RESEARCH ARTICLE

Data Mining for Identification of Molecular Targets in Ovarian Cancer

Vanessa Villegas-Ruiz[1,2], Sergio Juarez-Mendez[1]*

Abstract

Ovarian cancer is possibly the sixth most common malignancy worldwide, in Mexico representing the fourth leading cause of gynecological cancer death more than 70% being diagnosed at an advanced stage and the survival being very poor. Ovarian tumors are classified according to histological characteristics, epithelial ovarian cancer as the most common (~80%). We here used high-density microarrays and a systems biology approach to identify tissue-associated deregulated genes. Non-malignant ovarian tumors showed a gene expression profile associated with immune mediated inflammatory responses (28 genes), whereas malignant tumors had a gene expression profile related to cell cycle regulation (1,329 genes) and ovarian cell lines to cell cycling and metabolism (1,664 genes).

Keywords: Ovarian cancer - networks - systems biology

Asian Pac J Cancer Prev, **17 (4)**, 1691-1699

Introduction

Ovarian cancers (OC) is the sixth most common malignance in the worldwide, and represent the fourth leading cause of gynecological cancer death (Jemal et al., 2008), this is mainly more than 70% of patients are diagnosed in advanced stages, and the five year survival is less to 50% (Jemal et al., 2008). OC is classified according to the ovarian tissue of origin, the epithelial ovarian cancer (EOC) is the most common (Cannistra, 2004). EOC is further classified into serous, cell clear, mucinous and endometrioid types, with serous type being the most common. In Mexico the incidence is of 10.1 cases per 100,000 women (Globocan, 2012). Several factors are involved in prognosis of OC such as: early detection, age, tumor stage, and familiar history of ovarian/breast cancer, among others.

The identification of molecular signature has improved our understanding of the molecular mechanism associated with ovarian cancer pathogenesis has identified molecular markers useful for diagnosis, prognosis and even as target for treatment (Chen et al., 2015). Recent data indicates that certain deregulated genes are associated whit tumor progression (Liu et al., 2015).

Unregulated proliferation, migration, invasion, and treatment resistance characterize the ovarian cancer cell as well as point mutation in BRCA1/2, copy number amplification, over/under gene expression, genetics and epigenetic modification of DNA among others. The Omics studies have improved the approaches in cancer research; they provide large-scale genomics analyses of imbalances, gene expression, and proteomics profile. Our laboratory results, using high density microarrays, showed gene expression and alternative splicing profiles in non-malignant, malignant ovarian tumors and ovarian cell lines (Juarez-Mendez et al., 2013). However, it is not clear the molecular interaction of deregulates genes in OC.

Systems biology approach provides extraordinary tools to examine high complexity interaction of large gene expression data. Additionally, experimental evidence of proteins and RNA expression provided exceptionally information to search for molecular involved in prognosis, diagnosis and treatment of cancer. In this study we performed data-mining using high-density microarray and System biology using MetaCoreTM, Thomson Reuters to identify the most significant deregulated signaling pathways in non-malignant, malignant and ovarian cell lines.

Materials and Methods

Microarray gene expression

In this study we used microarray that included non-malignant ovarian tumors (NMOT, N=2), malignant ovarian tumors (MOT, N=4), ovarian cell lines (OCL, N=4) and healthy ovarian tissue (HOT, N=4) according to our previous report (Juarez-Mendez et al., 2013). Microarray data analyses were performed using Partek Genomics Suite v6.6 software (Partk Incorporated, Saint Louis, MO). In brief, microarray data was summarized using Median Polis, quantile normalization, the background noise correction was archived using RMA

*[1]Experimental Oncology Laboratory, Research Department, National Institute of Pediatrics, [2]Genetics Department-Research Unit, Institute of Ophthalmology, "Conde de Valenciana", Mexico D.F., Mexico *For correspondence: ser.mend@gmail.com*

and finally the data was log2 transformed. The microarray were compared as follows: NMOT vs HOT, MOT vs HOT and OCL vs HOT. The differential expressed genes were selected using cutoff fold change > 2 and < -2 and False Discovery Ratio (FDR) > 0.05.

