in 10mM Tris-Cl, pH 8.0.

Microarray data acquisition and processing. Human Promoter 1.0R Array and Human Exon 1.0 ST Array were scanned using GeneChip ® Scanner 3000 7G. We extracted signal intensities from the scanned images using the *R* package *Starr* for promoter array, and the Affymetrix software *Expression Console* (v.1.0) for exon array. Data analysis was carried in the *R* language. The *'"Starr"* and *"Siggenes"* packages were used to handle Affymetrix data, and the RMA algorithm for pre-processing, normalization and calculation of expression levels. The workflow of an UFEA-chip is summarized in Figure 1.

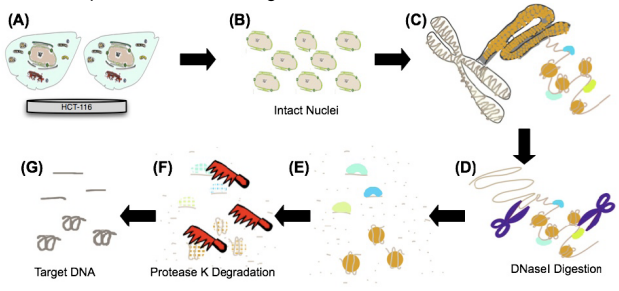

Figure 1. UFEA-chip Flowchart: Intact nuclei of HCT-116 are extracted (A-C) and digested with DNase I (D), proteins bound to DNA (E) are degraded by Protease K (F), and the remaining (target) DNA fragments (G) are purified.

Results

Enriched protein binding regions were detected through *Starr*. The normalized peak data were smoothed into bins of 250 bp. Using the quartile of the null distribution as an upper bound, a total 1709 enriched regions over the entire human genome were identified. The regions are distributed over the chromosomes more or less in proportion to the gene density in the chromosomes and, to a lesser extent, to chromosome size (Figure 2). Figure 3 shows the color coded mean intensity of each block over the 1709 enriched regions, 1681 over-enriched and 28 under-enriched (Table S1, Supporting Information (SI)). The intensities of blocks 1, 2 and 3, which lie within 500 bp of the TSS, are more than one order of magnitude greater than those of blocks 8, 9 and 10, which lie within 500 bp of the TTS. From exon array data1037 genes were identified as significant (differentially) expressed genes (SEG), including 611 up-regulated (uSEGs) and 426 down-regulated (dSEGs). The intensities of promoters in block three in Figure 3a from the promoter array and the intensities from the exon array of 16208 genes are plotted in Figure 4. The data display a slight skew towards up-regulation of the genes (mean intensity 0.03) and a strong skew towards over-enrichment (0.49). In the figure region I (II) contains 297 (267) over-enriched and up-regulated, or oP-uX (down-regulated, or oP-dX) genes. The four sets of

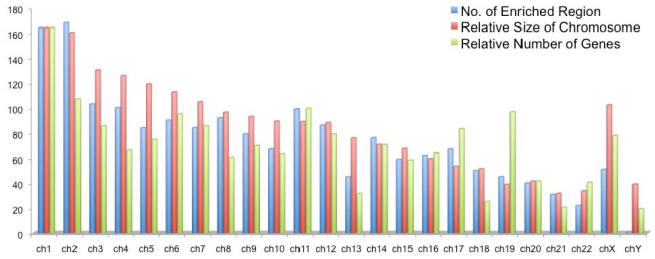

Figure 2. Distribution of 1709 enriched regions over individual chromosomes (blue) compared to relative sizes (red) of, and relative gene densities (green) in, chromosomes.

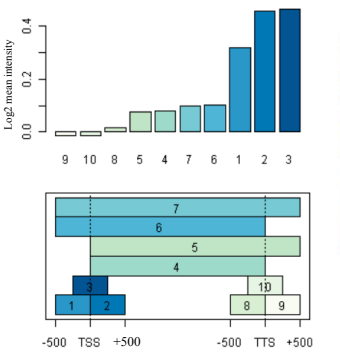

Figure 3. Lower panel: color coded mean intensities of blocks in averaged over the 1709 enriched regions. Upper panel: relative intensities of the ten blocks.

