Peer-reviewed version available at Genes 2019, 10, 560; doi:10.3390/genes10080560

Article

The Gene Master Regulators (GMR) Approach Provides Legitimate Targets for Personalized, Time-Sensitive Cancer Gene Therapy

Sanda Iacobas, Nneka Ede and Dumitru A Iacobas*

Personalized Genomics Laboratory, Center for Computational Systems Biology, Prairie View A&M University, Prairie View, TX, USA; sandaiacobas@gmail.com (S.I.), nede1@student.pvamu.edu (N.E.), daiacobas@pvamu.edu (D.A.I.)

* Correspondence: daiacobas@pvamu.edu; Tel.: +1(936) 261-9926

Abstract: The dynamic and never exactly repeatable tumor transcriptomic profile of people affected by the same form of cancer requires a personalized and time-sensitive approach of the gene therapy. The Gene Master Regulators (GMRs) were defined as genes whose highly controlled expression by the homeostatic mechanisms commands the cell phenotype by modulating major functional pathways through expression correlation with their genes. The Gene Commanding Height (GCH), a measure that combines the expression control and expression correlation with all other genes, is used to establish the gene hierarchy in each cell phenotype. We developed the experimental protocol, the mathematical algorithm and the computer software to identify the GMRs from transcriptomic data in surgically removed tumors, biopsies or blood from cancer patients. The GMR approach is illustrated with applications to our microarray data on human kidney, thyroid and prostate cancer samples, and on thyroid, prostate and blood cancer cell lines. We proved experimentally that each patient has his/her own GMRs, that cancer nuclei and surrounding normal tissue are governed by different GMRs, and that manipulating the expression has larger consequences for genes with higher GCH. Therefore, we launch the hypothesis that silencing the GMR may selectively kill the cancer cells from a tissue.

Keywords: papillary thyroid cancer; BCPAP cells; 8505C cells; prostate cancer; LNCaP cells; DU145 cells; kidney cancer; HL-60 cells; cancer gene software

1. Introduction

A very rich literature compared gene expression profiles in tissues collected from healthy and cancer donors to identify the transcriptomic signatures of various cancer phenotypes [e.g. 1-5] that are periodically organized in the atlas form [e.g. 6,7]. Nanostring launched recently a panel claiming to categorize the disease heterogeneity using 32 biological signatures involving 770 genes across 23 key breast cancer pathways (https://www.nanostring.com/products/gene-expression-panels/gene-expression-panels-overview/ncounter-breast-cancer-360-panel). There are also available platforms to compare the gene expression profiles of surgically removed tumors with publically available transcriptomes of cancer standard samples (e.g.: https://www.origene.com/products/tissues/tissuescan).

However, comparing samples collected from different persons may not be such a good idea, owing that, in addition to the disease itself, the gene expressions depend on several other risk factors making each human unique and with a unique life pathway. The never repeatable combination of factors affecting the gene expression profile is related to the person's race, sex, age, genetic background, diet (affecting the microbiome), environment (exposure to ionizing radiation, carcinogenic toxins, stress), bad habits (smocking, drugs, alcohol), medial history etc. Our gene expression studies on tissues from humans and animal models proved the transcriptomic profile dependence on strain and genetic background [8], sex [9], age [10], exposure to stress [11] and

Peer-reviewed version available at Genes 2019, 10, 560; doi:10.3390/genes10080560

carcinogenic toxins [12], medical history and treatments [13]. This is why numerous investigators (e.g. [14] on papillary thyroid cancer) started to pair the transcriptomes of the cancer region with the cancer free adjacent tissue of the same patient.

The legitimate question in the transcriptomic signature quest is how many of the tested genes should be found as regulated (and how many of them up and how many down) to assess the designated form of cancer. Since there are 1.9×10^{22} distinct sets of genes if "only" 10 hits are needed from the 770 candidates of the nanostring nCounter® breast cancer 360TM), there is no way to determine the predictive value of each of these sets from metadata. Moreover, in addition to the checked biomarkers, hundreds other genes are regulated and their (never repeatable) contributions to the cancer phenotype are neglected without knowing whether they are really negligible.

Still, let us suppose that a particular cancer form does have a transcriptomic signature as resulted from the meta-analysis of gene expression data from a large population of cancer patients. Are the signature genes valuable targets for the cancer gene therapy or they are good only for diagnosis (if the above supposition is true)? Being selected from the most frequently regulated genes in the population of cancer patients, the signature genes appeared as little protected by the cellular homeostatic mechanisms as are minor players. Therefore, restoring their right expression level may be of little consequence for the cell.

Instead, of genes whose altered sequence or expression allegedly triggers a particular form of cancer in everybody [15-20], we proposed [21,22] that the most legitimate targets for cancer gene therapy are what we call "gene master regulators" (GMRs) of cancer nuclei. We defined the GMR as the gene whose highly protected expression level by the cellular homeostatic mechanisms sets up the cell phenotype by controlling major functional pathways through expression correlation with their genes. The high protection makes the GMR less sensitive to the environmental oscillations and therefore less variably expressed among biological replicas. However, small oscillations of the GMR expression are amplified by in-phase (positive) or anti-phase (negative) oscillations of the expression of many other genes. The composite metric termed Gene Commanding Height (GCH) was introduced by us [21,22] to establish the gene hierarchy in each cell phenotype, with the GMR having the top GCH.

The idea of "master regulators" was floating in genomics for a long time, most investigators looking for transcription factors whose regulation might have large downstream effects on the expression of many genes [e.g.: 23,24]. In addition to defining the GMRs in quantitative terms (by the GCH), our procedure does not restrict the GMR's quest to transcription factors. Instead, we rank with respect to the GCH scores all coding AND non-coding RNAs whose abundance was adequately quantified via the used (RNA next generation sequence or microarray) platform. Moreover, in a recent paper [22], we have shown how five non-coding RNAs (ANKRD36BP2, FAM86B3P, H19, HCG11 and PMS2L2) regulate apoptosis in a surgically removed papillary thyroid cancer via expression correlation with apoptotic genes. Thus, our results are in line with other studies reporting the involvement of the non-coding RNAs in cancer development (e.g. [25-7]) and therapy (e.g. [28]).

The GMR approach is based on our Genomic Fabric Paradigm [e.g.: 29,30] and runs on the computer software package CANCER-GMR (coded in Python3) with statistical and graphical user interface packages SciPy (https://www.scipy.org/) and Tkinter GUI. Our procedure can be applied to the four quarters of >1mm diameter cancer nucleus identified in a biopsy or surgically removed tumor, or to four dishes with a cancer cell line.

The GMR targeting would be effective in selectively destroying the cancer cells from a tissue if: i) cancer nuclei and surrounding quasi-normal tissues are governed by different GCH hierarchies, ii) expression manipulation of a gene has larger consequences in cells where that gene has higher GCH and iii) the GCH of the GMR is well above the GCHs of the next genes in the hierarchy. In this report, the GMR approach is illustrated with applications to our microarray data on human kidney, thyroid and prostate cancer samples, and on thyroid, prostate and blood cancer cell lines.

2. Materials and Methods

2.1. Tumor samples

We have profiled the cancer nuclei (labeled CANCER1 and CANCER2, both Gleason score 4+5=9) and the surrounding normal tissues (NORM1 and NORM2) from two surgically removed, frozen prostate cancers. The study was part of Dr. DA Iacobas' project approved by the Institutional Review Boards (IRB) of the New York Medical College's (NYMC) and Westchester Medical Center (WMC) Committees for Protection of Human Subjects. The approved IRB (L11,376 from 02/10/2015) granted access to frozen cancer specimens from the WMC Pathology Archives and depersonalized pathology reports, waiving patient's informed consent. Four 2-8 mm^3 samples were collected from the cancer nuclei and normal tissues of each tumor. Although the selected regions were as homogeneous as possible, cells of different phenotypes were not completely eliminated and expression of their genes affected the reported results.

In this report, we reprocessed also our previously published microarray data from surgically removed, frozen preserved kidney (CCRCC - clear cell renal cell carcinoma) and papillary thyroid cancer tumors. Two primary cancer nuclei (labeled as PTA and PTB) and the cancer free (NOR) tissue from the right kidney, together with a chest metastasis (MET) were profiled from a 74 years old man with metastatic CCRCC, Fuhrman grade 3 [21]. The unilateral, single, papillary carcinoma (PAP-C), pathological stage pT3NOMx and the cancer free surrounding tissue were collected from a deceased 33y old woman [22].

