13

Functional Analysis of the Cervical Carcinoma Transcriptome: Networks and New Genes Associated to Cancer

Mauricio Salcedo et al.*
Laboratorio de Oncología Genómica, Unidad de Investigación en Enfermedades Oncológicas, Hospital de Oncología, CMN-SXXI, IMSS, México DF

1. Introduction

Cancer is one of the most important public health problem in Mexico and worldwide, especially for female population, breast and cervical cancer (CC) types are the most frequent. Incidence rates of CC are higher in developing countries 40/100,000 women per year vs. 10/100,000 in developed countries (1). In Mexico there are 12,000 new reports cases every year (2). The absence of the screening programs or comparatively ineffective screening programs lead to relatively late diagnosis of the disease and also in differences in the human papillomavirus (HPV) infection (3). Several types of HPV are associated with CC worldwide (4, 5), being the HPV16 the most frequent oncogenic type.

Epidemiological and experimental studies suggest that high risk HPV have an important role in cervical carcinogenesis. Persistent viral infection, genetic background in combination with constitutive expression of the viral oncogenes as E6 and E7, are decisive steps for malignant transformation, because these oncoproteins interact with the tumour suppressor proteins p53 and pRB, respectively for their degradation (6, 7). Finally, these interactions could induce cellular proliferation and genetic instability for example, which could promote the accumulation of mutations and aneuploidy (8). In conclusion, viral oncoproteins have a general impact in global profile of expressed genes, which could be analyzed by high-throughput methodologies. One of these techniques is DNA oligonucleotide-based microarray technology, which allows a rapid and high-throughput detection of thousands of transcripts simultaneously (9-11).

It has been published several studies about gene expression profiles in HPV infected cells. Mainly these reports are based on gene expression levels altered by E6 and E7 HPV oncoproteins (12-17). Regarding changes in gene expression profiles in cervical cancer

* Sergio Juarez-Mendez[1], Vanessa Villegas-Ruiz[1], Hugo Arreola[1], Oscar Perez[2], Guillermo Gómez[3], Edgar Roman-Bassaure[4], Pablo Romero[1], Raúl Peralta[1]
1 Laboratorio de Oncología Genómica, Unidad de Investigación en Enfermedades Oncológicas, Hospital de Oncología, CMN-SXXI, IMSS, México DF
2 Laboratorio de Oncología Experimental, Instituto Nacional de Pediatría, SS, México
3 Centro Nacional de Clínica de Displasias, Unidad de Oncología, Hospital General de México, SS.
4 Servicio de Oncología, Hospital General de México, SS.

samples, there are a few papers comparing normal cervical expressed genes versus tumors samples (12, 18), the major aim in those studies was to find potential tumor markers with clinical value. At present the list of the potential markers is short (p16, survivin).
We have published some works about alterations in gene expression in CC (19, 20). In those reports we observed that WNT pathway, calcium pathway and some cellular proteases (MMP11, cathepsin F) could be involved in cervical carcinogenesis. Thus, these findings are contributing to our knowledge about alterations in CC pathogenesis.
In the present work our microarray data obtained from microarray assay on CC samples (20) were newly managed and analyzed by using new bioinformatics suite programs and then, to get others genes altered in cervical cancer, as well as to define signaling pathways probably implicated in this type of cancer.

2. Material and methods

Biological samples. Eight squamous CC tissues stage IIB (according of International Federation of Obstetrics and Gynecology, FIGO) HPV16 positive were selected and two healthy "normal" cervical samples were originally studied. The normal samples were collected after hysterectomy by uterine myomatosis without HPV infection history. All DNAs from healthy and cervical cancer samples were subjected to PCR by using general oligonucleotides against to HPV; and to confirm the HPV type, the positive samples were then sequenced (data not shown). An important point to eliminate negative false, only the samples harboured at least 70% of tumour cells or normal epithelial cells were analyzed.
Total RNA was extracted from squamous CC and normal tissues using TRIzol reagent (GIBCO, BRL, USA). RNA was synthesized and labeled with CodeLink Express Assay Reagent Kit (Applied Microarrays, GE Healthcare, USA).

2.1 Microarray platform

CodeLink™ Human Whole Genome Microarrays offer comprehensive coverage of the Human genome, this array have ~57,000 probes and consider transcripts and ESTs (expression tagged sequences). In this system are included: 1,200 genes for oncogenesis process, 1,400 for cell cycle, 1,000 for cell-signaling, 3,000 for metabolism, 1,400 for developmental process, 2,700 of transcription and translation, 1,100 for immune and inflammation response, 800 for protein phosphorylation, 600 for apoptosis, 1,150 for ion transport, 400 for synaptic transmission, 200 for kinases, among other.
This array harbors 45,674 genes (based on unique UniGene IDs), 360 positive controls, 384 negative controls, 100 housekeeping genes, one specific and functionally validated probe and the oligonucleotide probe length of 30-mer.
This platform has been applied in different biological models (21-23). CodeLink Bioarray are recently introduced, single-color oligonucleotide microarrays, which differ from Affymetrix GeneChips in the following aspects: 1) this Bioarray use a single pre-synthesized, pre-validated 30-mer probe to detect each target transcript, whereas Gene-Chips use multiple in-situ synthesized, 25-mer probes; and 2) the surface of CodeLink Bioarrays is made of 3-dimensional aqueous gel matrix, whereas that of Affymetrix GeneChips is made of 2-dimensional glass matrix. These characteristics could suggest that CodeLink Bioarrays behave differently from GeneChips and may require different normalization strategies from the ones optimized for GeneChips (24)(Figure 1).

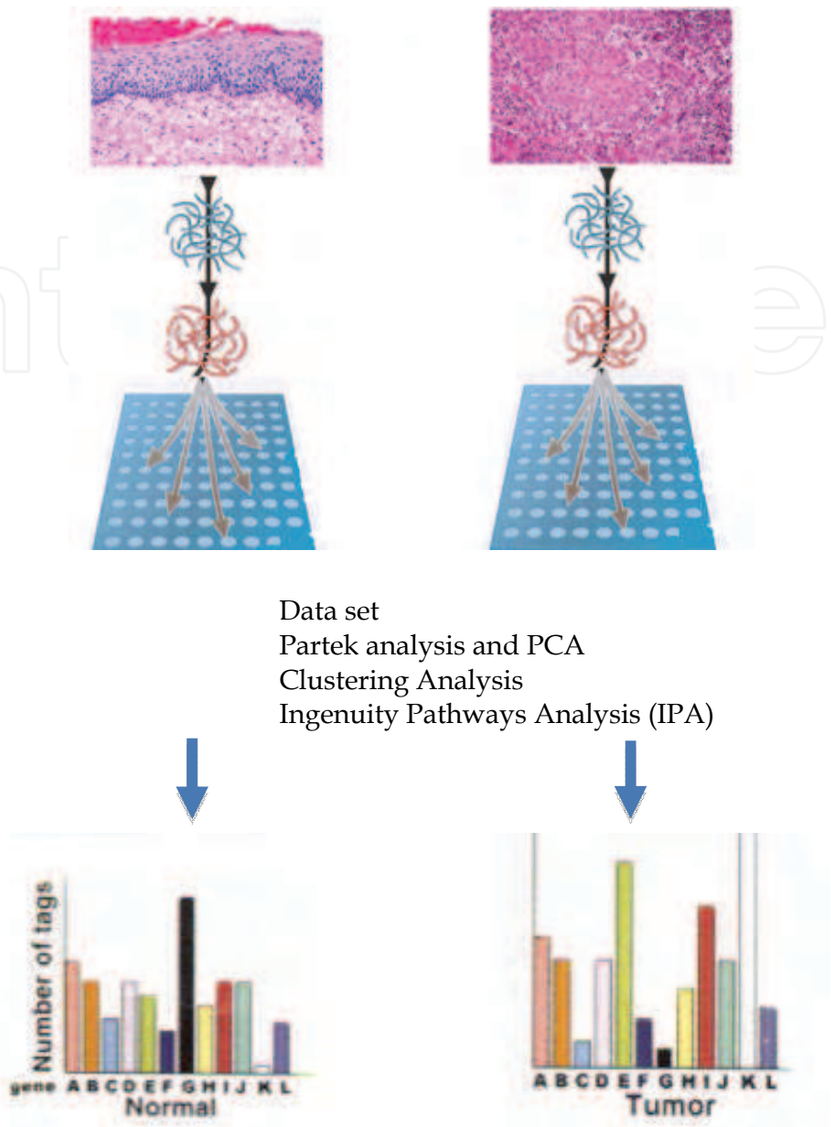


Fig. 1. General strategy of microarray analysis on cervical cancer samples. Originally RNA samples from normal epithelial cells or CC lesions were subjected to hybridization reaction on arrays. The data were analyzed by using different bioinformatics Tools to obtain differentially expressed genes.