Systems biology

The significant deregulated genes obtained by

Table 1. Gene Ontology Enrichment

Category	Process	pValue	FDR	Category	Ontology
NMOT	Arsenite metabolism and transport	2.53E-03	9.43E-02	NMOT	Pathway maps
NMOT	Immune response_Oncostatin M signaling via JAK-Stat in mouse cells	2.52E-02	9.43E-02	NMOT	Pathway maps
NMOT	Immune response_Oncostatin M signaling via JAK-Stat in human cells	2.98E-02	9.43E-02	NMOT	Pathway maps
NMOT	Nicotine signaling	2.93E-02	9.43E-02	NMOT	Pathway maps
NMOT	Development_Thrombopoetin signaling via JAK-STAT pathway	3.07E-02	9.43E-02	NMOT	Pathway maps
MOT	Cell cycle_Chromosome condensation in pro-metaphase	1.60E-11	5.83E-09	MOT	Pathway maps
MOT	Cell cycle_The metaphase checkpoint	1.57E-11	5.83E-09	MOT	Pathway maps
MOT	Cell cycle_Role of APC in cell cycle regulation	7.54E-09	1.86E-06	MOT	Pathway maps
MOT	Cell cycle_Spindle assemble and chromosome separation	1.14E-08	2.10E-06	MOT	Pathway maps
MOT	Cell cycle_Initiation of mitosis	6.72E-08	9.96E-06	MOT	Pathway maps
OCL	Cell cycle_The metaphase checkpoint	2.30E-19	1.83E-16	OCL	Pathway maps
OCL	Cell cycle_Role of APC in cell cycle	3.82E-18	1.52E-15	OCL	Pathway maps
OCL	Cell cycle_Chromosome condensation in pro-metaphase	1.84E-17	4.87E-15	OCL	Pathway maps
OCL	Cell cycle_Spindle assembly and chromosome separation	1.02E-13	2.02E-11	OCL	Pathway maps
OCL	Cell cycle_Initiation of mitosis	9.92E-09	1.57E-06	OCL	Pathway maps
NMOT	Development_Blood vessel morphogenesis	3.32E-04	1.29E-02	NMOT	Process Networks
NMOT	Chemotaxis	6.19E-03	1.21E-01	NMOT	Process Networks
NMOT	Reproduction_GnRH signaling pothway	1.05E-02	1.37E-01	NMOT	Process Networks
NMOT	Reproduction_Gonadotropin regulation	1.71E-02	1.58E-01	NMOT	Process Networks
NMOT	Neurophysiological process_Transmission of nerve impulse	2.02E-02	1.58E-01	NMOT	Process Networks
MOT	Cell cycle_G2-M	2.15E-15	3.43E-13	MOT	Process Networks
MOT	Cell cycle_Mitosis	2.65E-13	2.11E-11	MOT	Process Networks
MOT	Cytoskeleton_Spindle microtubules	4.70E-10	2.49E-08	MOT	Process Networks
MOT	Development_Blood vassel	1.84E-05	7.31E-04	MOT	Process Networks
MOT	Cell cycle_Core	2.34E-05	7.44E-04	MOT	Process Networks
OCL	Cell cycle_Core	4.77E-24	7.54E-22	OCL	Process Networks
OCL	Cell cycle_Mitosis	3.56E-20	2.82E-18	OCL	Process Networks
OCL	Cytoskeleton_Spindle microtubules	1.42E-17	7.5’1e-16	OCL	Process Networks
OCL	Cell cycle_G2-M	3.42E-17	1.35E-15	OCL	Process Networks
OCL	Cell cycle_S phase	3.44E-16	1.09E-14	OCL	Process Networks

means of microarray gene expression were loaded in the Metacore portal; the significant data were labeled using ID gene and fold change. The ontology were analyzed using Enrichment analysis workflow, p-values were calculated according to dataset activated ($p < 0.05$).

Results

Gene Expression

In order to identify deregulated genes associated to NMOT, MOT and OCL, we performed microarray analysis using a normal tissue (HOT) as a base line reference. The comparative microarray analysis showed differential expressed genes as follows: NMOT (N = 28), MOT (N = 1329) and OCL (N = 1664) Figure 1. Interestingly, in MOT and OCL we identified that ~60% genes were down regulated, unlike to NMOT in wich ~14% were down regulated. Our results showed an apparent progression of

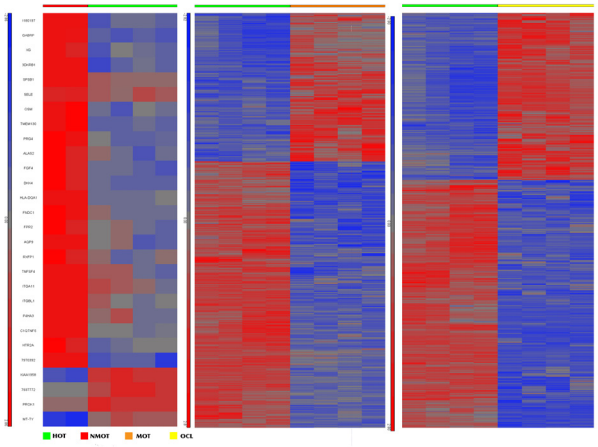

Figure 1. Hierarchical clustering. The hat map depics gene expression profile in non-malignant, malignant and ovarian cell lines. On the left side the heat map depicts gene expression profile of NMOT against HOT, in the middle MOT against HOT and on the right side OCL against HOT. The graphic was generated using Partek Genomics Suite v6.6.

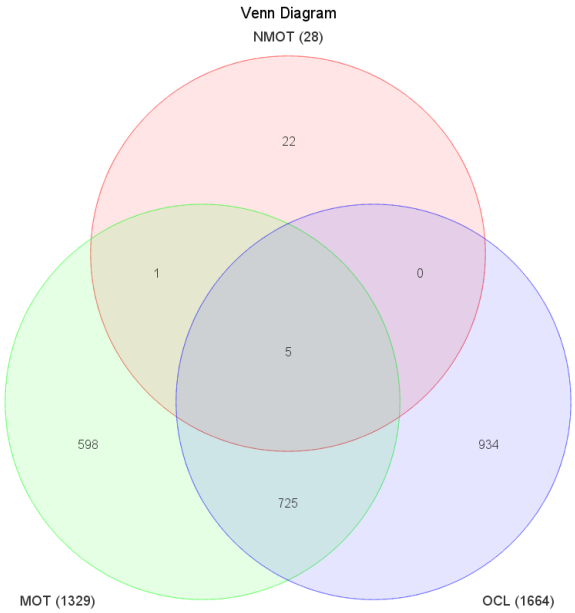

Figure 2. Correlations for expressed genes in ovarian tumors. The Venn diagram depicts the correlation of gene expression profile in NMOT, MOT and OCL against HOT. Five genes were correlated among NMOT, MOT and OCL cases, while 725 expressed genes were correlated between MOT and OCL.