2.2. Cell lines

The results for the surgically removed thyroid and prostate tumors were compared with those obtained from the commercially available standard human cancer cell lines: BCPAP, 8505C, LNCaP and DU145. We have also determined the GCH hierarchy in the human leukemia cell line HL-60. The HL-60 cell line was originally obtained at MD Anderson Cancer Center from a 36-year-old woman with acute promyelocytic leukemia [31].

The BCPAP cell line is a papillary thyroid carcinoma cell line isolated from a female patient, with a TP53 mutation in the codon 278 in heterozygosity (Pro→Leu) [32]. The 8505C cell line was established from undifferentiated thyroid carcinomas of a 78-year-old-female patient. Her tumor contained also residual well differentiated components, suggesting "well differentiated to undifferentiated carcinoma progression" [33]. We used both BCPAP and 8505C cell lines to test whether manipulation of the expression of a gene has large transcriptomic consequences if that gene has a higher GCH.

The LNCaP cells (Lymph Node Carcinoma of the Prostate) are androgen-sensitive adherent epithelial cells, obtained from a 50-year-old white male in 1977 [34]. The DU145 hormone insensitive cells were derived in 1976 from prostate adenocarcinoma metastatic to the brain of a 70 year old white male [35].

2.3. Biological replicas

The biological replicas (the quarters of a quadrisected homogeneous region of a tumor or four cell culture dishes of a cell line) can be considered as being the same system but subjected to slightly different environmental conditions. As such, the transcriptomic data provide valuable information on how much the genes resist or adapt to the external influences and how the variations of their expression levels are correlated to optimize the functional pathways. From the expression values of each coding and non-coding RNA in the biological replicas we derive three independent measures to be used in subsequent analyses: i) average level, ii) coefficient of variation and iii) expression correlation with each other RNA. The average expression level is used to identify what gene is up/down-regulated when compared two conditions. The coefficient of variation (CV) is used to estimate the control of the transcript abundance in each condition, and the expression correlation to identify and quantify the transcriptomic networks.

2.4. Microarray

We have used our standard protocol [36] for RNA extraction, purification, reverse transcription and fluorescent labeling, and hybridization with Agilent human 4x44k gene expression two-color G2519F microarrays. The chips were scanned with an Agilent G2539A dual laser scanner and primary analysis performed with (Agilent) feature extraction v. 12.0 software. All spots with corrupted or saturated pixels, or with forward fluorescence less than twice the background one in any of the four profiled biological replicas were removed from the analysis of that type of samples.

2.5. Relative Expression Variation (REV) and Relative Expression Stability (RES)

All microarray platforms probe transcripts redundantly by several (unfortunately not uniform numbers of) spots. Therefore, instead of the coefficient of variation for the expression level in biological replicas as determined by one spot we use the Relative Expression Variation (REV, [37]) that takes into account all spots R_i probing redundantly the same transcript i. REV is the mid-interval chi-square estimate of the pooled expression level CVs in biological replicates of that condition (cancer or normal) with a pre-established probability ε and number of degrees of freedom derived from the number of spots probing the same transcript:

$$REV_i^{(condition)}(\varepsilon) = \frac{1}{2}\left(\sqrt{\frac{r_i}{\chi^2(r_i;1-\varepsilon/2)}} + \sqrt{\frac{r_i}{\chi^2(r_i;\varepsilon/2)}}\right)\sqrt{\frac{1}{R_i}\sum_{k=1}^{R_i}\left(\frac{s_{ik}^{(condition)}}{\mu s_{ik}^{(condition)}}\right)^2}$$

s_{ki} = standard deviation of the expression level of gene i probed by spot k

μ_{ik} = average expression level of gene i probed by spot k

$r_i = \lambda R_i - 1 =$ number of degrees of freedom

λ = number of biological replicas ($\lambda \geq 4$)

R_i = number of microarray spots probing redundantly transcript i

(1)

In our experiments with Agilent 4x44k microarrays the number of spots probing redundantly the same transcript ranges from 1 to 28, so that r = 3, 4, … , 71. Our Cancer GMR software has uploaded the chi-square values for all these values of r and for ε = 0.010, 0.025, 0.050, 0.100, values less than 0.010 being useless because of the technical noise affecting the gene expression levels. However, in the applications presented here, we used λ = 4 and ε = 0.05 for which the correction coefficient of the CV ranges from 1.566 (R = 1) to 0.960 (R = 28). Therefore, λ and ε will be omitted from the next equations.

Relative Expression Stability (RES) is a measure we introduce to rank the priorities of the cell homeostatic mechanisms in controlling the right abundance of a particular transcript. RES is applied to all transcripts regardless of them translating into proteins or having only regulatory roles for the expression of other transcripts.

$$RES_i = \ln\left(\frac{\langle REV\rangle}{REV_i}\right) \quad , \quad \langle REV\rangle = \text{median REV for all transcripts} \qquad (2)$$

The log form was selected to assign positive values to the more stably expressed and negative ones to the less stably expressed genes than the median one.

2.6. Expression regulation

Instead of an arbitrarily introduced (e.g. 1.5x or 2.0x) absolute fold-change, we consider a gene as significantly regulated in cancer with respect to the normal counterpart if the absolute expression ratio |x| exceeds the cut-off calculated for that gene.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 June 2019 doi:10.20944/preprints201906.0198.v1

Peer-reviewed version available at Genes 2019, 10, 560; doi:10.3390/genes10080560

$$|x_i| > CUT_i = 1 + \sqrt{2\left(\left(REV_i^{(cancer)}\right)^2 + \left(REV_i^{(normal)}\right)^2\right)} \quad , \quad where:$$

$$|x_i| = \begin{cases} \dfrac{\mu_i^{(cancer)}}{\mu_i^{(normal)}} & , \text{ if } \mu_i^{(cancer)} > \mu_i^{(normal)} \\ -\dfrac{\mu_i^{(normal)}}{\mu_i^{(cancer)}} & , \text{ if } \mu_i^{(cancer)} < \mu_i^{(normal)} \end{cases} \quad , \quad with: \begin{cases} \mu_i^{(cancer)} = \overline{\mu_{ik}^{(cancer)}} \\ \mu_i^{(normal)} = \overline{\mu_{ik}^{(normal)}} \end{cases} \tag{3}$$

CUT observes the uncertainty about expression regulation by taking into account the contributions of both biological variability and technical noise. It is not uniform among the quantified transcripts and takes >1 values that may be even smaller than 1.5.

However, rather than the popular percentage of the regulated out of quantified genes, we measure the change in the transcriptional profiles by the Weighted Pathway Regulation (WPR). WPR is not restricted to the regulated but ponders all quantified genes. We have previously used WPR to quantify the remodeling and recovery of functional genomic fabrics in heart [38], hypothalamus [13] and hippocampus [39]:

$$WPR \equiv \left\langle \mu_i^{(normal)} \left(\left| x_i^{(cancer)} - 1 \right| \right) \left(1 - p_i^{(cancer)} \right) \right\rangle \quad , \quad where: \tag{4}$$

$p_i^{(cancer)}$ = p-val of the heteroscedastic *t*-test of the equlity of the mean expressions

The percentage of the regulated genes regards all regulated genes as equal contributors regardless of the fold-change and p-value. In contrast, WPR weights the genes by considering the expression ratio (x), the p-value of the t-test of equal expression, and the average expression level in the normal tissue ($\mu^{(normal)}$).

2.7. Expression Correlation

Pearson pair-wise correlation coefficient ρ_{ij} was computed for the $\log_2$ expressions of all $N(N-1)/2$ pairs that can be formed with the N adequately quantified distinct transcripts in all biological replicas. The correlation coefficient takes values from -1 to +1. Close to positive and negative unit ρ's indicate that the expression of one gene of the pair has strong synergistic or antagonistic consequences on the expression of the other, without specifying what gene comes first. Expressions of synergistic partners fluctuate in phase, those of antagonistic partners fluctuate in antiphase. We believe that strong synergism and antagonism occur when the expressing genes are linked in a functional pathway, providing the "transcriptomic stoichiometry" that rules the expression levels of the involved proteins [40]. Close to zero correlation coefficient means either that the two genes are independently expressed (not networked in any functional pathway) or (very unlikely) that their synergistic correlation in some pathways is balanced by their antagonistic correlation in other pathways.