3. Analysis of the data

Partek® Genomics Suite™ is a comprehensive suite of advanced statistics and interactive data visualization specifically designed to reliably extract biological signals from noisy data. The commercial software is unique in supporting all microarray and next generation sequencing technologies including gene expression and digital gene expression, exon/alternative splicing, RNA-Seq, copy number and association, ChIP-chip, ChIP-seq, and microRNAs in a single software package, allowing for analysis of multiple applications in one complete solution.

This kind of analysis will provide results with minimal noises generated from the internal controls. To perform this analysis is necessary to apply a software suite which is composed by three different statistical tests: Probe Statistical algorithm (MAS5), Probe Logarithmic Intensity Error (Plier) and Robust Multichip Analysis (RMA). The goal of these tests is to establish differences and similarities between internal controls and to get the most real data. In the present case, the normalized samples were analyzed RMA statistical tests eliminating the tags harboring variations.

In the global gene expression is difficult understand what happen with all of genes in different cellular process including the cancer, in this context a way of visualization data is a Principal Component Analysis or PCA (25). This method is a mathematical technique to reduction the effect of the gene expression sample in a small dimensional space, when there is less changes in the global gene expression the dimensional is smaller. Next for visualization of changes in gene expression in all samples we made a clustering analysis (26), in this method we used a K-means algorithm and let to classify to determine similitude and dissimilitude in all samples, to finish we applied methods of systems biology as Ingenuity Pathway Analysis) and gene classification to determine new list of candidates for subsequent lab verification and might help in the search for a cure for cancers.

3.1 Networks by Ingenuity Pathways Analysis (IPA)

This software (www.ingenuity.com) to help life science researchers explore, interpret, and analyze complex biological Systems, and is used to help researchers analyze 'omics data and model biological systems. This analysis was to identify Networks of interacting genes and other functional groups. A cut-off ratio of 2 was used to define genes.

4. Results

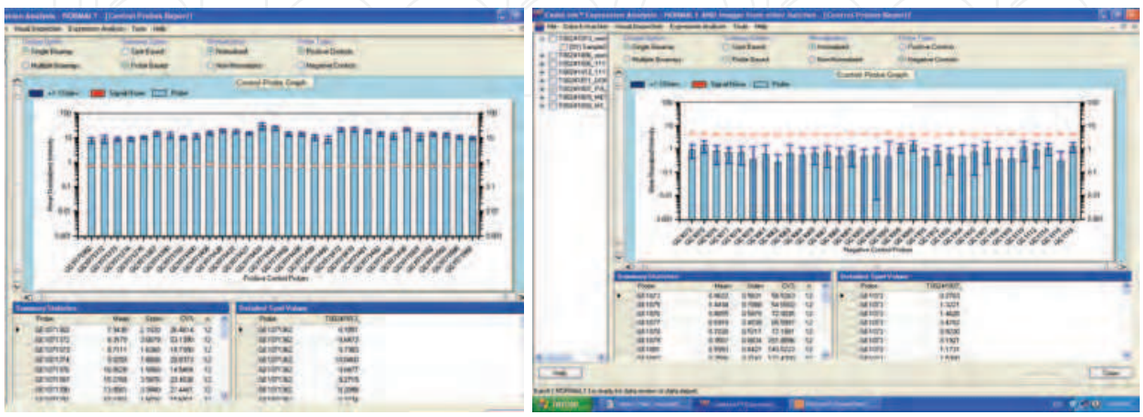
The best and most accurate method for identifying disease-causing genes is monitoring gene expression values in different samples using microarray technology. One of the shortcomings of microarray data is that they provide a small quantity of samples with respect to the number of genes. This problem reduces the classification accuracy of the methods, so gene selection is essential to improve the predictive accuracy and to identify potential marker genes for a disease. Among numerous existing methods for gene selection, PARTEK has become one of the leading methods, but its performance can be reduced because of the small sample size, noisy data and the fact that the methods remove redundant genes.

The original cervical dataset was already published and available in a previous report (20). This dataset was obtained by using CodeLink microarray platform and provides the expression levels of 57,000 probes for 2 normal tissues and 8 cervical cancers HPV16 positive. The data were pre-processed by carrying out a base 10 logarithmic transformation and normalized (see Figure 2). After first analysis of the data, using significance analysis of microarrays 3,248 genes well annotated were identified.

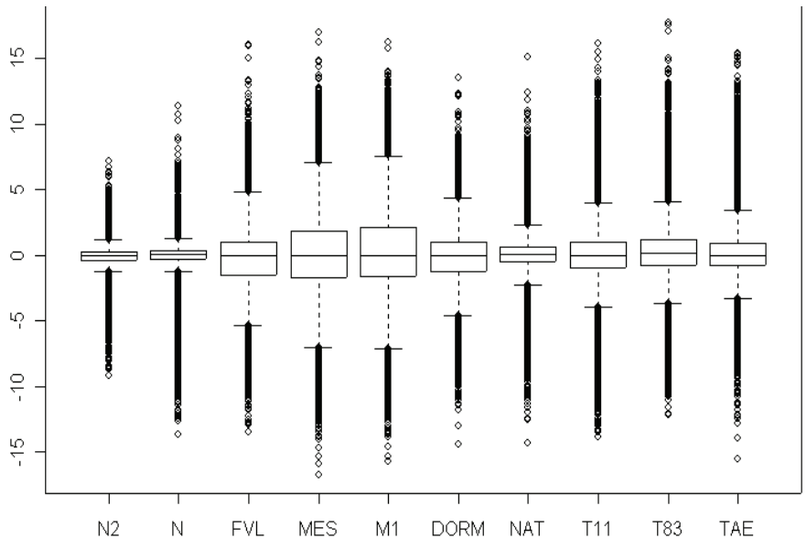
After that, the data already normalized were managed and analyzed using Partek genomics suite version 6.5 (Partek GS) obtaining a list of differentially expressed genes (Tumor versus Normal). Specific and new procedures are actually performed in analysis of microarray data. For instance, after internal control analysis, the raw data are normalized using the common software available on line or some other like Partek GS. This suite provides

rigorous and easy-to-use statistical tests for differential expression of genes or exons, and a flexible and powerful statistical test to detect alternative splicing based on a powerful mixed model analysis of variance.

By one-side ANOVA statistical test a small group of 208 genes were selected with a false discovery ratio < 10%, from these, 111 overexpressed genes with a fold change >2, and 97 downregulated genes with a fold change of <-2 genes were observed (Tables 1 and 2).



A)



B)

Fig. 2. Normalization data set. The values of gene expression were normalyzed by medians. A) Left image is showing the normalization of positive internal controls harboured in the bioarray. The right image is showing the normalization of negative internal controls of the array. As example, these figures corresponds to a normal samples.