Table 2. Significant network NMOT-associated

Name	GO Process	P-Value	zScore and gScore	Seed
FGF4, C1QTNF5, ITGA11, SP1, FGFR2	regulation of homotypic cell-cell adhesion	1.15E-26	62.55	8
	regulation of cell-cell adhesion	4.44E-26		
	regulation of cell activation	1.16E-23		
	regulation of cell adhesion	1.45E-22		
	positive regulation of T cell activation	5.99E-22		
EG-VEGF, ALAS2, HTR2A, SP1, MC4R	G-protein coupled receptor signaling pathway	2.63E-43	61.92	8
	G-protein coupled receptor signaling pathway, coupled to cyclic nucleotide second messenger	8.88E-34		
	positive regulation of cytosolic calcium ion concentration	1.32E-33		
	cytosolic calcium ion homeostasis	1.68E-31		
	cell surface receptor signaling pathway	3.26E-31		
BDKRB1, ATF-2, PI3K cat class IA (p110-alpha), STAT5, Cyclin D2	Fc-epsilon receptor signaling pathway	9.20E-20	31.21	4
	response to stress	1.85E-18		
	response to oxygen-containing compound	2.03E-18		
	regulation of cell death	3.39E-18		
	response to abiotic stimulus	4.60E-18		
E-selectin, Prokineticin 2, SMAD3, CXCR4, FosB	positive regulation of cellular metabolic process	1.05E-33	30.9	4
	positive regulation of macromolecule metabolic process	7.13E-33		
	response to external stimulus	6.30E-32		
	positive regulation of cellular component movement	2.64E-31		
	cellular response to organic substance	4.55E-31		
SP1, Alpha-2B adrenergic receptor, HTR4, Alpha-1D adrenergic receptor, GABA-A receptor alpha-4 subunit	G-protein coupled receptor signaling pathway	1.14E-48	24.13	3
	synaptic transmission	1.55E-47		
	cell-cell signaling	1.48E-43		
	gamma-aminobutyric acid signaling pathway	1.88E-40		
	chloride transmembrane transport	7.71E-39		

Table 3. Significant network MOT-associated

Name	GO Process	P-Value	zScore and gScore	Seed
ATP2C2, UCK2, CRISPLD2, OLFML3, KIAA0240	translational elongation	1.76E-29	40.91	25
	translation	3.59E-23		
	cellular protein metabolic process	5.34E-11		
	protein metabolic process	1.72E-09		
	gene expression	6.64E-08		
CDK1 (p34), Chk1, ATR, Claspin, Cyclin B	cell cycle checkpoint	1.57E-18	29.8	23
	DNA damage checkpoint	4.17E-18		
	DNA integrity checkpoint	1.03E-17		
	negative regulation of cell cycle	2.87E-17		
	negative regulation of mitotic cell cycle	8.29E-17		
CMKLR1, LTBP3, CSPG4 (NG2), Prokineticin 2, XYLT1	activation of adenylate cyclase activity	1.53E-18	27.14	21
	regulation of adenylate cyclase activity	2.85E-18		
	regulation of cAMP biosynthetic process	4.19E-18		
	positive regulation of cAMP biosynthetic process	5.76E-18		
	positive regulation of adenylate cyclase activity	6.66E-18		
Importin (karyopherin)-alpha, CSE1L, Fe65, micro-RNA 15a, Myopodin	establishment of localization in cell	8.23E-13	26.67	20
	protein localization to nucleus	7.41E-12		
	protein import into nucleus	1.10E-11		
	protein targeting to nucleus	1.10E-11		
	single-organism nuclear import	1.10E-11		
CDO1, TTYH2, HIST1H2BL, PI52B, APH-1	cellular response to glucagon stimulus	4.13E-60	24.48	19
	response to glucagon	4.17E-59		
	energy reserve metabolic process	5.59E-44		
	energy derivation by oxidation of organic compounds	8.81E-36		
	generation of precursor metabolites and energy	1.20E-33		

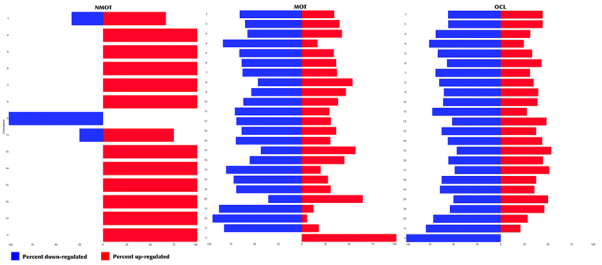

Figure 3. Percent and distribution of deregulated genes in the human genome. Gene expression profile associated to NMOT, MOT and OCL. NMOT showed more up regulated genes distributed in 14 chromosomes, while, MOT and OCL more than 50% of differential expressed genes were down regulated. The most representative difference between MOT and OCL was Y chromosome. MOT showed up regulation in contrast to OCL wich showed down regulation.

NMOT to MOT. In order to identify common deregulated genes in NMOT, MOT and OCL, we performed a Venn diagram Figure 2. The gene expression correlation was as a follows: NMOT only N= 22, MOT only N= 598, OCL only N= 937, NMOT+ MOT N= 1, NMOT+ OCL N= 0, MOT+OCL N= 725 and NMOT+MOT+OCL N= 5.

On the other hand, the up and down regulated genes were mapped by chromosomal. The deregulated genes NMOT-associated were mapped to only 14 chromosomes: up regulated (1, 4, 5, 6, 7, 8, 11, 13, 14, 15, 19, 22, and X) and down regulated (1, 9 and 11). In addition, MOT and OCL showed equal distribution of deregulated genes. Interestingly, MOT showed suppression in chromosome Y while in OCL was up-regulated Figure 3.

Enrichment of deregulated genes

The cell has a high level of complexity in molecular interaction. In order to identify the gene ontology associated to NMOT, MOT and OCL, we performed a systems biology analysis using MetaCoreTM, Thomson Reuters. The deregulated genes were loaded in MetaCore portal, after that, we performed an enrichment analysis. The expressed genes were ontology-based classified the top five are ranked in Table 1.