2.8. Gene Commanding Height (GCH)

The Gene Commanding Height (GCH) was introduced by us [21, 22] to quantify the importance of each gene for the cell phenotype. We consider that the expression level of a critical gene for the cell phenotype should be under a stricter control/protection of the homeostatic mechanisms and therefore should have higher expression stability among biological replicas. The same gene should also have a major regulatory role by coordinating the expression of many other genes. If we are allowed a comparison, the most protected persons in the UK are the Queen and the Prime Minister. However, the Prime Minister not the Queen is the Master Regulator owing to the power of the office to oversee all major sectors in the UK policy and economy.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 June 2019 doi:10.20944/preprints201906.0198.v1

Peer-reviewed version available at Genes 2019, 10, 560; doi:10.3390/genes10080560

$$GCH_i = \exp\left(RES_i + \frac{\frac{1}{N-1}\sum_{j=1,\, j\neq i}^{N} \rho_{ij}^2}{\underbrace{\frac{1}{N(N-1)}\sum_{k=1}^{N}\left(\sum_{j=1,\, j\neq k}^{N} \rho_{kj}^2\right)}_{\text{relative coordination power}}} \right) \simeq \frac{\langle REV \rangle}{REV_i} \times \exp\left(4\,\overline{\rho_{ij}^2}\Big|_{\forall j\neq i} \right) \quad (5)$$

2.9. Gene ontology and functional pathways

Whenever available through Gene Ontology Consortium [41, 42, www.geneontology.org] and/or Kyoto Encyclopedia of Genes and Genomes (KEGG, [43, http://www.genome.jp]) functional pathways are assigned to the genes. This can be done either by using the KEGG Ontology number as in GAEV [44], (https://github.com/UtaDaphniaLab/Gene_Annotation_Easy_Viewer), either by the gene symbol as in our GMR-Pathway (see below).

In this report we considered the following KEGG pathways: APO (map hsa4210 apoptosis), BTF (hsa03022 basal transcription factors), CCY (hsa04110 cell cycle), CSP (hsa04062 chemokine signaling), OXP (hsa00190 oxidative phosphorylation), RCC (hsa05211 renal cell carcinoma) and RPO (hsa03020 RNA polymerase).

2.10. CANCER-GMR software

Programs of our CANCER-GMR software package were designed using the Anaconda distribution of Python 3 with statistical and graphical user interface packages such as SciPy (https://www.scipy.org/) and Tkinter GUI (Graphical User Interface). The GMR Software Package includes executable programs to determine the absolute fold-change cut-offs (**CUT**) when comparing gene expression average levels in cancer and healthy tissues, identify functional pathways within KEGG dbase (**PATHWAY**), the Weighted Pathway Regulation (**WPR**), the Pearson correlation coefficients between the expression levels of all gene pairs (**CORRELATION**) and the Gene Commanding Height (**GCH**).

```
#CUT#
REV1 = CORRECTION * std1/mean1
REVV2 = CORRECTION * std2/mean2
CUT = 1 + np.sqrt( 2 * ( np.square(df["REV1"]) + np.square(df["REV2"]) ) )

#PATHWAY#
gene = REST.kegg_get(species+':'+gene).read()
### find all pathways in KEGG
if current_section == "PATHWAY":
gene_identifiers= line[12:].split("; ") #Splits each line based on ';'

#WPR#
avg_gch = avg_gch / length #avg of GCH of the genes in a certain pathway
wpr = np.mean(avg_gch * (abs(fc)-1) * (1-p_val))

#CORRELATION#
df = pd.read_csv(df,  header=0, na_values = "NaN")
data = pd.concat([df[condition_input+"1"],df[condition_input+"2"],
df[condition_input+"3"], df[condition_input+"4"]], axis=1)
logvalues =  np.log2(data)
results = (logvalues.T).corr(method='pearson') #transpose and find correlation
pearsons_df = pd.DataFrame(results.values, columns = df['GeneName'], index =
df['GeneName'].values) ## change to dataframe
```

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 June 2019 doi:10.20944/preprints201906.0198.v1

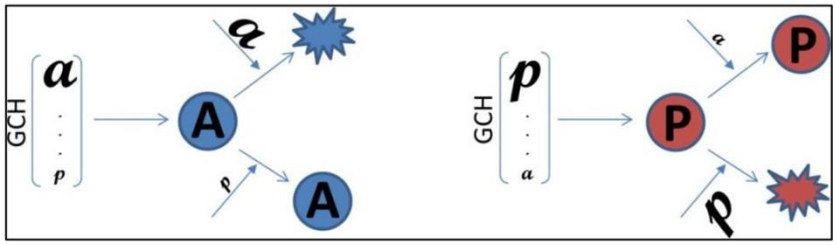
Peer-reviewed version available at Genes 2019, 10, 560; doi:10.3390/genes10080560


```
#GENE COMMANDING HEIGHT#
expcon = median / cov #expression control
results = (logvalues.T).corr(method='pearson') #transpose and find correlation
thesum = (results * results).sum(axis = 0)
### gene commanding height calculations
controlcoor = np.exp((4*thesum-1)/len(pearsons_df.index)-1)
gch = expcon.values * np.exp((4*thesum-1)/len(pearsons_df.index)-1).values
    gch = pd.DataFrame(gch, dtype='float')
```

2.11. Experimental design to validate the GMR theory

One way to validate the GMR theory is to establish the GCH hierarchy in two cell lines, transfect each line with genes having same expression level but different GCHs and compare the transcriptomic alterations (Fig.1A). The theory is validated if the alterations of the same gene transfection are higher in the cells where that gene has a higher GCH. We tested the usefulness of the GMR approach for cancer gene therapy by stable lentiviral transfecting the human BCPAP (papillary) and 8505C (anaplastic) thyroid cancer cell lines with four genes NEMP1, PANK2, DDX19B and UBALD1. As presented in Fig.1B, the selected genes have similar average expression levels (AVE, normalized to the median of all quantified transcripts) but significantly different GCH scores in the two lines.

A



B

GENE	DESCRIPTION	CELLS			
		BCPAP		8505C	
		AVE	GCH	AVE	GCH
NEMP1	nuclear envelope integral membrane protein 1	1.78	31.18	1.70	2.07
PANK2	pantothenate kinase 2, transcript variant 1	7.03	44.63	7.11	1.30
DDX19B	DEAD (Asp-Glu-Ala-Asp) box polypeptide 19B	1.72	0.82	1.55	15.61
UBALD1	UBA-like domain containing 1	1.79	1.12	1.26	10.31

Figure 1: A. Experimental design to validate the GMR theory. a is a gene with higher GCH in the **A**-cells (e.g.: 8505C anaplastic thyroid cancer cell line) than in the **P**-cells (e.g.: BCPAP papillary thyroid cancer cell line), while **p** is a gene with higher GCH in the **P**-cells than in the **A**-cells. The theory is verified if transfection of a will induce significantly larger WPR in the **A**-cells than in the **P**-cells and transfection of **p** will induce larger WPR in the **P**-calls than in the **A**-cells.
B. The genes used to test the GMR Theoryexpression data from GSE72304, [22]. The grey background indicates the larger GCH scores.

3. Results and Discussion

3.1. Experimental data

The GMR approach is illustrated here with applications to our gene expression datasets from surgically removed human (kidney, thyroid, prostate) cancer tissues and commercially available human cancer cell lines deposited by us in the Gene Expression Omnibus of the National Center for

Biotechnology Information (www.ncbi.nlm.nih.gov/gds). Data used in this report were from GSE72304 (a case of metastatic clear cell renal cell carcinoma (CCRCC), GSE97001 (a case of papillary thyroid cancer), GSE97002 (BCPAP papillary and 8505C anaplastic thyroid cancer cell lines) and from two (GSE to be provided prior to publication) cases of prostate cancer. Expression data for the GMR Theory validation on thyroid cancer cell lines BCPAP and 8505C were collected from GSE97031 (transfection with NEMP1), GSE97028 (DDX19B), GSE97030 (PANK2) and GSE97427 (UBALD1). Other expression data from cancer cell lines were collected from GSE72333 (DU145), GSE72414 (LNCaP) and GSE72415 (HL-60).