Id	Gene Name	Fold-Change
RASGRP3	RAS guanyl releasing protein 3 (calcium and DAG-regulated)	495.214
BG193781	-	263.988
AI738482	-	189.157
KNG1	kininogen 1	101.811
BM978180	-	98.8303
BG571599	-	86.5592
AHCYL1	adenosylhomocysteinase-like 1	85.9297
LOC196993	-	84.3881
BE869762	-	83.4775
BE217873	-	74.0974
PDC	phosducin	62.2193
BU180997	-	58.1258
AI911397	-	57.9307
T71115	-	51.9557
BI460109	-	51.0251
AI807757	-	50.3766
AA533894	-	49.0591
AI149692	-	48.6084
IL26	interleukin 26	48.4676
C6orf167	MMS22-like, DNA repair protein	41.5497
ZC3HC1	zinc finger, C3HC-type containing 1	40.224
KIAA0367	prune homolog 2 (Drosophila)	38.7778
BU618645	-	38.5742
AI681317	-	37.1748
C10orf6	family with sequence similarity 178, member A	36.6882
AI699935	-	36.6513
BX100783	-	36.5138
GABRB2	gamma-aminobutyric acid (GABA) A receptor, beta 2	36.3921
AA179557	-	35.8407
CA773815	-	34.9108
CFTR	cystic fibrosis transmembrane conductance regulator (ATP-binding cassette sub-family C, member 7)	33.2819
HIST1H4B	histone cluster 1, H4b	31.9941
AI809950	-	31.7055
MAN1C1	mannosidase, alpha, class 1C, member 1	31.6513
ROR2	receptor tyrosine kinase-like orphan receptor 2	27.7424
BF432424	-	27.2275
AA676547	-	26.9584
BU852916	-	26.6901
AI028168	-	26.5172
BX111626	-	26.4598
BX103473	-	26.3868
PEG10	paternally expressed 10	26.2898
BF513637	-	26.279
BM702800	-	25.9976
AI802092	-	25.4967
BF130283	-	24.1707
CNGA3	cyclic nucleotide gated channel alpha 3	24.1082
LOC377064	-	24.0827
KIAA1045	KIAA1045	24.0353
ST7OT2	ST7 overlapping transcript 2 (non-protein coding)	23.2047
KLHDC1	kelch domain containing 1	23.0623
FLJ00012	ATG16 autophagy related 16-like 2 (S. cerevisiae)	22.9406
AI382167	-	22.7438
FLJ27365	-	22.5175
AA203713	-	20.951
INS	insulin	20.9348
MST1	macrophage stimulating 1 (hepatocyte growth factor-like)	20.9245
AI201147	-	20.8481
DEPDC2	phosphatidylinositol-3,4,5-trisphosphate-dependent Rac exchange factor 2	19.1111
BX104034	-	19.0765
AV652440	-	18.7252
SOAT	sterol O-acyltransferase 1	18.6437
LOC113386	-	18.306
CGI-69	solute carrier family 25, member 39	17.8068
FLJ32312	pseudouridylate synthase 10	17.7345
SMCY	lysine (K)-specific demethylase 5D	17.6437
MYOCD	myocardin	17.6379
BQ019477	-	17.2211
BCMO1	beta-carotene 15,15'-monooxygenase 1	17.1736
GABPB2	GA binding protein transcription factor, beta subunit 2	17.1521
TRY1	protease, serine, 1 (trypsin 1)	17.1496
AW194927	-	17.0466
FOLR3	folate receptor 3 (gamma)	16.8971
ZNF608	zinc finger protein 608	16.426
BM996042	-	16.35
KCNH3	potassium voltage-gated channel, subfamily H (eag-related), member 3	16.084
LOC389072	-	15.7376
LRRC7	leucine rich repeat containing 7	14.923
R39119	-	14.9203
AA227844	-	14.1789
SYMPK	symplekin	13.4518
BX119469	-	13.281
FLJ13941	MORN repeat containing 1	13.0872
AA977184	-	13.0121
AKAP5	A kinase (PRKA) anchor protein 5	12.861
BM993877	-	12.7869
AW140064	-	12.5985
AI809909	-	12.3174
MYOHD1	myosin XIX	11.958
AI821300	-	11.9086
AK123295	-	11.507
C14orf46	lin-52 homolog (C. elegans)	11.1136
AW265107	-	10.8928
BC041899	-	10.7334
AA609479	-	10.5942
AK055332	-	10.071
PARK2	parkinson protein 2, E3 ubiquitin protein ligase (parkin)	9.50627
MAPK1	mitogen-activated protein kinase 1	8.59309
AI792305	-	8.47742
CAMK2G	calcium/calmodulin-dependent protein kinase II gamma	8.30591
SNTB1	syntrophin, beta 1 (dystrophin-associated protein A1, 59kDa, basic component 1)	8.03086
ABLIM1	actin binding LIM protein 1	7.67597
AI139817	-	6.63686
PIP3-E	interaction protein for cytohesin exchange factors 1	6.27178
C22orf1	metallophosphoesterase domain containing 1	6.25449
ARHGAP22	Rho GTPase activating protein 22	5.97438
BC037923	-	5.32521
BI850933	-	5.24072
ADAMTS9	ADAM metallopeptidase with thrombospondin type 1 motif, 9	4.85455
GAS2L2	growth arrest-specific 2 like 2	4.74955
GRIA1	glutamate receptor, ionotropic, AMPA 1	2.5043

Table 1. List of genes and EST's (without gene name) up-regulated

Id	Gene Name	Fold-Change
RASGRP3	RAS guanyl releasing protein 3 (calcium and DAG-regulated)	495.214
KNG1	kininogen 1	101.811
AHCYL1	adenosylhomocysteinase-like 1	85.9297
PDC	phosducin	62.2193
IL26	interleukin 26	48.4676
C6orf167	MMS22-like, DNA repair protein	41.5497
ZC3HC1	zinc finger, C3HC-type containing 1	40.224
KIAA0367	prune homolog 2 (Drosophila)	38.7778
C10orf6	family with sequence similarity 178, member A	36.6882
GABRB2	gamma-aminobutyric acid (GABA) A receptor, beta 2	36.3921
CFTR	cystic fibrosis transmembrane conductance regulator (ATP-binding cassette sub-family C, member 7)	33.2819
HIST1H4B	histone cluster 1, H4b	31.9941
MAN1C1	mannosidase, alpha, class 1C, member 1	31.6513
ROR2	receptor tyrosine kinase-like orphan receptor 2	27.7424
PEG10	paternally expressed 10	26.2898
CNGA3	cyclic nucleotide gated channel alpha 3	24.1082
KIAA1045	KIAA1045	24.0353
ST7OT2	ST7 overlapping transcript 2 (non-protein coding)	23.2047
KLHDC1	kelch domain containing 1	23.0623
FLJ00012	ATG16 autophagy related 16-like 2 (S. cerevisiae)	22.9406
INS	insulin	20.9348
MST1	macrophage stimulating 1 (hepatocyte growth factor-like)	20.9245
DEPDC2	phosphatidylinositol-3,4,5-trisphosphate-dependent Rac exchange factor 2	19.1111
SOAT	sterol O-acyltransferase 1	18.6437
CGI-69	solute carrier family 25, member 39	17.8068
FLJ32312	pseudouridylate synthase 10	17.7345
SMCY	lysine (K)-specific demethylase 5D	17.6437
MYOCD	myocardin	17.6379
BCMO1	beta-carotene 15,15'-monooxygenase 1	17.1736
GABPB2	GA binding protein transcription factor, beta subunit 2	17.1521
TRY1	protease, serine, 1 (trypsin 1)	17.1496
FOLR3	folate receptor 3 (gamma)	16.8971
ZNF608	zinc finger protein 608	16.426
KCNH3	potassium voltage-gated channel, subfamily H (eag-related), member 3	16.084
LRRC7	leucine rich repeat containing 7	14.923
SYMPK	symplekin	13.4518
FLJ13941	MORN repeat containing 1	13.0872
AKAP5	A kinase (PRKA) anchor protein 5	12.861
MYOHD1	myosin XIX	11.958
C14orf46	lin-52 homolog (C. elegans)	11.1136
PARK2	parkinson protein 2, E3 ubiquitin protein ligase (parkin)	9.50627
MAPK1	mitogen-activated protein kinase 1	8.59309
CAMK2G	calcium/calmodulin-dependent protein kinase II gamma	8.30591
SNTB1	syntrophin, beta 1 (dystrophin-associated protein A1, 59kDa, basic component 1)	8.03086
ABLIM1	actin binding LIM protein 1	7.67597
PIP3-E	interaction protein for cytohesin exchange factors 1	6.27178
C22orf1	metallophosphoesterase domain containing 1	6.25449
ARHGAP22	Rho GTPase activating protein 22	5.97438
ADAMTS9	ADAM metallopeptidase with thrombospondin type 1 motif, 9	4.85455
GAS2L2	growth arrest-specific 2 like 2	4.74955
GRIA1	glutamate receptor, ionotropic, AMPA 1	2.5043

Table 1. List of genes up-regulated

Id	Gene Name	Fold-Change
SCGB1D2	secretoglobin, family 1D, member 2	-186.686
KRT4	keratin 4	-173.876
FOS	FBJ murine osteosarcoma viral oncogene homolog	-138.062
MAL	mal, T-cell differentiation protein	-133.173
WISP2	WNT1 inducible signaling pathway protein 2	-90.6242
UNQ698	suprabasin	-88.3632
FLJ22655	RERG/RAS-like	-88.2086
SPARCL1	SPARC-like 1 (hevin)	-86.9887
MGC45780	scavenger receptor class A, member 5 (putative)	-78.2205
RGS5	regulator of G-protein signaling 5	-73.238
MAMDC2	MAM domain containing 2	-69.9879
IGFBP6	insulin-like growth factor binding protein 6	-68.8765
C1orf10	cornulin	-67.0584
FHL1	four and a half LIM domains 1	-64.6338
TNA	C-type lectin domain family 3, member B	-59.3405
W57655	-	-57.5288
PLAC9	placenta-specific 9	-51.9896
DF	complement factor D (adipsin)	-50.8317
UNQ467	keratinocyte differentiation-associated protein	-50.2533
PTGDS	prostaglandin D2 synthase 21kDa (brain)	-50.1348
SCGB2A2	secretoglobin, family 2A, member 2	-49.633
DPT	dermatopontin	-47.7048
CRISP3	cysteine-rich secretory protein 3	-46.6048
BNC2	basonuclin 2	-46.1538
OGN	osteoglycin	-44.5308
EDN3	endothelin 3	-42.5161
DPT	dermatopontin	-38.5712
FXYD1	FXYD domain containing ion transport regulator 1	-36.7655
PPP1R3C	protein phosphatase 1, regulatory (inhibitor) subunit 3C	-35.3073
OSR2	odd-skipped related 2 (Drosophila)	-34.3512
ANKRD25	KN motif and ankyrin repeat domains 2	-32.6406
BF941677	-	-30.9503
MITF	microphthalmia-associated transcription factor	-30.7002
BQ007074	-	-30.1011
AL137566	-	-29.833
FBLN5	fibulin 5	-29.4494
PGM5	phosphoglucomutase 5	-28.8192
CRYAB	crystallin, alpha B	-28.3094
MFAP4	microfibrillar-associated protein 4	-27.6357
AA101632	-	-27.2784
COLEC12	collectin sub-family member 12	-26.3813
PCOLCE2	procollagen C-endopeptidase enhancer 2	-26.1851
COL14A1	collagen, type XIV, alpha 1	-25.1745
ARHGAP6	Rho GTPase activating protein 6	-24.9415
AEBP1	AE binding protein 1	-24.5738
TENC1	tensin like C1 domain containing phosphatase (tensin 2)	-22.8413
ANGPTL2	angiopoietin-like 2	-22.7692
COL4A1	collagen, type IV, alpha 1	-22.2278
LMO3	LIM domain only 3 (rhombotin-like 2)	-22.0822
NDRG2	NDRG family member 2	-21.5607
FLJ10970	transmembrane protein 100	-21.5054
CLCA4	chloride channel accessory 4	-21.3412
COX7A1	cytochrome c oxidase subunit VIIa polypeptide 1 (muscle)	-20.3295
MITF	microphthalmia-associated transcription factor	-20.1976
GNG11	guanine nucleotide binding protein (G protein), gamma 11	-20.1211
TACC1	transforming, acidic coiled-coil containing protein 1	-19.9851
ECG2	serine peptidase inhibitor, Kazal type 7 (putative)	-19.9095
AI734212	-	-19.527
AI765637	-	-19.1416
EBF	early B-cell factor 1	-18.9832
BAI3	brain-specific angiogenesis inhibitor 3	-18.8985
ARHGAP28	Rho GTPase activating protein 28	-18.5654
KCNAB1	assium voltage-gated channel, shaker-related subfamily, beta membe	-18.2379
AA578982	-	-18.1596
CECR6	cat eye syndrome chromosome region, candidate 6	-18.0476
AF321976	-	-17.2499
EDN3	endothelin 3	-16.2896
MYLK	myosin light chain kinase	-16.0175
C7	complement component 7	-15.7473