The enrichment analysis in NMOT showed processes associated to immune response, inflammation, vessel morphogenesis and chemotaxis among others. On the other hand, we observed in MOT and OCL several processes associated to cell cycle such as: chromosome condensation, metaphase checkpoint, mitosis initiation, spindle assembly, G2-M and S Phase among others Table 1.

Several marks give the malignant phenotype in cancer cell such as: cell proliferation, angiogenesis, and self-survival. The transcriptome analysis in ovarian

Table 4. Significant network OCL-associated

Name	GO Process	P-Value	zScore and gScore	Seed
MRPS28, MURC, FAM54A, GBGT1, POP1 (RNase P/ MRP subunit)	mitochondrial translational initiation	6.75E-14	29.04	25
	mitochondrial translational elongation	8.30E-14		
	mitochondrial translational termination	9.19E-14		
	mitochondrial translation	4.80E-13		
	translational termination	5.63E-11		
JAK2, SNRPD1 (SMD1), SNRP116, SLC25A6, DOCK10	establishment of protein localization to organelle	4.74E-14	26.94	24
	protein targeting	5.30E-13		
	intracellular protein transport	5.52E-13		
	mitochondrion organization	6.40E-12		
	protein localization to organelle	1.13E-11		
HIST2H2AC, RNMT, SLC13A3, TRIP13, Wfikkn2	antigen processing and presentation of exogenous peptide antigen	2.08E-16	26.35	23
	immunoglobulin production involved in immunoglobulin mediated immune response	3.14E-16		
	antigen processing and presentation of exogenous antigen	3.57E-16		
	antigen processing and presentation of peptide antigen	1.53E-15		
	antigen processing and presentation of exogenous peptide antigen via MHC class II	6.28E-15		
ALY, CIAPIN1, PICT-1, RFC4, PUR1	micturition	3.82E-14	25.78	23
	positive regulation of catecholamine secretion	2.62E-13		
	positive regulation of amine transport	1.23E-12		
	positive regulation of dopamine secretion	1.37E-12		
	behavioral response to nicotine	2.05E-12		
SEH1L, CAT-3, RFC3, HIST3H2A, POLR3K	protein targeting to mitochondrion	1.33E-13	25.45	22
	establishment of protein localization to mitochondrion	2.35E-13		
	protein localization to mitochondrion	4.95E-13		
	synaptic transmission	1.42E-12		
	neuropeptide signaling pathway	4.19E-10		

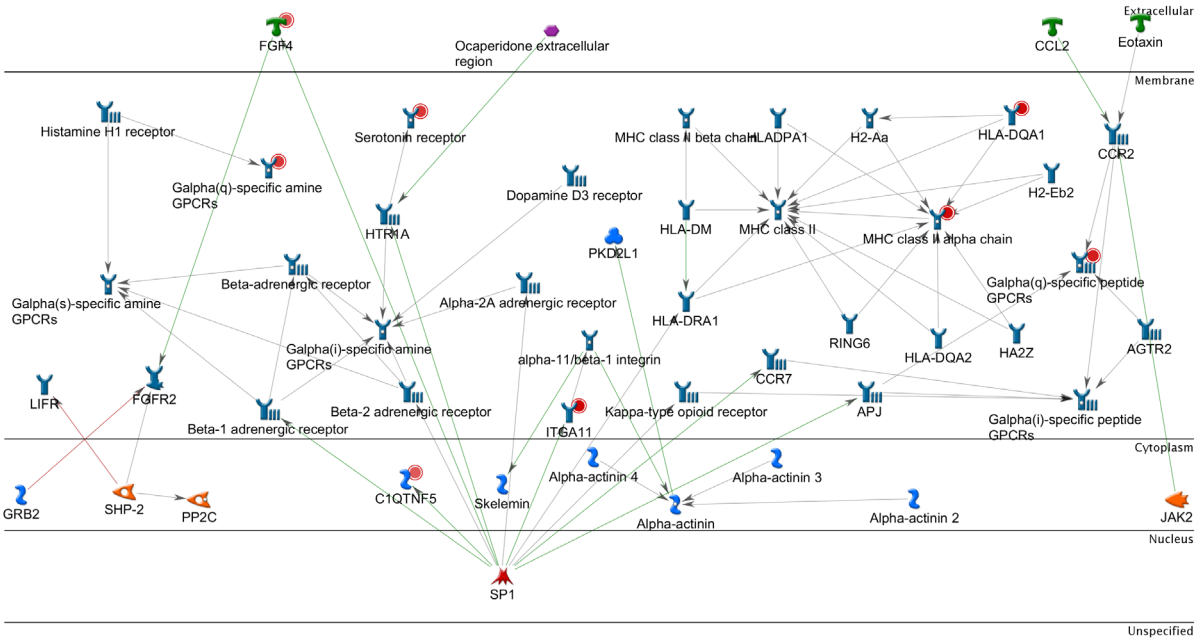

Figure 4. Significant pathway associated with NMOT. The circle red indicates up regulation. Eight up regulated genes were integrated in signaling pathways, involved in cell-cell adhesion.

cancer and ovarian cell lines showed that cell cycle is the most significant cellular process deregulated. In order to integrate signaling pathways of deregulated genes in NMOT, MOT and OCL, we built a network based on gene expression profile.

Network reconstruction

The reconstructed network was performed using a curate data by means of MetaCoreTM, Thomson Reuters system biology (SB) approach. The SB analysis reveal 15 significant networks associated to NMOT. We selected the top five based on significant and number of seed Table 2. The seed were deregulated genes observed in microarray.

According to number of seed, we used the top network, including: FGF4, C1QTNF5, ITGA11, SP1 and FGFR2. Additionally, eight genes were significant and included in regulation of cell-cell process Figure 4.