3.2. Expression stability, expression correlation and weighted pathway regulation

The Relative Expression Stability (RES) can be used not only to establish the hierarchy of individual genes but also the hierarchy of functional pathways. Fig. 2A presents the average RES scores of several pathway analyzed in each of the four regions profiled from the CCRCC samples. The averages were determined for 107 APO, 37 BTF, 131 CSP, 91 CCY, 100 OXP, 54 RCC and 31 RPO genes out of the 12610 distinct genes whose expression was adequately quantified in all four regions of the CCRCC samples. WPR analysis of the CCRCC samples for the same pathways returned the results from Fig.2B. Note that (by far) the most affected pathway was (as expected) RCC (the renal cell carcinoma) followed by OXP (oxidative phosphorylation).

Fig.2C illustrates the correlation analysis with examples of synergistically, antagonistically and independently expressed partners of *NEMP1* (nuclear envelope integral membrane protein 1) in the BCPAP cells. In our opinion, the strong correlations of the (not yet assigned to a pathway) *NEMP1* gene with the oncogenes *TFG* and *HRAS* indicate its potential role in the papillary thyroid cancer. *HRAS* is among the most documented genes whose mutations have been associated with thyroid cancer [45]. *TFG* gene is described in https://www.ncbi.nlm.nih.gov/gene/10342 as partially encoding several fusion oncoproteins and participating in several "oncogenic rearrangements resulting in anaplastic lymphoma and mixoid chondrosarcoma". TFG-MET (MET proto-oncogene receptor tyrosine-kinase) translocation was reported in a follicular variant of the papillary thyroid carcinoma [46]. Interestingly, NEMP1 was recently shown as promoting tamoxifen resistance in breast cancer cells [47]. Thus, the correlation analysis may be used to refine the maps of the functional pathways by determining the gene pairs whose correlated expression may result from a functional relationship between their encoded proteins.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 June 2019 doi:10.20944/preprints201906.0198.v1

Peer-reviewed version available at Genes 2019, 10, 560; doi:10.3390/genes10080560

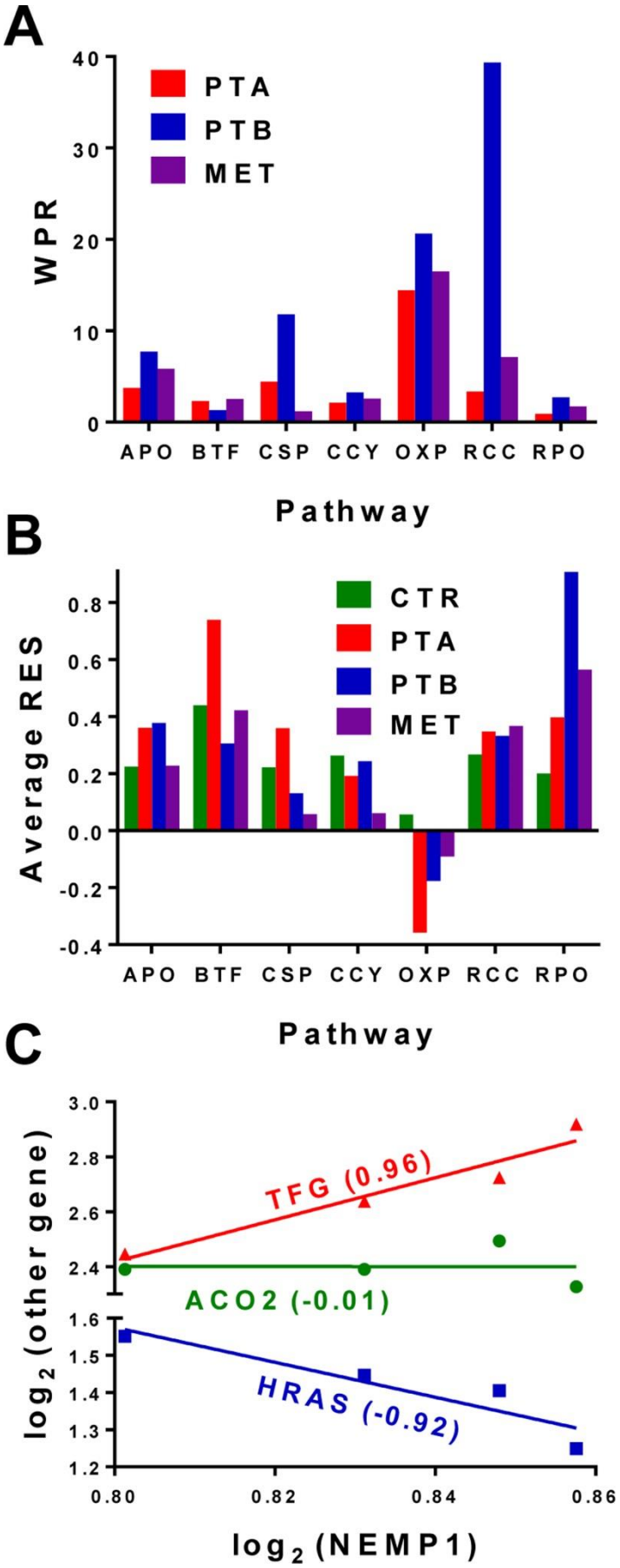


Figure 2reprocessed expression data from [21]**: A. Average Relative Expression Stability (RES) of several functional pathways in the four regions profiled from CCRCC samples.** Note that the RNA

polymerase (RPO) and basal transcription factors (BTF) are the most protected pathways, while the control of the oxidative phosphorylation (OXP) is relaxed, presumably to allow the cells to adapt the environmental conditions. **B. Weighted Pathway Regulation (WPR) analysis of several functional pathways.** Pathways: APO = apoptosis, BTF = basal transcription factors, CCY = cell cycle, CSP = chemokine signaling, OXP = oxidative phosphorylation, RCC = renal cell carcinoma, RPO = RNA polymerase. **C. Example of the correlation analysis.** NEMP1 is synergistically expressed with TFG (trafficking from ER to golgi regulator), antagonistically expressed with HRAS (Harvey rat sarcoma viral oncogene homolog) and independently expressed with ACO2 (aconitase 2, mitochondrial) in BCPAP cells.

3.3. Cancer nuclei and surrounding normal tissue are governed by distinct GMRs

Tables 1-3 present the GCH scores of the top 3 genes in cancer nuclei and surrounding normal tissue in a case of metastatic clear cell renal cell carcinoma, a case of papillary thyroid cancer and two cases of prostate cancer.

GENE	DESCRIPTION	CHR	CTR	PTA	PTB	MET
DAPK3	death-associated protein kinase 3	19	**30.31**	4.73	1.15	2.52
PMPCA	peptidase (mitochondrial processing) alpha	9	**28.35**	6.82	3.24	4.26
COA1	cytochrome c oxidase assembly factor 1 homolog	7	**22.40**	4.83	3.94	1.42
TASOR	transcription activation suppressor	3	3.08	**63.97**	1.59	5.40
BCR	breakpoint cluster region	22	1.15	**57.43**	1.14	1.22
C2orf81	chromosome 2 open reading frame 81	2	2.24	**51.24**	3.19	1.84
FAM27C	family with sequence similarity 27, member C	9	1.75	6.03	**57.19**	3.73
GTPBP3	GTP binding protein 3 (mitochondrial)	19	2.07	29.80	**40.06**	14.01
MIR1915HG	MIR1915 host gene (AKA: CASC 10 = cancer susceptibility candidate 10)	10	2.57	5.55	**31.14**	4.06
ALG13	ALG13, UDP-N-acetylglucosaminyltransferase subunit	X	3.64	9.97	2.12	**82.95**
NUDT18	nudix (nucleoside diphosphate linked moiety X)-type motif 18	8	1.64	2.69	1.89	**48.40**
RAD54B	RAD54 homolog B (S. cerevisiae)	8	0.96	6.10	4.09	**40.02**

Table 1reprocessed expression data from GSE72304, [21]: **Gene Commanding Heights of the top three genes (grey background) in the two primary tumor (PTA, PTB) regions from the right kidney and chest wall (MET) of a patient with metastatic clear cell renal cell carcinoma and their GCH scores in the other regions from the analyzed CCRCC sample. CHR** = chromosomal location. Note that, although different from one cancer region to the other, the top genes of cancer nuclei (PTA, PTB and MET) have substantially lower GCH scores in the control (normal, CTR) tissue and that the top genes of CTR region have lower GCH scores in the PTA, PTB and MET regions.