Table 2. List of genes and EST's (without gene name) down-regulated

Id	Gene Name	Fold-Change
SCGB1D2	secretoglobin, family 1D, member 2	-186.686
KRT4	keratin 4	-173.876
FOS	FBJ murine osteosarcoma viral oncogene homolog	-138.062
MAL	mal, T-cell differentiation protein	-133.173
WISP2	WNT1 inducible signaling pathway protein 2	-90.6242
UNQ698	suprabasin	-88.3632
FLJ22655	RERG/RAS-like	-88.2086
SPARCL1	SPARC-like 1 (hevin)	-86.9887
MGC45780	scavenger receptor class A, member 5 (putative)	-78.2205
RGS5	regulator of G-protein signaling 5	-73.238
MAMDC2	MAM domain containing 2	-69.9879
IGFBP6	insulin-like growth factor binding protein 6	-68.8765
C1orf10	cornulin	-67.0584
FHL1	four and a half LIM domains 1	-64.6338
TNA	C-type lectin domain family 3, member B	-59.3405
PLAC9	placenta-specific 9	-51.9896
DF	complement factor D (adipsin)	-50.8317
UNQ467	keratinocyte differentiation-associated protein	-50.2533
PTGDS	prostaglandin D2 synthase 21kDa (brain)	-50.1348
SCGB2A2	secretoglobin, family 2A, member 2	-49.633
DPT	dermatopontin	-47.7048
CRISP3	cysteine-rich secretory protein 3	-46.6048
BNC2	basonuclin 2	-46.1538
OGN	osteoglycin	-44.5308
EDN3	endothelin 3	-42.5161
DPT	dermatopontin	-38.5712
FXYD1	FXYD domain containing ion transport regulator 1	-36.7655
PPP1R3C	protein phosphatase 1, regulatory (inhibitor) subunit 3C	-35.3073
OSR2	odd-skipped related 2 (Drosophila)	-34.3512
ANKRD25	KN motif and ankyrin repeat domains 2	-32.6406
MITF	microphthalmia-associated transcription factor	-30.7002
FBLN5	fibulin 5	-29.4494
PGM5	phosphoglucomutase 5	-28.8192
CRYAB	crystallin, alpha B	-28.3094
MFAP4	microfibrillar-associated protein 4	-27.6357
COLEC12	collectin sub-family member 12	-26.3813
PCOLCE2	procollagen C-endopeptidase enhancer 2	-26.1851
COL14A1	collagen, type XIV, alpha 1	-25.1745
ARHGAP6	Rho GTPase activating protein 6	-24.9415
AEBP1	AE binding protein 1	-24.5738
TENC1	tensin like C1 domain containing phosphatase (tensin 2)	-22.8413
ANGPTL2	angiopoietin-like 2	-22.7692
COL4A1	collagen, type IV, alpha 1	-22.2278
LMO3	LIM domain only 3 (rhombotin-like 2)	-22.0822
NDRG2	NDRG family member 2	-21.5607
FLJ10970	transmembrane protein 100	-21.5054
CLCA4	chloride channel accessory 4	-21.3412
COX7A1	cytochrome c oxidase subunit VIIa polypeptide 1 (muscle)	-20.3295
MITF	microphthalmia-associated transcription factor	-20.1976
GNG11	guanine nucleotide binding protein (G protein), gamma 11	-20.1211
TACC1	transforming, acidic coiled-coil containing protein 1	-19.9851
ECG2	serine peptidase inhibitor, Kazal type 7 (putative)	-19.9095
EBF	early B-cell factor 1	-18.9832
BAI3	brain-specific angiogenesis inhibitor 3	-18.8985
ARHGAP28	Rho GTPase activating protein 28	-18.5654
KCNAB1	potassium voltage-gated channel, shaker-related subfamily, beta member 1	-18.2379
CECR6	cat eye syndrome chromosome region, candidate 6	-18.0476
EDN3	endothelin 3	-16.2896
MYLK	myosin light chain kinase	-16.0175
C7	complement component 7	-15.7473
CCNG1	cyclin G1	-15.5806
SLC16A9	solute carrier family 16, member 9 (monocarboxylic acid transporter 9)	-15.5025
EMCN	endomucin	-15.2301
NOPE	immunoglobulin superfamily, DCC subclass, member 4	-15.0347
CLST11240	HIG1 hypoxia inducible domain family, member 1B	-14.9221
PRLR	prolactin receptor	-14.8317
EPHA3	EPH receptor A3	-14.6125
ZNF404	zinc finger protein 404	-14.4942
C21orf34	chromosome 21 open reading frame 34	-13.8307
NDN	necdin homolog (mouse)	-13.5665
A2M	alpha-2-macroglobulin	-13.2787
HPSE2	heparanase 2	-13.2025
ZNF471	zinc finger protein 471	-12.8793
C9orf13	sushi, von Willebrand factor type A, EGF and pentraxin domain containing 1	-12.4415
GYPC	glycophorin C (Gerbich blood group)	-12.1484
DSCR1L1	regulator of calcineurin 2	-11.7187
IL17D	interleukin 17D	-11.3345
RGS5	regulator of G-protein signaling 5	-10.2954
OSBPL1A	oxysterol binding protein-like 1A	-10.2234
C21orf34	chromosome 21 open reading frame 34	-9.41903
LIMS2	LIM and senescent cell antigen-like domains 2	-9.37894

Table 2. List of genes down-regulated

4.1 Analyzing microarray data with novel software suite

If we want to display the data in just two dimensions, we want as much of the variation in the data as possible captured in just two dimensions. Principal component analysis or PCA has been developed for this purpose. Applying this PCA method in the cervical data we observed some expected differences (Figure 3).