On the other hand, 30 signaling pathways were

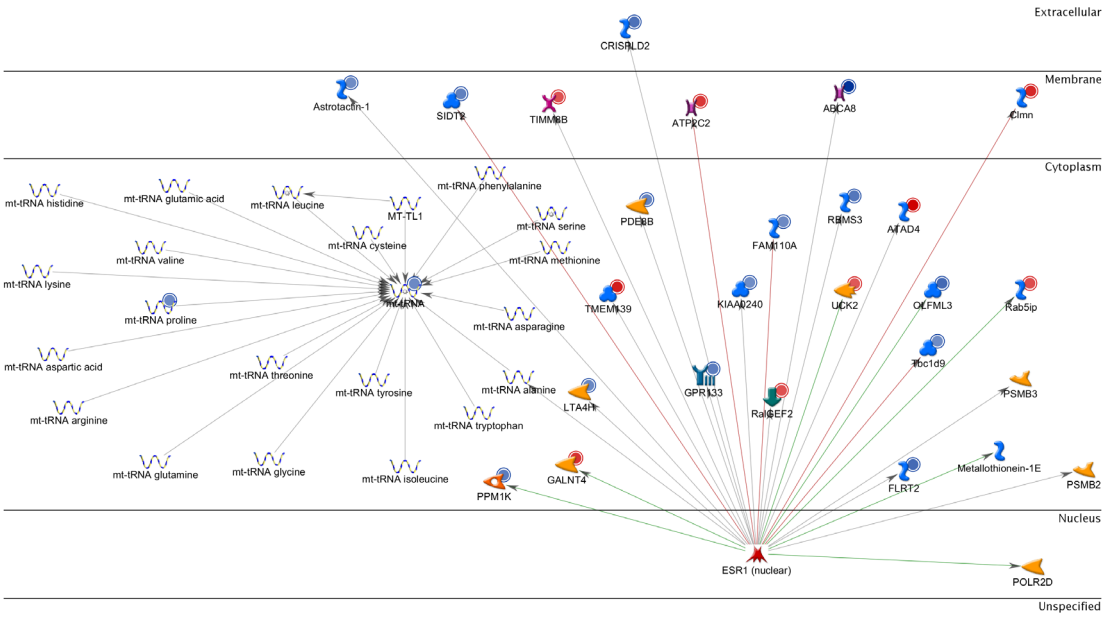

Figure 5. Estrogen receptor 1 is the most significant pathway in MOT. The circle blue and red indicate down and up regulate, respectively. Intensity colors of circle indicate level expression. 15 and nine components of ERS1 signaling pathway were down and up regulate, respectively. ERS1 expression profile is involved in cell cycle, metabolisms and gene expression regulation showed in Table 3.

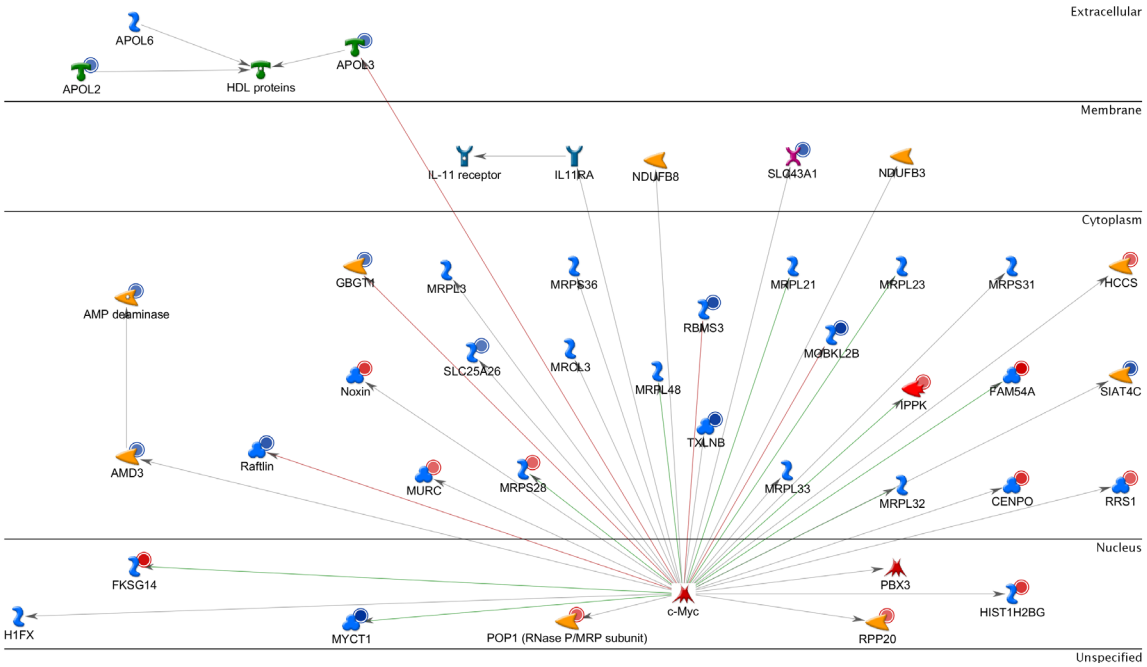

Figure 6. The significant signaling pathway c-Myc proto-oncogene was associated with OCL. 13 (blue circle) and 11 (red circle) components of c-Myc signaling pathway were down and up regulated, respectively. Deregulated genes were associated significantly with mitochondrial processes.

identified in MOT. The top five are shown in the table 3. We focused in the most significant related genes, including 25 deregulated genes such as: ATP2C2, UCK2, CRISPLD2, OLFML3 and KIAA0240 among others. The target gene in this signaling pathway is the estrogen receptor protein ESR1; 16 genes were down regulated and nine up regulated. The GO processes were associated to cell cycle process, including: translation, elongation, gene expression and cell cycle checkpoint, among others. After that, we build the network of the most significant and related genes deregulated Figure 5.