In Table 1, only *Alg13* is an actionable GMR (for MET region) owing to its significantly higher GCH (82.95) with respect to the second gene, *NUDT18* (GCH = 48.40). Interestingly, ALG13 (an early target of miR-34a) was reported as correlated with worse clinical outcomes for neuroblastoma [48]. Results in Table 1 also indicate that distinct cancer nuclei of the same tumor (PTA, PTB) may have distinct gene hierarchy that explains their phenotypic diversity.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 June 2019 doi:10.20944/preprints201906.0198.v1

Peer-reviewed version available at Genes 2019, 10, 560; doi:10.3390/genes10080560

GENE	DESCRIPTION	CHR	NORM	PAP-C	BCPAP	8505C
RASD1	RAS, dexamethasone-induced 1	17	41.51	4.50	5.70	7.31
POTEF	POTE ankyrin domain family, member F	2	31.17	8.50	6.90	6.36
RCN2	reticulocalbin 2, EF-hand calcium binding domain	15	31.09	5.53	7.99	10.38
SPINT2	serine peptidase inhibitor, Kunitz type, 2	19	1.93	54.97	18.83	5.88
RPAP3	RNA polymerase II associated protein 3	12	5.33	51.74	3.25	12.69
BZW1	basic leucine zipper and W2 domains 1	2	2.67	44.32	12.77	26.73
RPF1	ribosome production factor 1 homolog	1	8.36	2.22	135.50	22.11
TIMP2	TIMP metallopeptidase inhibitor 2	17	2.68	6.36	110.45	18.04
ECT2	epithelial cell transforming 2	3	6.93	8.16	100.98	28.15
SENP5	SUMO1/sentrin specific peptidase 5	3	9.93	6.32	100.37	13.71
RPL13A	ribosomal protein L13a	19	13.16	8.73	63.26	83.02
ALDOA	aldolase A, fructose-bisphosphate	16	7.00	28.05	2.59	67.30
TIPIN	TIMELESS interacting protein	15	3.11	9.15	37.04	56.85

Table 2reprocessed expression data from GSE97001, GSE97002, [22]: **Gene Commanding Heights of the top genes (grey background) in normal (NORM) and cancer (PAP-C) regions of the unilateral tumor removed from a 33y old Asian female, with papillary thyroid cancer (pathological stage pT3NOMx).** For comparison, we present also the results for the standard BCPAP (papillary) and 8505C (anaplastic) human thyroid cancer cell lines. Owing to their close (and over 100) GCHs, four instead of three genes are listed for the BCPAP cells. Note the differences in GCH scores between NORM and PAP-C. Note also the differences between the GCH scores of genes in the surgically removed carcinoma and the BCPAP cell line, even both are reported as papillary thyroid cancers.

In a previous publication [22], we have presented the GCH scores of 78 cancer biomarkers, 44 oncogenes, 55 apoptosis genes and 120 ncRNAs in the cancer and normal areas of a surgically removed papillary thyroid tumor. In that selection of 297 genes, all but *RAB15* ($GCH^{(cancer)}$ = 26.14) had GCH scores below 20. The complete GCH analysis of the expression data in the same tumor sample revealed that the GMR of the cancer area is *SPINT2* ($GCH^{(cancer)}$ = 54.97). Interestingly, SPINT2 a transmembrane protein that inhibits serine proteases implicated in cancer progression [49], acts as a putative tumor suppressor when hypermethylated [50].

GENE	DESCRIPTION	CHR	NORM1	CANCER1	NORM2	CANCER2	LNCaP	DU145
TOR1A	torsin family 1, member A	9	84.24	1.91	3.27	10.94	6.57	16.47
MRPS12	mitochondrial ribosomal protein S12	19	80.71	4.09	3.50	3.04	12.16	15.71
GTF2H1	general transcription factor IIH, polypeptide 1	11	42.66	5.71	5.27	5.83	4.34	17.34
BAIAP2L1	BAI1-associated protein 2-like 1	7	2.06	49.38	0.86	2.56	3.72	15.95
FAM71E1	family with sequence similarity 71, member E1	19	0.93	48.21	1.08	4.49	3.59	16.26
MAP6D1	MAP6 domain containing 1	3	1.29	45.26	1.34	2.05	7.50	16.61
SFR1	SWI5-dependent recombination repair 1	10	2.66	1.36	40.10	4.64	5.01	17.10
EDF1	endothelial differentiation-related factor 1	9	2.52	1.86	29.51	5.75	5.18	17.25
RHOD	ras homolog family member D	11	1.25	1.15	27.90	2.50	3.92	14.89
LOC145474	uncharacterized long non-coding RNA	14	2.23	1.64	1.16	126.75	1.01	12.33
PRRG1	proline rich Gla (G-carboxyglutamic acid) 1	X	N/A	N/A	1.82	87.53	5.10	14.33
ASAP3	ArfGAP with SH3 domain, ankyrin repeat and PH domain 3	1	1.19	1.73	1.97	76.23	4.15	16.28
WFDC3	WAP four-disulfide core domain 3	20	3.57	1.90	1.33	11.61	173.58	15.89
RPL31	60S ribosomal protein L31	2	1.30	0.94	2.19	8.25	39.11	18.16
ALX4	ALX homeobox 4	11	N/A	N/A	N/A	6.32	35.18	18.16
VIM	vimentin	10	1.28	1.97	N/A	2.85	3.51	33.95
POTEM	POTE ankyrin domain family, member M	14	1.18	2.40	0.54	4.18	2.51	33.25
EXOC5	exocyst complex component 5	14	1.32	1.32	1.03	5.28	2.75	32.19

392

393 **Table 3** data from GSE… to be communicated prior to publication**: Gene Commanding Heights of the top three genes**
394 **(grey background) in normal (NORM) and cancerous (CANCER) regions of surgically removed**
395 **prostate tumors from a 65y old black male and of a 47y old white male, both with prostatic**
396 **adenocarcinoma (Gleason score 4+5=9/10) and negative for adenocarcinoma resection margins.**
397 Note the differences in GCH scores between NORMAL and CANCER. Note that the two men have
398 different GMRs in both normal and cancer regions. For comparison, we show also the GCHs of the
399 same genes (and of their own GMRs) for two standard prostate cancer cell lines: the
400 androgen-sensitive LNCaP and the hormone insensitive DU145. Note also the large GCH gap
401 between the first two genes (*WFDC3* and *RPL31*) in the LNCaP cells.

402

403 It is notable that *WFDC3*, the far above GMR of the androgen-sensitive prostate cancer LNCaP
404 cells, was also reported as one of the most down-regulated gene in the ventral prostate of aged (18
405 moths) estrogen receptor $\beta^{-/-}$ mouse [51]. Interestingly, the GMR of the cancer nucleus of the second
406 man (*LOC145474*) is a non-coding RNA, confirming that both coding and non-coding RNAs may
407 play dominant roles in prostate tumorigenesis [52]. To our knowledge, this is the first time that
408 *LOC145474* is reported as related to a prostate cancer.

409

410 *3.4. Experimental validation of the GMR theory*

411 Our experimental results (summarized in Fig. 3) on the 8505C (anaplastic, 3A & 3C) and BCPAP
412 (papillary, 3B & 3D) human thyroid cancer cell lines stably transfected with DDX19B, NEMP1,
413 PANK2 or UBALD1 (characteristics in Fig.1) indicate that:

414

$$\begin{aligned}
GCH_{NEMP1}^{(BCPAP)} &> GCH_{NEMP1}^{(8505C)} \Rightarrow WPR_{NEMP1}^{(BCPAP)} > WPR_{NEMP1}^{(8505C)} \\
GCH_{PANK2}^{(BCPAP)} &> GCH_{PANK2}^{(8505C)} \Rightarrow WPR_{PANK2}^{(BCPAP)} > WPR_{PANK2}^{(8505C)} \\
GCH_{DDX19B}^{(BCPAP)} &< GCH_{DDX19B}^{(8505C)} \Rightarrow WPR_{DDX19B}^{(BCPAP)} < WPR_{DDX19B}^{(8505C)} \\
GCH_{UBALD1}^{(BCPAP)} &< GCH_{UBALD1}^{(8505C)} \Rightarrow WPR_{UBALD1}^{(BCPAP)} < WPR_{UBALD1}^{(8505C)}
\end{aligned} \quad (6)$$

415

416

417

418 We have also observed that transfections of NEMP1 and PANK2 significantly slowed down
419 multiplication of BCPAP cells, while transfection of DDX19B or UBALD1 had little effect on these
420 cells. By contrast, both transfections of DDX19B and UBALD1 significantly slowed down grow of
421 8505C cells, while transfection of NEMP1 and PANK2 had little effect. Together, these observations
422 confirms that expression manipulation of a gene has larger consequences in cells where that gene
423 has higher GCH.