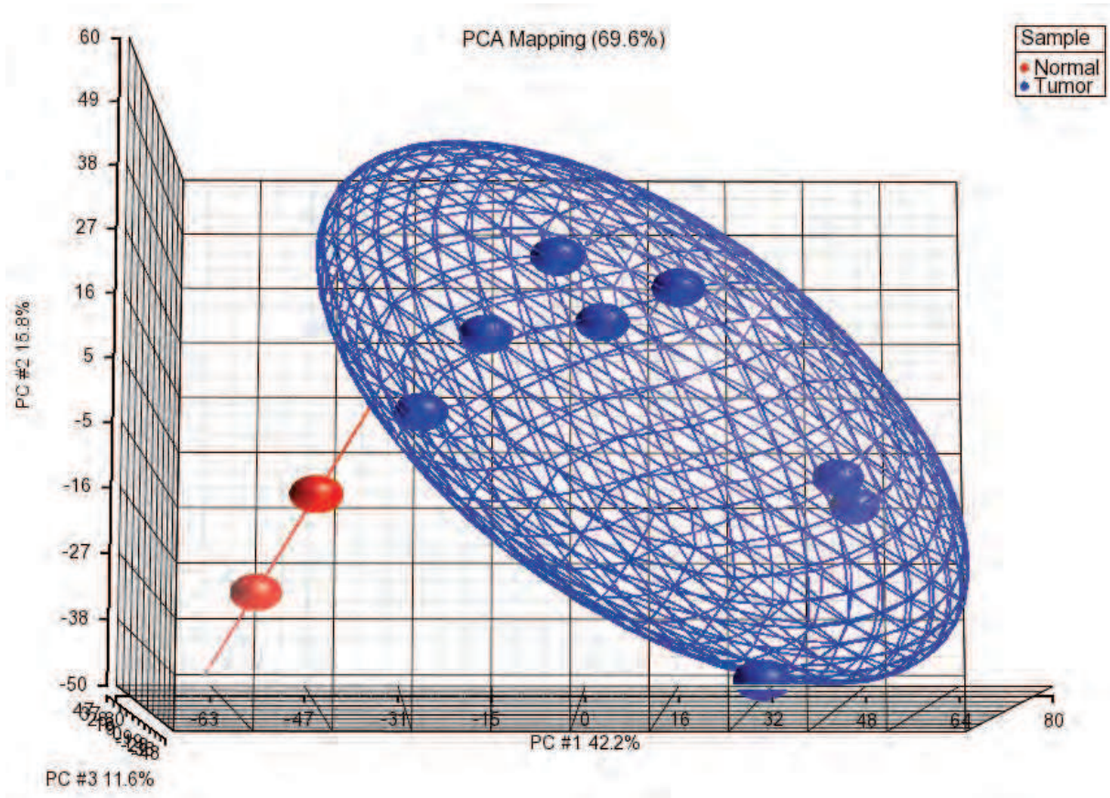


Fig. 3. Principal component analysis of cervical cancer samples. As expected, after perform PCA analysis, the normal samples (red balls) were grouped out of CC Group (blue balls) in a 3-D image.

In order to obtain a graphical representation of the differences between normal and tumor tissues, hierarchical cluster analysis was performed on all samples with a pseudo-color visualization matrix of the 208 selected genes grouping with greater intra-group similarity and differences between groups. The phylogenetic tree resulting from the hierarchical complete linkage-clustering algorithm is shown in figure 4. The figure shows those genes that are changing respect to health cervical tissue. In this method of clustering, allows to do relationships among objects (genes or samples) and are represented by a tree whose branch lengths reflect the degree of similarity between the objects, as assessed by a pairwise similarity function. The computed trees can be used to arrange individual samples in the original data table; this allows the samples or groups of samples with similar expression patterns to be shown adjacent to each other.

In general, tumor samples showed heterogeneity among them compared with normal samples, which had a more homogeneous gene expression profile. So, clustering analysis in CC failed to show significant segregation of patients based on expression profiling possibly

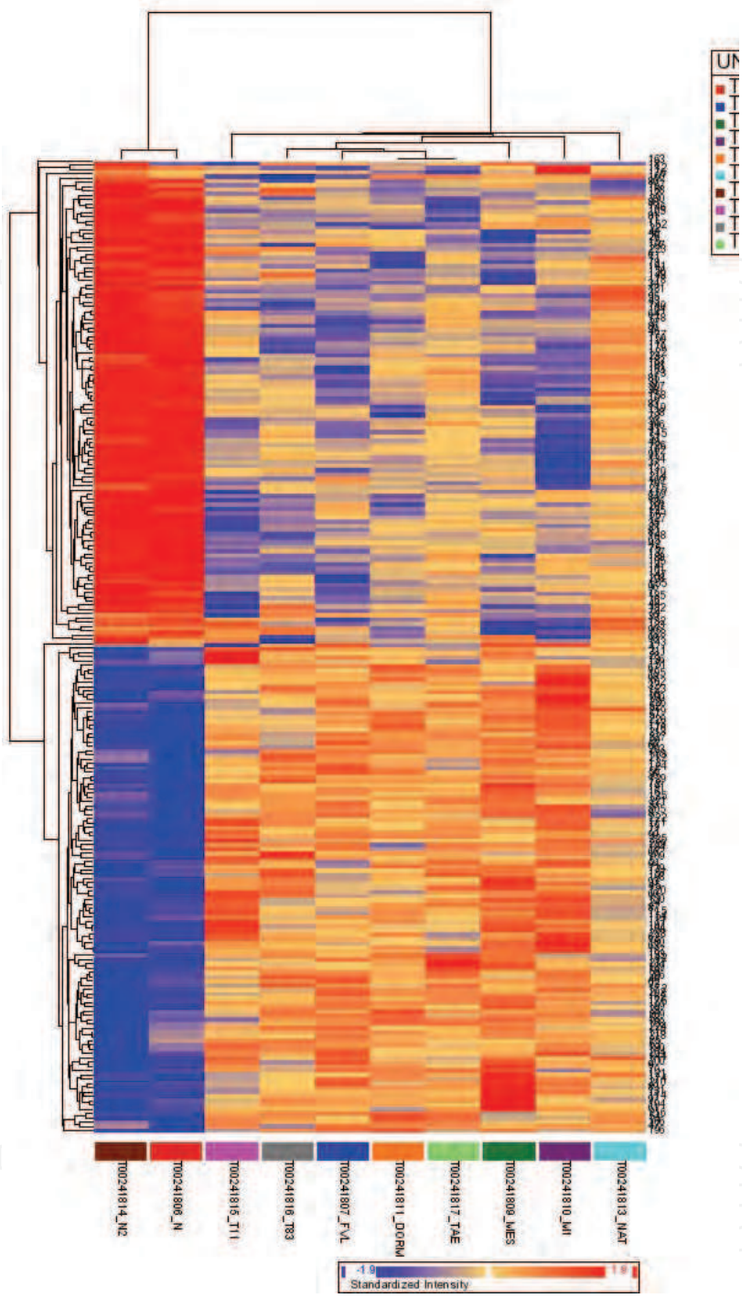
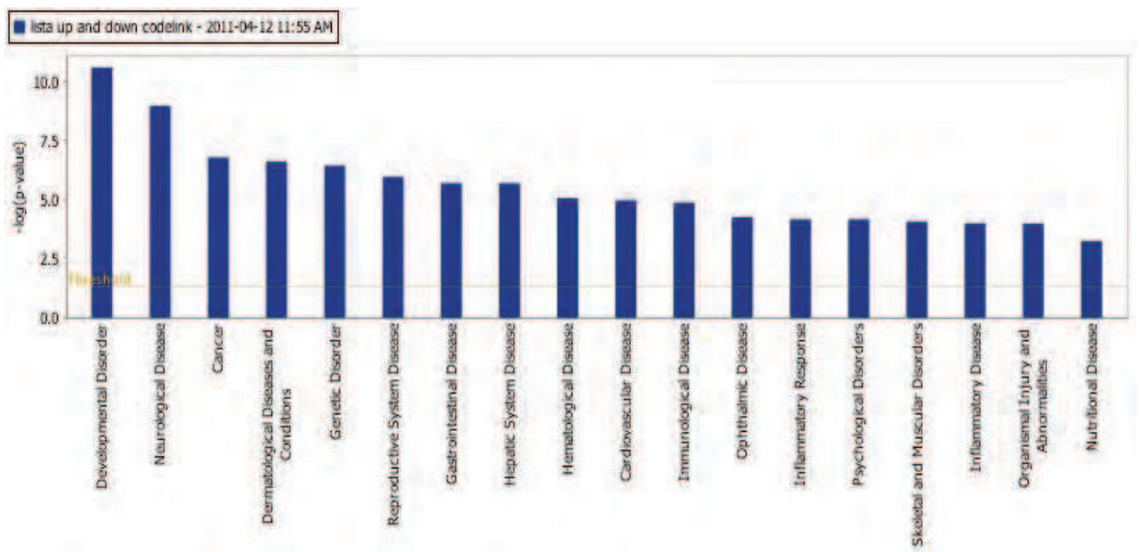


Fig. 4. Hierarchical clustering of the gene expression data for cervical tissues. Clustering analysis were performed for all tumours and normal samples. The data were clustered using the standard hierarchical method with ward linkage and using the Perason correlation to determine distance tumour. Before clustering, the data was filtered to remove genes that were scored absent in 75% or more of the samples as they are likely to be measuring noise in the system. The cluster of normal samples exhibit nearly identical patterns of gene expression changes, on contrary, as expected the invasive samples grouped in a different branch showing a heterogeneous gene expression. Blue color indentify downregulation and red color an overexpression status.

due to the heterogeneous nature of the samples as well as the relatively small numbers of samples in this study. Even when the samples were subjected to rigorous procedures of analysis, the special selection of the patients, including age, clinical stage, HPV16 positive and contraceptive oral status avoiding any bias, in this stage of the carcinogenesis process (stage IIb) the pattern of gene expression is quite different between samples.