Finally, we analyzed 1664 deregulated genes OCL-associated and 30 networks were identified; the top five networks are shown in the Table 4. The most significant network contains 25 seed including: MRPS28, MURC, FAM54A, GBGT1 among others. The most significant ontology was associated to mitochondrial process including: initiation, elongation and translation. The significant network is shown in Figure 6.

Discussion

A great challenge in cancer research is the understanding of such a complex trait as well as the identification of

molecular markers that could help to predict treatment response, better classification of tumors and the identification of druggable targets. The microarray gene expression is an extraordinary tool that provides a wealth of data about differentially expressed genes. In cancer, several cellular processes are involved such as: cell cycle, proliferation, apoptosis evasion, inflammation, migration and metastasis, among others (Hanahan and Weinberg, 2000; Hanahan and Weinberg, 2011).

In Mexico, the OC is the third most common gynecological cancer (Globocan, 2012). The diagnosis is at advanced stages and the prognosis is poor. In the recent year, several comprehensive tools have been developed to understand the complex molecular interaction in human disease, including cancer. Our results using microarray gene expression revealed a tissue-associated profile. The lowest differences were observed in NMOT (N=28) Figure 1-2. Moreover, eight genes were the most significant and integrated to specific signaling pathways and they are related to cancer: FGF4 in breast (Saint-Ruf et al., 1990; Schmitt et al., 1996), colorectal (Ikeda et al., 2008), melanoma (Adelaide et al., 1988), stomach (Ikeda et al., 2008) and ovarian cancer (Schmitt et al., 1996; Mayr et al., 2006); ITGA11 in kidney (Dalgliesh et al., 2010), laryngeal (Stransky et al., 2011) and lung (Young et al., 2009); MHC class ll alpha chain in non-small cell lung carcinoma (Ohri et al., 2009); HLA-DQA1 in hepatocellular carcinoma (Donaldson et al., 2001), renal cell carcinoma (Ellerhorst et al., 2003), melanoma(D'Alessandro et al., 1987; Nagore et al., 2002; Ugurel et al., 2004); C1QTNF5 is associated to adrenocortical carcinoma (Fonseca et al., 2012), endometrial and lung neoplasms.

On the other hand, alpha(q)-specific peptide GPCRs, alpha(q)-specific amine GPCRs and serotonin receptor were associated to schizophrenia. Our results could be suggesting that non-malignant ovarian tumor, share elements with malignant ovarian tumors. However, theses molecules are not integrated in cancer signaling pathways.

The estrogen receptor protein was the most significant signaling pathway deregulated in malignant ovarian tumors. The down-regulated genes were associated to metabolism (PPM1K, LTA4H, GPR133, PDE8B, ABCA8), signaling (GPR133, FLRT2, Tbc1d9), tumor suppressor (KIAA0240), expression regulation (RBMS3) and transport (ABCA8). The potential target could be Tbc1d9: this gene is regulated by HNF3 (FoxA1 and FoxA2) or FOXM1 mediated ESR1. FoxA1, FoxA2 and FOXM1 were over expressed 2.145, 2.092 and 7.416 fold change, respectively. HNF3 has been reported in several types of cancer including: breast (Albergaria et al., 2009; Davidson et al., 2011; Davidson et al., 2012; Varadi et al., 2012) (Shah et al., 2012), non small cell lung carcinoma (Sakaeda et al., 2013), neuroblastoma (Shimizu et al., 2002), pancreatic (Song et al., 2010), prostate (Barbieri et al., 2012; Grasso et al., 2012; Imamura et al., 2012) and small cell lung carcinoma (Sakaeda et al., 2013).

The over expressed genes included in the most significant signaling pathway were GALNT4, TMEM139, RalGEF2, ATAD4, UCK2, TIMM8B, ATP2C2, they have been associated with several cancers such as: melanoma (Berger et al., 2012), breast, skin (Durinck et al., 2011; Shah et al., 2012), prostate (Grasso et al., 2012), larynx (Stransky et al., 2011), lung, (Durinck et al., 2011), pancreas, glioblastoma (Parsons et al., 2008), leukemia (Quesada et al., 2012), medulloblastoma (Robinson et al., 2012), and ovarian cancer (Jones et al., 2012).

Several models are used to investigate the molecular basis of the phenomena in cancer research; we included cancer cell lines to investigate in vitro cancer. Our results showed a differential gene expression profile, as expected (Figure 1-3). Additionally, we identified 730 genes with correlation between MOT and OCL, 599 and 934 were exclusively for MOT and OCL, respectively. These data indicate large differences between the two models of cancer we used.

In addition, OCL the most significant signaling pathway was associated with mitochondrial processes, high-level expression could lead to deregulated metabolism caused by means of in vitro culture. Thirteen genes were down regulated including: APOL2, APOL3, SLC43A1, GBGT1, RBMS3, AMPD2, SLC25A26, MOBKL2B, SIAT4C, AMD3, Faftlin, TXLNB, MYCT1. Moreover, they have been associated with several cancers such as: prostate (Johanneson et al., 2010; Barbieri et al., 2012; Grasso et al., 2012), hepatocellular carcinoma (Guichard et al., 2012), mouth (Stransky et al., 2011), kidney (Pena-Llopis et al., 2012), larynx (Stransky et al., 2011), skin (Durinck et al., 2011), lung (Ogawa et al., 1997) and medulloblastoma (Pugh et al., 2012) among others.