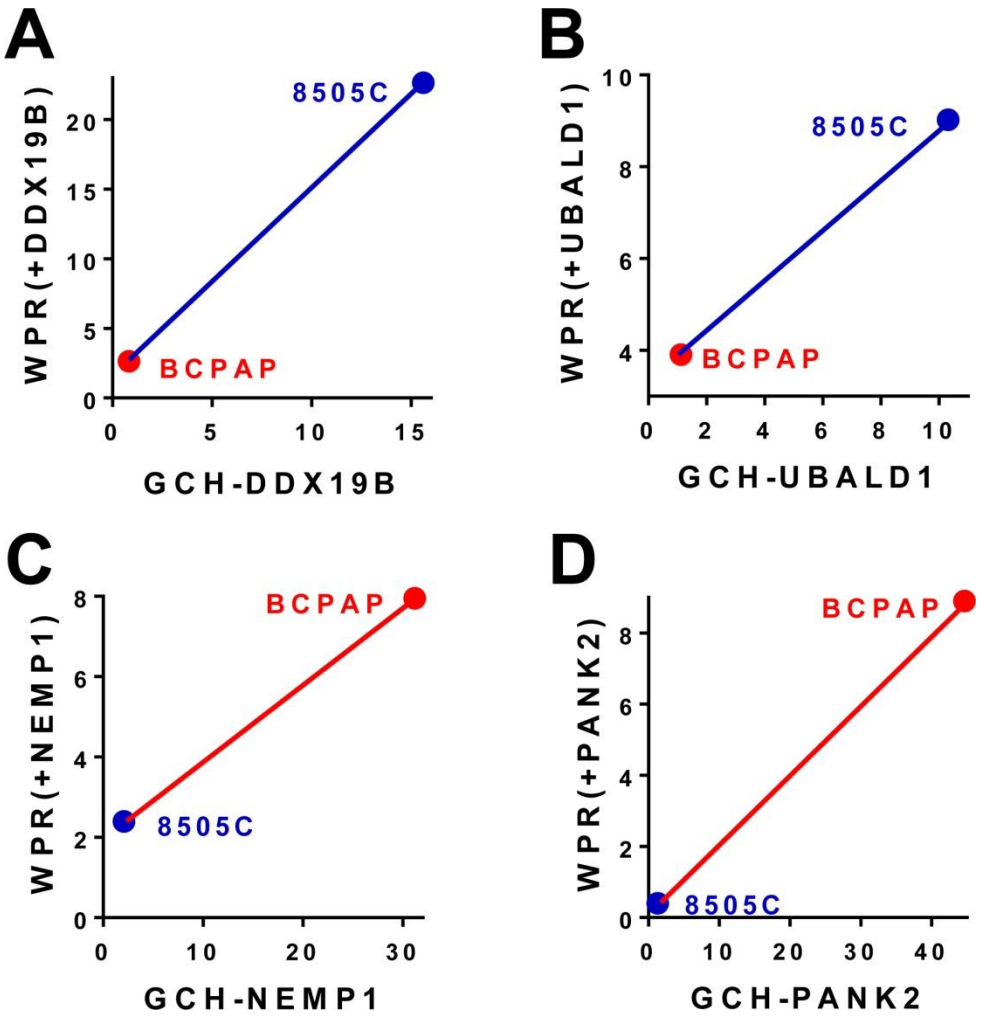


Figure 3expression data from GSE97031, GSE97028, GSE97030, GSE97427, [22]**: Validation of the GMR Theory. A & B.** Stable transfection of genes with higher GCH in the 8505C cells than in BCAP cells had larger transcriptomic effects in 8505C cells as measured by the Weighted Pathway Regulation (WPR). **C & D.** Stable transfection of genes with higher GCH in the BCPAP cells than in 8505C cells had larger transcriptomic effects in BCPAP cells as measured by the Weighted Pathway Regulation (WPR).

3.5. Predicted transcriptomic alteration by GMR manipulation

At the time, we had no possibility to alter the expression of the GMRs identified in the thyroid cancer cell lines 8505C and BCPAP. Fig. 4 presents the predicted Weighted Pathway Regulation if significantly altering the expressions of the top three genes in the 8505C and BCPAP cells.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 June 2019 doi:10.20944/preprints201906.0198.v1

Peer-reviewed version available at Genes 2019, 10, 560; doi:10.3390/genes10080560

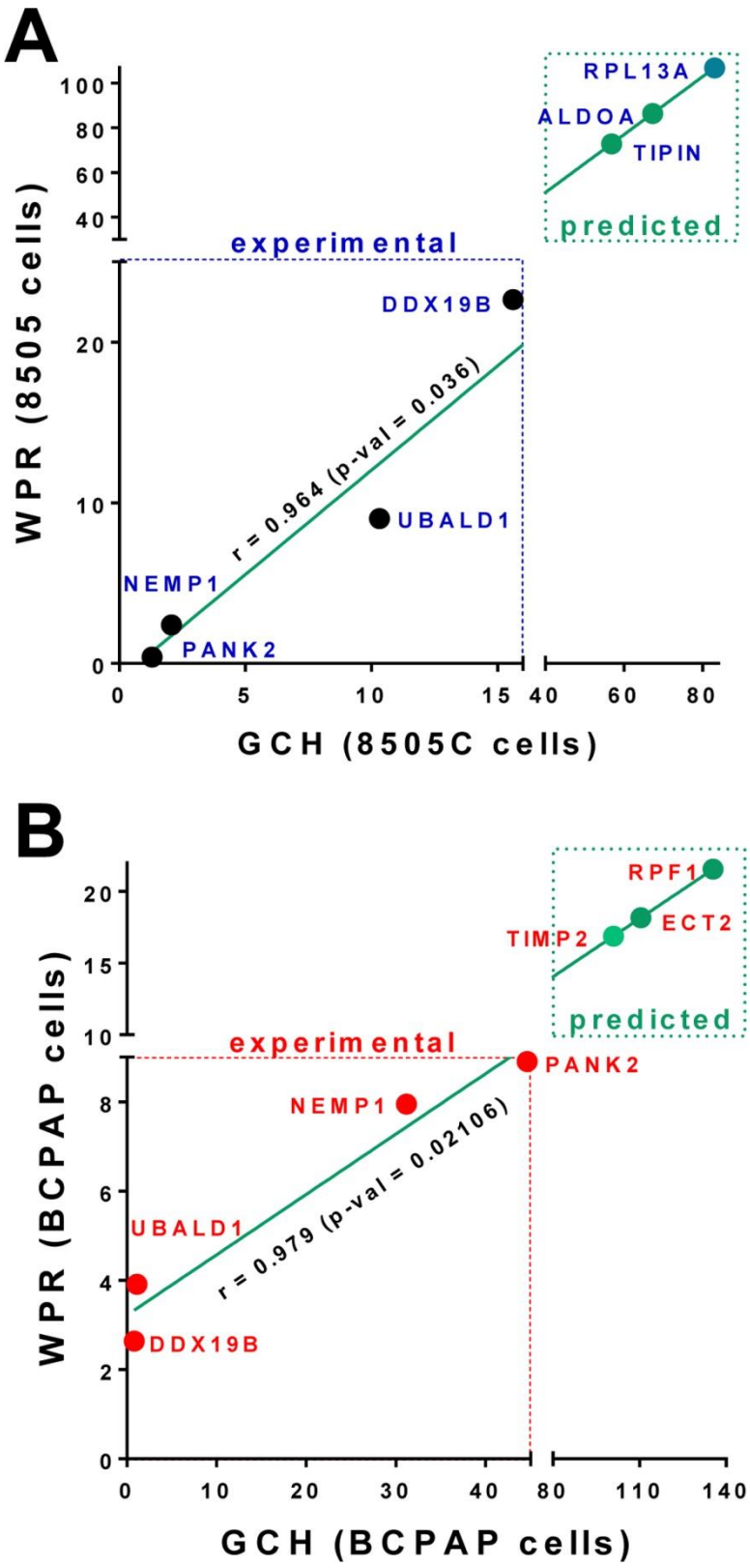


Figure 4: A. Experimentally measured and theoretically predicted effects of stably transfecting PANK2, NEMP1, UBALD1, DDX19B and the top three genes (RPL13A, ALDOA and TIPIN) in the 8505C. **B.** Experimentally measured and theoretically predicted effects of stably transfecting PANK2, NEMP1, UBALD1, DDX19B and the top three genes (RPF1, ECT2 and TIMP2) in the BCPAP cells.