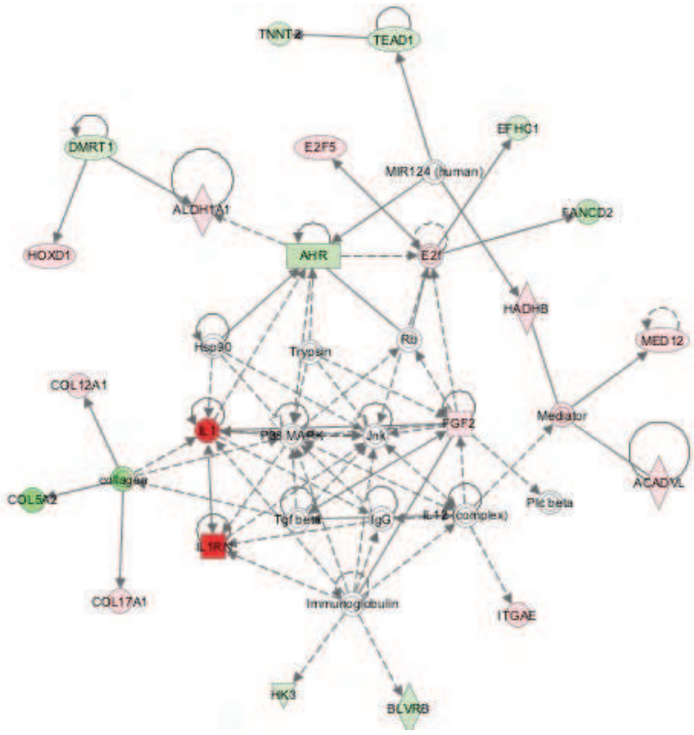
Status	Genes
Developmental disorder	
Up	*ALDH1A1, ALDOC, APAF1, BMP4, COMP, CREB1, CRHR1,COK1, DSP, FGF2, GHSR, GLI2, HPRT1, HRH, IL11, PTEN, ROR2, SLC16A2*
Down	*AGER, AHR, AGR1, ASCL1, ASPH, ATM, CAV3, DCC, DRD3, FANCD2, FGF8, FSL2, GATA4, GPX3, H2FZ, IKBKB, PTK2B, RMRP, TEAD1, TNNT2*
Neurological disease	
Up	*ALDH1A1, CD5, CENTFR, CREB1, CXCL10, E3F5, FGF2, GFI1, HLA-B, HPRT1, HRH1, HSPA5, IL1RN, IMPA2, KCNC3, MED12, PTEN, SCLC16A2,* ***PARK2, PEG10***
Down	*ADCYAP1, ADCYAP1R, AGER, ATM, CHI3L1, CHRNB2, DRD3, GJB2, IKBK, LPAR1, NDP, S1PR1*
Cancer tumorigenesis	
Up	*AKR1C1/2, ALDH1A1, ALDOC, APAF1, BMP4, BNC1, CALB2, CASP7, CD36, CD3G, CD247, CRABP1 CREB1, CTSL1, CXCL2, DAB2, DNASE1, DOK1, DSP, ENPEP, EPS15, FDXR, FGF2, GFI1, GLI2, HLA-B, HLA-E. HMX2,* ***HOXA9****, HPRT1, HR, HRH1, HSPA5, IL1RN, IL7R, INPP5D, ITGAE, KLF9, PTEN, RHOG, TXL1*
Down	*AGER, AHR, AMFR, AMH, ASCL1, ASPH, ATM, ATP2A2, BMP3, CFH, CFHR1, CHI3L1, CHRNA3, CHRBN2, CHRNE, CHRNG, CST6, DCC, DRDE, EIF4A1, EPJB6, FANCD2, FGF8, FHL2, FOLR1, GATA2, GJB2, GPX3, IGLL1, IKBK, LPAR, LTB4R, S1PR1, ADCYAPP1R1*

Table 3. Disease and Disorders functions in cervical cancer

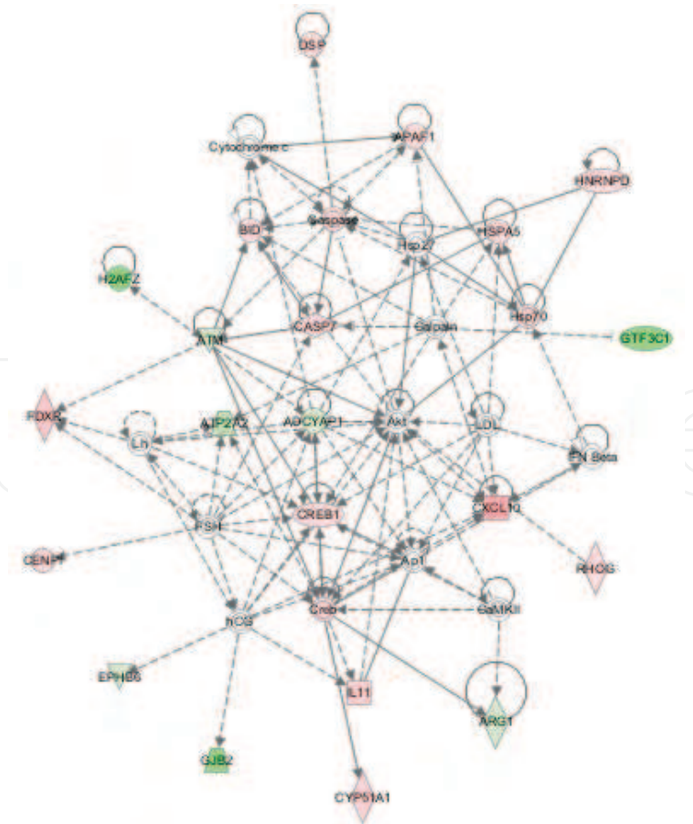
We used IPA to investigate the biological relevant of the observed genome-wide expressed gene changes by categorizing our data set into biological functions and/or diseases Ingenuity Pathway analysis was applied. The 208 genes annotated list by PARTEK analysis was submitted to the visualization IPA tool. This bioinformatics tool is employed for visualizing expression data in the context of KEGG biological pathways; the importance IPA is that retrieves an impact factor (IF) of genes that entire pathway involved, which can help to obtain a clearer notion of the alteration level in each biological pathway, and understand the complexity of these different process of the cancer cell. We imported a list of significantly up and down regulated genes (with extension .txt) into the program to convert the expression data into illustrations in an attempt to explore altered mechanisms in CC. To overcome any possible incorrect IF in altered pathways due to different size of samples, we submitted a similar quantity of up and down regulated genes. This allowed confirming that genes involved in several metabolic pathways were altered in CC (see networks).
We were able to associate biological functions and diseases to the experimental results. Fifteen pathways were obtained with a high score. Table 3 is showing the genes and the top three disorders/disease of "small networks" based in the analysis of the data. As can be seen, a clear route in cancer as it is known was not observed but some genes have been previously associated; however, these data give important information involving "*non canonical*" pathways in cancer.
Finally, in the Figure 5 is showed a "hypothetical network in CC" based from the 15 small networks. In addition to gene expression values, the proposed method uses Gene Ontology, which is a reliable source of information on genes. The use of Gene Ontology can compensate, in part, for the limitations of microarrays, such as having a small number of samples and erroneous measurement results.

5. Discussion

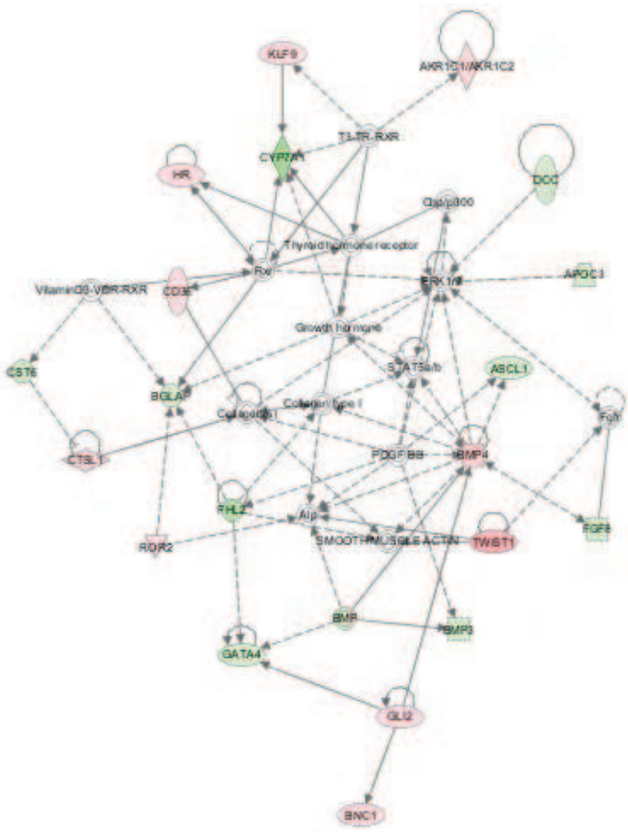
In our results, non classical "cancer genes" were conserved, respect to expected genes as MYC, FOS, RB, P53, HIF, etc. However, in the "strict sense of the word" when is considered a cancer gene? By instance, over-expression, down-regulation, point mutation, amplification, loss of heterozygosity, polymorphisms, epigenetic changes, etc. Thus, any gene could be considered like cancer gene, if they are following special criteria as recently was reported (27).
In this context, we decided to explore two non-related genes in cervical cancer *PARK2* gene. Interestingly, *PARK2* gene mutations (point mutations and exonic deletions) were first identified in autosomal recessive juvenile-onset parkinsonism. This gene is mapped to 6q25.2-q27 containing 12 small exons, and encodes parkin protein which functions as an E3 ligase, ubiquitinating proteins for destruction by the proteosome. Several substrates for parkin have been identified, including a 22kD glycosolated form of synuclein, parkin-associated endothelin receptor-like receptor (Pael-R), and CDCrel-1. Over-expression of Pael-R causes it to become ubiquinated, insoluble, and unfolded, and lead to endoplasmic reticulum stress and cell death (for review see 28). The location of Parkin is in a chromosomal region that is frequently deleted in multiple tumor types, including hepatocellular carcinoma (HCC), ovarian cancer, and breast cancer. The Parkin gene is within FRA6E, the third most active common fragile site (29,30). Interestingly, all three fragile sites regions were found consistently deleted in HCC (31) as well as in ovarian, breast, and prostate cancers. Further PARKIN protein overexpression did not lead to