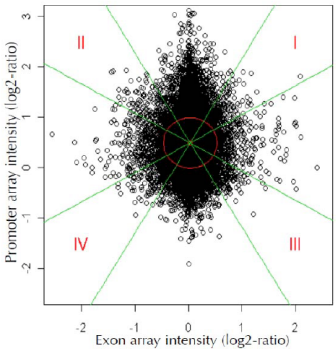

Figure 4. Promoter and exon array intensities of 16208 genes. The number of genes in the sections are: I, 297; II, 267; III, 358; IV, 336.

Table 1. KEGG pathways in which PEGs and SEGs are over-represented. uPEG and dPEG, over- and under enriched PEG, respectively; uSEG and dSEG, up- and down-regulated SEG, respectively. All pathways have P values (for over-representation) less than 0.05. Only the top 13 of uSEG and dSEG (both have 17) pathways are listed.

uSEG	KEGGID	P-value	Odds Ratio	Exp Count	Count	Size	Term
1	190	1.03E-13	8.65	4.15	25	106	Oxidative phosphorylation
2	5012	3.76E-12	7.95	4.03	23	103	Parkinson's disease
3	5016	2.39E-10	5.45	6.23	26	159	Huntington's disease
4	5010	3.55E-08	4.79	5.76	22	147	Alzheimer's disease
5	3050	2.01E-06	8.61	1.57	10	40	Proteasome
6	3040	7.81E-06	4.26	4.5	16	115	Spliceosome
7	1100	1.70E-04	1.85	39.19	60	1001	Metabolic pathways
8	4115	1.06E-03	3.96	2.62	9	67	p53 signaling pathway
9	3020	3.48E-03	5.71	1.06	5	27	RNA polymerase
10	240	8.71E-03	2.78	3.56	9	91	Pyrimidine metabolism
11	4110	2.04E-02	2.26	4.78	10	122	Cell cycle
oP-uX							
1	190	2.82E-04	5.41	1.73	8	106	Oxidative phosphorylation
2	5012	1.29E-03	4.67	1.68	7	103	Parkinson's disease
3	5016	3.98E-03	3.47	2.59	8	159	Huntington's disease
4	5010	9.42E-03	3.23	2.39	7	147	Alzheimer's disease
5	3040	3.84E-02	2.87	1.87	5	115	Spliceosome
dSEG							
1	4520	1.31E-06	6.84	2.2	12	74	Adherens junction
2	4510	1.12E-05	3.75	5.7	18	192	Focal adhesion
3	5200	4.46E-05	2.88	9.41	23	317	Pathways in cancer
4	4120	2.88E-04	3.71	3.71	12	125	Ubiquitin mediated proteolysis
5	4310	1.11E-03	3.14	4.31	12	145	Wnt signaling pathway
6	5412	1.30E-03	4.21	2.17	8	73	Arrhythmogenic right ventricular cardiomyopathy
7	4330	1.51E-03	5.5	1.28	6	43	Notch signaling pathway
8	4350	2.75E-03	3.69	2.44	8	82	TGF-beta signaling pathway
9	5222	2.75E-03	3.69	2.44	8	82	Small cell lung cancer
10	5215	4.28E-03	3.41	2.61	8	88	Prostate cancer
11	5212	4.76E-03	3.71	2.11	7	71	Pancreatic cancer
12	5223	4.91E-03	4.23	1.6	6	54	Non-small cell lung cancer
13	4810	7.05E-03	2.35	6.06	13	204	Regulation of actin cytoskeleton
oP-dX							
1	4150	5.44E-03	6.29	0.72	4	51	mTOR signaling pathway
2	4114	1.67E-02	3.64	1.51	5	107	Oocyte meiosis
3	4810	2.35E-02	2.67	2.87	7	204	Regulation of actin cytoskeleton
4	4120	3.04E-02	3.08	1.76	5	125	Ubiquitin mediated proteolysis
5	5213	3.61E-02	4.45	0.73	3	52	Endometrial cancer

genes, uSEG, dSEG, oP-uXs, and oP-dXs, were mapped to KEGG for over-representation according to biological function as in Table 1. All cancer related pathways, 10 of 13 in the dSEG set and 2 in 5 in set oP-dX, are associated with down-regulation of genes. The most prominent activated pathways, in both the uSEG and oP-uX sets, are related to mitochondria dysfunction: the metabolic oxidative phosphory-lation pathway and three of the five neurodegenerative disease pathways in KEGG database, Huntington's, Parkinson's, and Alzheimer's (HPA). This suggests neurodegenerative disease could be a significant side-effect risk in Oxaliplatin treatment.