Interestingly, the GMR of the anaplastic cell line (*RPL13A*) was found as the most stably expressed gene in two ovarian cancer cell lines (UACC-1598 and SKOV3) subjected to two widely used

anticancer treatment [53], confirming the major role played by this gene in stabilizing the cancer phenotype.

3.6. Ribosomal genes top the hierarchy in the acute promyelocytic leukemia HL-60 cell line

An interesting gene hierarchy was obtained for the HL-60 cells (Fig. 5), where most of the ribosomal genes top all KEGG identified genes as associated with acute myeloid leukemia (AML, https://www.kegg.jp/kegg-bin/show_pathway?map=has05221&show_description=show). Thus, the top two AML genes, *MPO* (myeloperoxidase) and *PML* (promyelocytic leukemia) have the GCHs 22.96 and 21.95, below those of the 56th and the 59th ranked ribosomal proteins RPL29 (23.64) and RPL13 (22.12). As presented in Fig. 5C, genes from both large (RPL) and small (RPS) ribosomal subunits are among the highest ranked genes in the HL-60 cells. According to our results, certain ribosomal genes (RPL13A, RPS5) are more influential in dictating/preserving the HL-60 phenotype than *RARA* (retinoic acid receptor alpha, GCH = 19.67), the gene whose translocation t(15;17) [53,54] is associated with 98% of acute promyelocytic leukemia cases. Interestingly, *RPL13A* was also found as the most influential gene in 8505C (anaplastic thyroid cancer) cells and with high GCH (63.26) in the BCPAP cells (Table 2 above). For comparison, Fig. 5 presents also the GCH scores for the genes associated with the immune system KEGG pathway of leukocyte transendothelial migration (https://www.kegg.jp/kegg-bin/show_pathway?map=hsa04670&show_description=show).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 June 2019 doi:10.20944/preprints201906.0198.v1

Peer-reviewed version available at Genes 2019, 10, 560; doi:10.3390/genes10080560

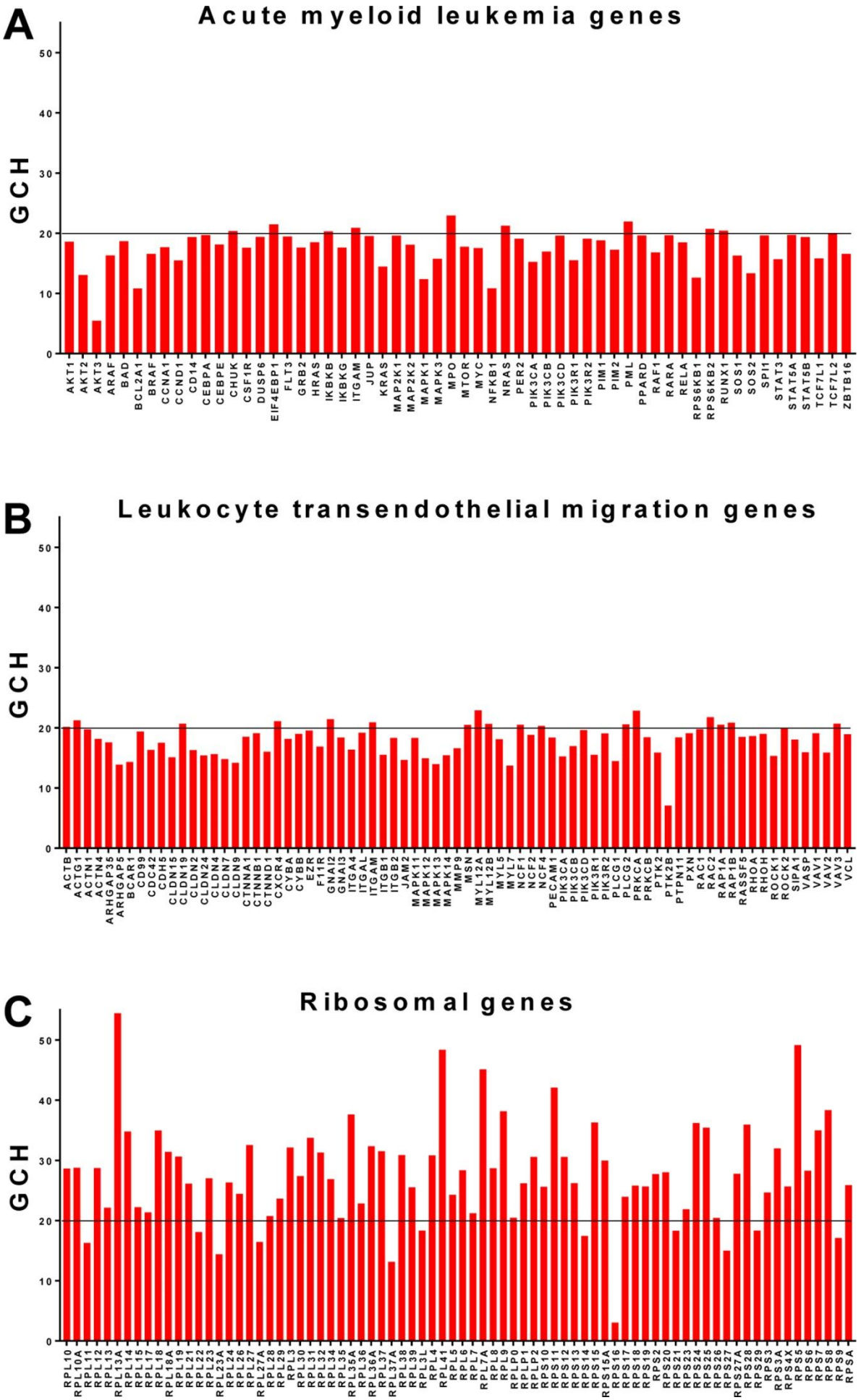


Figure 5expression data from GSE72415**: Gene Commanding Height (GCH) scores in the acute promyelocytic leukemia cell line HL-60. A. GCH scores of genes associated with acute myeloid leukemia. B. GHC scores of genes associated with KEGG pathway of leukocyte transendothelial migration. C. GCH scores of the ribosomal genes.**

4. Conclusions

In this report, we proposed that the most legitimate targets for the cancer gene therapy are the genes whose highly controlled expressions by the homeostatic mechanisms are the most influential by being correlated with expressions of many other genes. We termed these targets Gene Master Regulators (GMRs), and developed and used the necessary experimental protocol, mathematical algorithm and computer software to identify them from gene expression studies.

The GMR approach was applied to our microarray data on five standard (anaplastic thyroid, papillary thyroid, prostate androgen-sensitive and insensitive, and blood) cancer cell lines and ten profiled regions from surgically removed CCRCC (4), and prostate (4) and papillary thyroid (2) cancers. The studies revealed that the GMRs may differ even for patients with the same form of cancer (like in the above two prostate cases and in comparisons with cell lines), justifying the necessity of a personalized approach of cancer gene therapy. We found that the GMRs can be located in any chromosome, that their transcripts can be both coding and non-coding RNAs and that the encoded proteins may be involved in a wide diversity of biological processes. Although we don't have yet the experimental evidence, most likely the gene hierarchy changes in time, so that the GMRs should be targeted as soon as possible after their identification.

Importantly, we found that the cancer nuclei and the surrounding normal tissues are governed by different GMRs and that manipulation of the expression of a gene has consequences in line with its Gene Commanding Height. Based on these findings, we launch the hypothesis that silencing the GMR (using CRISPR or shRNA) may selectively kill the cancer cells with little effect on the normal cells of the tissue. However, not always we found the GMR being well above the rest of the genes, with the most notable exception in this report for WFDC3 (GCH = 173.58) in LNCaP cells (next gene is RPL31 at 39.11). Therefore, this hypothesis may work for only cases where the GMR has a significantly dominant GCH over the other genes in the cancer nucleus and a very low GCH in the normal cells of the tissue.