TNNT2
TEAD1
DMRT1
EFHC1
E2F5
MIR124 (human)
ALDH1A1
FANCD2
HOXD1
AHR
E2f
HADHB
Hsp90
Trypsin
Rb
MED12
COL12A1
IL1
PGF2
Mediator
collagen
ACADVL
COL5A2
Tgf beta
IgG
IL12 (complex)
Pkc beta
IL1R
COL17A1
Immunoglobulin
ITGAE
HK3
BLVRB

A


DSP
Cytochrome c
APAF1
HNRNPD
BID
Caspase
Hsp27
HSPA5
H2AFZ
ATM
CASP7
Calpain
Hsp70
GTF3C1
FOXO
ATP2A2
ADCYAP1
Akt
DL
Lh
EN Beta
CXCL10
FSH
CREB1
CENPF
Ap1
RHOG
Creb
CaMKII
NOS
EPHB6
IL11
ARG1
GJB2
CYP51A1

B



C

D

Fig. 5. Networks in cervical cancer. Top three networks diagrams generated as graphical representations of the molecular relationship between genes and gene product. The gene

products are represented as nodes (shapes) and the biological relationship between two nodes is represented as an edge (line). A) Network 1. skeletal and muscular system development and function, embryonic development, tissue development, B) network 2; dermatological diseases and conditions, cardiovascular disease, organismal injury and abnormalities, C) network 3; cell cycle, organismal functions, organismal injury and abnormalities, D) Ingenuity Pathways Analysis network of genes associated with CC. This network diagram shows the biological association of 8 focus networks associated with cervical cancer as graphical representation of the molecular relationship between genes/gene produtcs.

increased sensitivity to all pro-apoptotic induction but may show specificity for a certain type of cellular stress. At present, *Parkin* gene could be considered as new tumor suppressor gene (32). In our case, *per se* Parkin expression in CC is interesting. It is widely described that p53 master gene is constitutively expressed but under stress conditions as HPV infection, DNA damage or point mutations its half-time life of the protein is increased. Similar situation might be observed for *Parkin* gene due to increased expression in CC. This could be supported because the 6q25 cytogenetic region in CC is not altered as happen for *TP53* gene (33). To this respect, we could hypothesize that a new overexpression of *Parkin* gene could be involved in invasion cervical carcinogenesis. These findings could demonstrate that the genetic context in which a mutation occurs can play a significant role in determining the type of illness produced or associated.

It has been established that although we inherit two copies of all genes (except those that reside on the sex chromosomes), there is a subset of these genes in which only the paternal or maternal copy is functional. This phenomenon of monoallelic, parent-of-origin expression of genes is termed genomic imprinting. Imprinted genes are normally involved in embryonic growth and behavioral development, but occasionally they also function inappropriately as oncogenes and tumor suppressor genes (34). Furthermore, it is well know that a variety of genetic changes influence the development and progression of cancer. These changes may result from inherited or spontaneous mutations that are not corrected by repair mechanisms prior to DNA replication. It is increasingly clear that so called epigenetic effects that do not affect the primary sequence of the genome also play an important role in tumorigenesis (35).

Other gene overexpressed seen in this analysis was *PEG10* gene. This gene is mapped in chromosme 7q21. PEG10 protein prevents apoptosis in hepatocellular carcinoma cells through interaction with SIAH1, a mediator of apoptosis. May also have a role in cell growth promotion and hepatoma formation. Inhibits the TGF-beta signaling by interacting with the TGF-beta receptor ALK1. This is a paternally expressed imprinted gene that encodes transcripts containing two overlapping open reading frames (ORFs), RF1 and RF1/RF2, as well as retroviral-like slippage and pseudoknot elements, which can induce a -1 nucleotide frame-shift. Increased expression of this gene is associated with hepatocellular carcinomas. These findings link to cancer genetics and epigenetic by showing that a classic proto-oncogene, *MYC*, acts directly upstream of a proliferation-positive imprinted gene, PEG10 (36,37).

The *HOX* genes are a family of transcription factors that bind to specific sequences of DNA in target genes regulating their expression. The role of *HOX* genes in adult cell differentiation is still obscure, but growing evidence suggests that they may play an important role in the development of cancer. We have previously reported that some *HOX*

genes could be related to CC. Specifically, *HOXA9* was observed expressed in cervical cancer by RT-PCR end point. In the present work, the data are showing that statistically significative *HOXA9* gene is differentially expressed in CC. Together to *HOXB13, D9, D10*, and *HOXC* cluster (*HOXC9, C11–C13*) genes this family of genes might be an important factor involved in CC (35).

It is clear that the most altered genes in CC are not commonly associated to cancer process. This fact could suggest: 1) the "classic genes of cancer" are statistically significant altered with tiny values, but there are some exceptions and specific tumor types as neuroblastomas and *N-MYC* gene, *Her2/neu* in breast cancer. 2) At least in stage IIb of cervical carcinogenesis could be involved genes related to "cellular economy" but not belonging to genes of cancer. This is supported by recent reports showing molecular alterations in genes not previously related to cancer (27). 3) The extreme values (high or low) in microarray analysis not always represent strong candidates of markers in the models performed. What about the most frequent?. (4) Integrative genomics, multicentre protocols in well selected samples, stratified stages and clinical follow-up, will be the clue to get cancer hallmarks. In addition, the study of the molecular function of selected genes strengthened the hypothesis that these genes are involved in the process of cancer growth.

The data information obtained from microarray analysis should be validated because can appear errors in positive and negative false due to nature of the massive assays. In this context, in order to confirm the microarray data, additional molecular tool as end point PCR, real time PCR, northern blot, immunohistochemistry should be performed and to obtain results. (Mendez S. An Integrative microarray gene expression analysis, approach identifies candidates' array multi-experiments in Ovary Tumours, submitted to publication 2011).

6. Acknowledgements

This work was partially supported by CONACYT (México) grants 69719 AND 87244 from FONDOS SECTORIALES. We appreciate the technical assistance of Laboratorio de Oncología Genómica, CIS, HO-IMSS. Sergio JUAREZ and Mauricio SALCEDO made similar efforts in the present work.

7. References

[1] Parkin DM, Bray F, Ferlay J, Pisani P. Global cancer statistics, 2002. CA Cancer J Clin. 2005, 55(2):74-108

[2] Secretaría de Salud (Méx). Registro Histopatológico de Neoplasias en México. México: Secretaría de Salud, México D.F; 1999.

[3] Munoz N, Xavier Bosch F. Cervical cancer and human papillomavirus: epidemiological evidence and perspectives for prevention. Salud Publica Mex 1997; 39:274–282.

[4] Boffetta P, Parkin DM. Cancer in developing countries. CA Cancer J Clin 1994 44:81-90

[5] Walboomers JM, Jacobs MV, Manos MM, Bosch FX, Kummer JA, Shah K, et al. Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. J Pathol 1999, 189: 12–19.

[6] Scheffner, M., B. A. Werness, J. M. Huibregtse, A. J. Levine, and P. M. Howley. The E6 oncoprotein encoded by human papillomavirus types 16 and 18 promotes the degradation of p53. Cell 1990, 63:1129–1136.

[7] Boyer SN, Wazer DE, Band V. E7 protein of human papilloma virus-16 induces degradation of retinoblastoma protein through the ubiquitin-proteasome pathway. Cancer Res. 1996, 56(20):4620-4624.