On the other hand, 11 transcripts were up regulated in the most significant signaling pathway of OCL, including: Noxin, IPPK, FAM54A, MURC, MRPS28, CENPO, RRS1, FKSG14, POP1, RPP20, HIST1H2BG, its expression has been altered in several types of cancer such as: melanoma (Berger et al., 2012), nervous system neoplasms (Molenaar et al., 2012), leukemia (Zhang et al., 2012), prostate (Grasso et al., 2012), stomach (Cui et al., 2011; Hong et al., 2011), skin (Durinck et al., 2011) mouth neoplasm (Stransky et al., 2011), lung (Peifer et al., 2012), endometrial (Kuhn et al., 2012), laryngeal (Stransky et al., 2011), among others.

In conclusion, the great challenges in cancer are the early detection prognosis and treatment. Using microarray gene expression and systems biology approaches we could identify the most significant signaling pathways in non-malignant, malignant and ovarian cancer cell lines. The significant genes identified in non-malignant and malignant ovarian tumors could be useful as potential markers of disease.

Acknowledgements

This work was supported in part by grant Basic Science CONACyT 243233 and Federal Founds, National Institute of Pediatrics.

References

Adelaide J, Mattei MG, Marics I, et al (1988). Chromosomal localization of the hst oncogene and its co-amplification with the int.2 oncogene in a human melanoma. *Oncogene*, **2**, 413-6.

Albergaria A, Paredes J, Sousa B, et al (2009). Expression of FOXA1 and GATA-3 in breast cancer: the prognostic significance in hormone receptor-negative tumours. *Breast Cancer Res*, **11**, R40.

Barbieri CE, Baca SC, Lawrence MS, et al (2012). Exome sequencing identifies recurrent SPOP, FOXA1 and MED12 mutations in prostate cancer. *Nat Genet*, **44**, 685-9.

Berger MF, Hodis E, Heffernan TP, et al (2012). Melanoma genome sequencing reveals frequent PREX2 mutations. *Nature*, **485**, 502-6.

Cannistra SA (2004). Cancer of the ovary. *N Engl J Med*, **351**, 2519-29.

Chen S, Gou WF, Zhao S, et al (2015). The role of the REG4 gene and its encoding product in ovarian epithelial carcinoma. *BMC Cancer*, **15**, 471.

Cui J, Chen Y, Chou WC, et al (2011). An integrated transcriptomic and computational analysis for biomarker identification in gastric cancer. *Nucleic Acids Res*, **39**, 1197-207.

D'Alessandro G, Zardawi I, Grace J, et al (1987). Immunohistological evaluation of MHC class I and II antigen expression on nevi and melanoma: relation to biology of melanoma. *Pathol*, **19**, 339-46.

Dalgliesh GL, Furge K, Greenman C, et al (2010). Systematic sequencing of renal carcinoma reveals inactivation of histone modifying genes. *Nature*, **463**, 360-3.

Davidson B, Stavnes HT, Holth A, et al (2011). Gene expression signatures differentiate ovarian/peritoneal serous carcinoma from breast carcinoma in effusions. *J Cell Mol Med*, **15**, 535-44.

Davidson B, Stavnes HT, Risberg B, et al (2012). Gene expression signatures differentiate adenocarcinoma of lung and breast origin in effusions. *Hum Pathol*, **43**, 684-94.

Donaldson PT, Ho S, Williams R, et al (2001). HLA class II alleles in Chinese patients with hepatocellular carcinoma. *Liver*, **21**, 143-8.

Durinck S, Ho C, Wang NJ, et al (2011). Temporal dissection of tumorigenesis in primary cancers. *Cancer Discov*, **1**, 137-43.

Ellerhorst JA, Hildebrand WH, Cavett JW, et al (2003). Heterozygosity or homozygosity for 2 HLA class II haplotypes predict favorable outcomes for renal cell carcinoma treated with cytokine therapy. *J Urol*, **169**, 2084-8.

Fonseca AL, Kugelberg J, Starker LF, et al (2012). Comprehensive DNA methylation analysis of benign and malignant adrenocortical tumors. *Genes Chromosomes Cancer*, **51**, 949-60.

Globocan (2012). http://globocan.iarc.fr/Pages/fact_sheets_population.aspx.

Grasso CS, Wu YM, Robinson DR, et al (2012). The mutational landscape of lethal castration-resistant prostate cancer. *Nature*, **487**, 239-43.

Guichard C, Amaddeo G, Imbeaud S, et al (2012). Integrated analysis of somatic mutations and focal copy-number changes identifies key genes and pathways in hepatocellular carcinoma. *Nat Genet*, **44**, 694-8.

Hanahan D, Weinberg RA (2000). The hallmarks of cancer. *Cell*, **100**, 57-70.

Hanahan D, Weinberg RA (2011). Hallmarks of cancer: the next generation. *Cell*, **144**, 646-74.

Hong CS, Cui J, Ni Z, et al (2011). A computational method for prediction of excretory proteins and application to identification of gastric cancer markers in urine. *PLoS One*, **6**, 16875.

Ikeda S, Sasazuki S, Natsukawa S, et al (2008). Screening of 214 single nucleotide polymorphisms in 44 candidate cancer susceptibility genes: a case-control study on gastric and colorectal cancers in the Japanese population. *Am J Gastroenterol*, **103**, 1476-87.

Imamura Y, Sakamoto S, Endo T, et al (2012). FOXA1 promotes tumor progression in prostate cancer via the insulin-like growth factor binding protein 3 pathway. *PLoS One*, **7**, 42456.

Jemal A, Siegel R, Ward E, et al (2008). Cancer statistics, 2008. *CA Cancer J Clin*, **58**, 71-96.