[8] Duensing S, Munger K. The human papillomavirus type 16 E6 and E7 oncoproteins independently induce numerical and structural chromosome instability. Cancer Res. 2002, 62(23):7075-7082.

[9] Nees M, van Wijngaarden E, Bakos E, Schneider A, Dürst M. Identifcation of novel molecular markers which correlate with HPVinduced tumor progression. Oncogene 1998, 16:2447–2458.

[10] Brown P, Botstein R. Exploring the new world of the genome with DNA microarrays. Nature Genet 1999, 21 (suppl):33-37.

[11] Lockhart DJ, Winzeler EA. Genomics, gene expression and DNA arrays. Nature, 2000, 405:827-836.

[12] Ruutu M, Peitsaro P, Johansson B, Syrjanen S: Transcriptional profiling of a human papillomavirus 33-positive squamous epithelial cell line which acquired a selective growth advantage after viral integration. Int J Cancer 2002, 100: 318-326.

[13] Duffy CL, Phillips SL, Klingelhutz AJ: Microarray analysis identifies differentiation-associated genes regulated by human papillomavirus type 16 E6. Virol 2003, 314: 196-205.

[14] Thomas JT, Oh ST, Terhune SS, Laimins LA: Cellular changes induced by low-risk human papillomavirus type 11 in keratinocytes that stably maintain viral episomes. J Virol 2001, 75: 7564-7571.

[15] Garner-Hamrick PA, Fostel JM, Chien WM, Banerjee NS, Chow LT, Broker TR, Fisher C: Global effects of human papillomavirus type 18 E6/E7 in an organotypic keratinocyte culture system. J. Virol 2004, 78: 9041-9050.

[16] Toussaint-Smith E, Donner DB, Roman A: Expression of human papillomavirus type 16 E6 , E7 oncoproteins in primary foreskin keratinocytes is sufficient to alter the expression of angiogenic factors. Oncogene 2004, 23: 2988-2995.

[17] Nees M, Geoghegan JM, Hyman T, Frank S, Miller L, Woodworth CD: Papillomavirus type 16 oncogenes downregulate expression of interferon-responsive genes and upregulate proliferation-associated, NF-kappaB-responsive genes in cervical keratinocytes. J Virol 2001, 75: 4283-4296.

[18] Shim C, Zhang, Hun C, Lee. Profiling of differentially expressed genes in human primary cervical cancer by complementary DNA expression array. Clin Cancer Res 1998, 4:3045-3050.

[19] Vazquez-Ortiz G, García JA, Ciudad CJ, Noé V, Peñuelas S, López-Romero R, Mendoza-Lorenzo P, Piña-Sánchez P, Salcedo MDifferentially expressed genes between high-risk human papillomavirus types in human cervical cancer cells. Int J Gynecol Cancer. 2007, 17:484-491

[20] Pérez-Plasencia C, Vázquez-Ortiz G, López-Romero R, Piña-Sanchez P, Moreno J, Salcedo M. Genome wide expression analysis in HPV16 cervical cancer: identification of altered metabolic pathways. Infect Agent Cancer. 2007, 6;2:16

[21] Davidson LA, Nguyen DV, Hokanson RM, Callaway ES, Isett RB, Turner ND, Dougherty ER, Wang N, Lupton JR, Carroll RJ, Chapkin RS. Chemopreventive n-3 polyunsaturated fatty acids reprogram genetic signatures during colon cancer initiation and progression in the rat. Cancer Res. 2004, 64:6797-6804.

[22] Ramakrishnan R, Dorris D, Lublinsky A, Nguyen A, Domanus M, Prokhorova A, Gieser L, Touma E, Lockner R, Tata M, Zhu X, Patterson M, Shippy R, Sendera TJ, Mazumder A. An assessment of Motorola CodeLink microarray performance for gene expression profiling applications. Nucleic Acids Res. 2002, 30:e30.

[23] Nie AY, McMillian M, Parker JB, Leone A, Bryant S, Yieh L, Bittner A, Nelson J, Carmen A, Wan J, Lord PG. Predictive toxicogenomics approaches reveal underlying molecular mechanisms of nongenotoxic carcinogenicity. Mol Carcinog. 2006, 45:914-933.

[24] Wu W, Dave N, Tseng G, Richards T, Xing EP, Kaminski N. Comparison of normalization methods for CodeLink Bioarray data. BMC Bioinformatics 2005, 6:309

[25] Reich D, Price Al, Pattterson N. Principal component analysis of genetic data. Nature Genetics. 2008, 40(5): 491-492

[26] Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci U S A 1998, 95: 14863–14868

[27] Seltzer MJ, Bennett BD, Joshi AD, Gao P, Thomas AG, Ferraris DV, Tsukamoto T, Rojas CJ, Slusher BS, Rabinowitz JD, Dang CV, Riggins GJ. Inhibition of glutaminase preferentially slows growth of glioma cells with mutant IDH1. Cancer Res. 2010, 70:8981-8987.

[28] Schapira A. Etiology of Parkinson's disease Neurol 2006, 4:S10-S23

[29] Kahkonen M. Population cytogenetics of folate-sensitive fragile sites. Common fragile sites. Hum Genet 1988, 80:344–348.

[30] Denison SR, Wang F, Mueller B, Kock N, Phillips LA, Klein C, Smith DI. Genetic alteration in the common fragile site gene PARK2 in ovarian and other cancers. Oncogene 2003, 22:8370– 8378.

[31] Zhao P, Song X, Nin YY, Lu YL, Li XH. Loss of fragile histidine triad protein in human hepatocellular carcinoma. World J Gastroenterol 2003, 9:1216–1219.

[32] Wang F, Denison S, Lai JP, Philips LA, Montoya D, Kock N, Schule B, Klein C, Shridhar V, Roberts LR, Smith D. Parkin Gene Alterations in Hepatocellular Carcinoma. Genes Chromosome & cancer 2004, 40:85–96

[33] Hidalgo A, Baudis M, Petersen I, Arreola H, Piña P, Vazquez G, Hernandez D, Gonzalez JL, Lazos M, Lopez R, Salcedo M. Microarray comparative genomic hybridization detection of chromosomal imbalances in uterine cervix carcinoma. BMC Cancer 2005, 5, 77

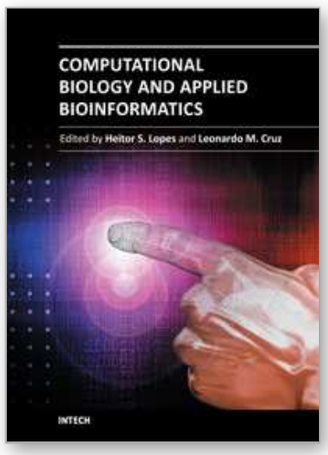
[34] Jirtle RL. Genomic imprinting and cancer. Exp Cell Res. 1999, 248:18-24.

[35] Plass C, Soloway PD. DNA methylation, imprinting and cancer. Eur J Hum Genet 2002, 10:6-16

[36] Li CM, Margolin AA, Salas M, Memeo L, Mansukhani M, Hibshoosh H, Szabolcs M, Klinakis A, Tycko B. PEG10 Is a c-MYC Target Gene in Cancer Cells. Cancer Res 2006, 66: 665-672

[37] Tsuji K, et al. PEG10 is a probable target for the amplification at 7q21 detected in hepatocellular carcinoma. Cancer Genet Cytogenet, 2010, In press.

[38] López R, Garrido E, Vázquez G, Piña P, Pérez C, Alvarado I, Salcedo M. A subgroup of HOX Abd-B gene is differentially expressed in cervical cancer. Int J Gynecol Cancer. 2006, 16:1289-96



Computational Biology and Applied Bioinformatics

Edited by Prof. Heitor Lopes

ISBN 978-953-307-629-4

Hard cover, 442 pages

Publisher InTech

Published online 02, September, 2011

Published in print edition September, 2011

Nowadays it is difficult to imagine an area of knowledge that can continue developing without the use of computers and informatics. It is not different with biology, that has seen an unpredictable growth in recent decades, with the rise of a new discipline, bioinformatics, bringing together molecular biology, biotechnology and information technology. More recently, the development of high throughput techniques, such as microarray, mass spectrometry and DNA sequencing, has increased the need of computational support to collect, store, retrieve, analyze, and correlate huge data sets of complex information. On the other hand, the growth of the computational power for processing and storage has also increased the necessity for deeper knowledge in the field. The development of bioinformatics has allowed now the emergence of systems biology, the study of the interactions between the components of a biological system, and how these interactions give rise to the function and behavior of a living being. This book presents some theoretical issues, reviews, and a variety of bioinformatics applications. For better understanding, the chapters were grouped in two parts. In Part I, the chapters are more oriented towards literature review and theoretical issues. Part II consists of application-oriented chapters that report case studies in which a specific biological problem is treated with bioinformatics tools.

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820

Fax: +385 (51) 686 166
www.intechopen.com

Fax: +86-21-62489821

IntechOpen