Johanneson B, McDonnell SK, Karyadi DM, et al (2010). Family-based association analysis of 42 hereditary prostate cancer families identifies the Apolipoprotein L3 region on chromosome 22q12 as a risk locus. *Hum Mol Genet*, **19**, 3852-62.

Jones S, Wang TL, Kurman RJ, et al (2012). Low-grade serous carcinomas of the ovary contain very few point mutations. *J Pathol*, **226**, 413-20.

Juarez-Mendez S, Zentella-Dehesa A, Villegas-Ruiz V, et al (2013). Splice variants of zinc finger protein 695 mRNA associated to ovarian cancer. *J Ovarian Res*, **6**, 61.

Kuhn E, Wu RC, Guan B, et al (2012). Identification of molecular pathway aberrations in uterine serous carcinoma by genome-wide analyses. *J Natl Cancer Inst*, **104**, 1503-13.

Liu X, Gao Y, Zhao B, et al (2015). Discovery of microarray-identified genes associated with ovarian cancer progression. *Int J Oncol*, **46**, 2467-78.

Mayr D, Kanitz V, Anderegg B, et al (2006). Analysis of gene amplification and prognostic markers in ovarian cancer using comparative genomic hybridization for microarrays and immunohistochemical analysis for tissue microarrays. *Am J Clin Pathol*, **126**, 101-9.

Molenaar JJ, Koster J, Zwijnenburg DA, et al (2012). Sequencing of neuroblastoma identifies chromothripsis and defects in neuritogenesis genes. *Nature*, **483**, 589-93.

Nagore E, Planelles MD, Ledesma E, et al (2002). Molecular genetic analysis of HLA-DR and -DQ alleles in Spanish patients with melanoma. *Acta Derm Venereol*, **82**, 90-3.

Ogawa JI, Inoue H, Koide S (1997). alpha-2,3-Sialyltransferase type 3N and alpha-1,3-fucosyltransferase type VII are related to sialyl Lewis(x) synthesis and patient survival from lung carcinoma. *Cancer*, **79**, 1678-85.

Ohri CM, Shikotra A, Green RH, et al (2009). Macrophages within NSCLC tumour islets are predominantly of a cytotoxic M1 phenotype associated with extended survival. *Eur Respir J*, **33**, 118-26.

Parsons DW, Jones S, Zhang X, et al (2008). An integrated genomic analysis of human glioblastoma multiforme. *Science*, **321**, 1807-12.

Peifer M, Fernandez-Cuesta L, Sos ML, et al (2012). Integrative genome analyses identify key somatic driver mutations of small-cell lung cancer. *Nat Genet*, **44**, 1104-10.

Pena-Llopis S, Vega-Rubin-de-Celis S, Liao A, et al (2012). BAP1 loss defines a new class of renal cell carcinoma. *Nat Genet*, **44**, 751-9.

Pugh TJ, Weeraratne SD, Archer TC, et al (2012). Medulloblastoma exome sequencing uncovers subtype-specific somatic mutations. *Nature*, **488**, 106-10.

Quesada V, Conde L, Villamor N, et al (2012). Exome sequencing identifies recurrent mutations of the splicing factor SF3B1 gene in chronic lymphocytic leukemia. *Nat Genet*, **44**, 47-52.

Robinson G, Parker M, Kranenburg TA, et al (2012). Novel mutations target distinct subgroups of medulloblastoma. *Nature*, **488**, 43-8.

Saint-Ruf C, Gerbault-Seureau M, Viegas-Pequignot E, et al (1990). Proto-oncogene amplification and homogeneously staining regions in human breast carcinomas. *Genes Chromosomes Cancer*, **2**, 18-26.

Sakaeda M, Sato H, Ishii J, et al (2013). Neural lineage-specific homeoprotein BRN2 is directly involved in TTF1 expression

in small-cell lung cancer. *Lab Invest*, **93**, 408-21.
Schmitt JF, Susil BJ, Hearn MT (1996). Aberrant FGF-2, FGF-3, FGF-4 and C-erb-B2 gene copy number in human ovarian, breast and endometrial tumours. *Growth Factors*, **13**, 19-35.
Shah SP, Roth A, Goya R, et al (2012). The clonal and mutational evolution spectrum of primary triple-negative breast cancers. *Nature*, **486**, 395-9.
Shimizu S, Kondo M, Miyamoto Y, et al (2002). Foxa (HNF3) up-regulates vitronectin expression during retinoic acid-induced differentiation in mouse neuroblastoma Neuro2a cells. *Cell Struct Funct*, **27**, 181-8.
Song Y, Washington MK, Crawford HC (2010). Loss of FOXA1/2 is essential for the epithelial-to-mesenchymal transition in pancreatic cancer. *Cancer Res*, **70**, 2115-25.
Stransky N, Egloff AM, Tward AD, et al (2011). The mutational landscape of head and neck squamous cell carcinoma. *Science*, **333**, 1157-60.
Ugurel S, Uhlig D, Pfohler C, et al (2004). Down-regulation of HLA class II and costimulatory CD86/B7-2 on circulating monocytes from melanoma patients. *Cancer Immunol Immunother*, **53**, 551-9.
Varadi V, Bevier M, Grzybowska E, et al (2012). Genetic variation in ALCAM and other chromosomal instability genes in breast cancer survival. *Breast Cancer Res Treat*, **131**, 311-9.
Young RP, Hopkins RJ, Hay BA, et al (2009). Lung cancer susceptibility model based on age, family history and genetic variants. *PLoS One*, **4**, e5302.
Zhang J, Ding L, Holmfeldt L, et al (2012). The genetic basis of early T-cell precursor acute lymphoblastic leukaemia. *Nature*, **481**, 157-63.