Author Contributions: Conceptualization, D.A.I.; methodology, D.A.I., S.I.; software, N.E.; validation, S.I., D.A.I.; formal analysis, S.I., D.A.I.; investigation, S.I., D.A.I.; resources, D.A.I.; data curation, D.A.I.; writing—original draft preparation, D.A.I.; writing—review and editing, D.A.I.; visualization, D.A.I.; supervision, D.A.I.; project administration, D.A.I.; funding acquisition, D.A.I.

Funding: This research was funded in part by the Texas A&M University System Chancellor's Research Initiative and by the PVAMU Undergraduate scholarly research award.

Conflicts of Interest: The authors declare no conflict of interest. the results".

References

1. Erstad, D.J.; Fuchs, B.C.; Tanabe, K.K. Molecular signatures in hepatocellular carcinoma: A step toward rationally designed cancer therapy. *Cancer* **2018**, 124(15):3084-3104. doi: 10.1002/cncr.31257.
2. Kretschmer, A.; Tilki, D. Biomarkers in prostate cancer - Current clinical utility and future perspectives. *Crit Rev Oncol Hematol*. **2017**, 120:180-193. doi: 10.1016/j.critrevonc.2017.11.007.
3. Lam, M.; Roszik, J.; Kanikarla-Marie, P.; Davis, J.S.; Morris, J.; Kopetz, S.; et al. The potential role of platelets in the consensus molecular subtypes of colorectal cancer. *Cancer Metastasis Rev* **2017**, 36(2):273-288. doi: 10.1007/s10555-017-9678-9.
4. Sacco, A.; Fenotti, A.; Affò, L.; Bazzana, S.; Russo, D.; Presta, M.; et al. The importance of the genomic landscape in Waldenström's Macroglobulinemia for targeted therapeutical interventions. *Oncotarget* **2017**, 8(21):35435-35444.
5. Santosh, A.B.; Jones, T.; Harvey, J. A review on oral cancer biomarkers: Understanding the past and learning from the present. *J Cancer Res Ther* **2016**, 12(2):486-92. doi: 10.4103/0973-1482.176414

6. The Cancer Genome Atlas (TCGA) Research Network et al. The Cancer Genome Atlas Pan-Cancer analysis project. *Nat Genet* **2013,** 45, 1113–1120. doi: 10.1038/ng.2764
7. Uhlen, M.; Zhang, C.; Lee, S.; Sjöstedt, E.; Fagerberg, L.; Bidkhori, G.; et al. A pathology atlas of the human cancer transcriptome. *Science* **2017,** 357(6352). pii: eaan2507. doi: 10.1126/science.aan2507
8. Iacobas, S.; Iacobas, D.A.; Spray, D.C.; Scemes, E. The connexin43 transcriptome during brain development: importance of genetic background. *Brain Research* **2012,** 1487: 131-139. doi: 10.1016/j.brainres.2012.05.062.
9. Iacobas, D.A.; Iacobas, S.; Thomas, N.; Spray, D.C. Sex-dependent gene regulatory networks of the heart rhythm. *Funct Integr Genomics* **2010**, 10(1):73-86. doi: 10.1007/s10142-009-0137-8.
10. Iacobas, D.A.; Fan, C.; Iacobas, S.; Spray, D.C.; Haddad, G.G. Transcriptomic changes in developing kidney exposed to chronic hypoxia. *Biochem Biophys Res Comm* **2006**, 349(1), 329-338. DOI:10.1016/j.bbrc.2006.08.056.
11. Iacobas, D.A.; Fan, C.; Iacobas, S.; Haddad, G.G. Integrated transcriptomic response to cardiac chronic hypoxia: translation regulators and response to stress in cell survival. *Funct Integr Genomics* **2008**, 8(3):265-75. doi: 10.1007/s10142-008-0082-y.
12. Kobets, T.; Iatropoulos, M.J.; Duan, J.D.; Brunnemann, K.D.; Iacobas, D.A.; Iacobas, S.; et al. Effects of Nitrosamines on the Expression of Genes Involved in Xenobiotic Metabolism in the Chicken Egg Alternative Genotoxicity Model. *Toxicol Sci* **2018**; 166(1), 82–96, doi: 10.1093/toxsci/kfy197.
13. Iacobas, D.A.; Chachua, T.; Iacobas, S.; Benson, M.J.; Borges, K.; Veliskova, J.; Velisek, L. ACTH and PMX53 recover the normal synaptic transcriptome in a rat model of infantile spasms. *Sci Rep* **2018**, 8:5722, DOI:10.1038/s41598-018-24013-x.
14. Teng, H.; Mao, F.; Liang, J.; Xue, M.; Wei, W.; Li, X.; et al. Transcriptomic signature associated with carcinogenesis and aggressiveness of papillary thyroid carcinoma. *Theranostics* **2018**, 8(16):4345-4358. doi:10.7150/thno.26862.
15. Coghlin, C.; Murray, G.I. Biomarkers of colorectal cancer: Recent advances and future challenges. *Proteomics Clin Appl* **2015**, 9(1-2):64-71. doi: 10.1002/prca.201400082.
16. Vogelstein, B.; Papadopoulos, N.; Velculescu, V.E.; Zhou, S.; Diaz, L.A. Jr.; Kinzler, K.W. Cancer genome landscapes. *Science* **2013**, 339(6127):1546-58.5. doi: 10.1126/science.1235122.
17. Estevez-Garcia, P.; Rivera, F.; Molina-Pinelo, S.; Benavent, M.; Gómez, J.; Limón, M.L.; et al. Gene expression profile predictive of response to chemotherapy in metastatic colorectal cancer. *Oncotarget* **2015**, 6(8):6151-9. DOI:10.18632/oncotarget.3152
18. Li, H.; Samawi, H.; Heng, D.Y. The use of prognostic factors in metastatic renal cell carcinoma. *Urol Oncol* **2015**, 33(12):509-16 Review. doi: 10.1016/j.urolonc.2015.08.003
19. Martínez-Bosch, N.; Rodriguez-Vida, A.; Juanpere, N.; Lloreta, J.; Rovira, A.; Albanell, J.; et al. Galectins in prostate and bladder cancer: tumorigenic roles and clinical opportunities. *Nat Rev Urol* **2019**, [Epub ahead of print]. doi: 10.1038/s41585-019-0183-5.
20. Tang, K.; Xu, H. Prognostic value of meta-signature miRNAs in renal cell carcinoma: an integrated miRNA expression profiling analysis. *Sci Rep* **2015**, 14; 5:10272. doi: 10.1038/srep10272.
21. Iacobas, D.A.; Iacobas, S. Towards a personalized cancer gene therapy: a case of clear cell renal cell carcinoma. *Cancer & Oncol Res* **2017**, 5(3):45:52. DOI:10.13189/cor.2017.050301
22. Iacobas, D.A.; Tuli, N.; Iacobas, S.; Rasamny, J.K.; Moscatello, A.; Geliebter, J.; et al. Gene master regulators of papillary and anaplastic thyroid cancer phenotypes. *Oncotarget* **2018**, 9(2), 2410-2424. doi: 10.18632/oncotarget.23417
23. Clevers, H.C.; Owen, M.J. Towards a molecular understanding of T-cell differentiation. *Immunol Today* **1991**, 12(2):86-92. DOI: 10.1016/0167-5699(91)90163-N
24. Doevendans, P.A.; van Bilsen, M. Transcription factors and the cardiac gene programme. *Int J Biochem Cell Biol* **1996**, 28(4):387-403.
25. Liyanarachchi, S.; Li, W.; Yan, P.; Bundschuh, R.; Brock, P.; Senter, L.;, et al. Genome-Wide Expression Screening Discloses Long Noncoding RNAs Involved in Thyroid Carcinogenesis. *J Clin Endocrinol Metab* **2016**; 101(11):4005-4013. DOI: 10.1210/jc.2016-1991
26. Wang, Y.Y.; Lin, X.D.; Fu, X.H.; Yan, W.; Lin.; F.S.; Kuang, P.H.; et al. Long non-coding RNA BANCR regulates cancer stem cell markers in papillary thyroid cancer via the RAF/MEK/ERK signaling pathway. *ONCOLOGY REPORTS* **2018**; 40(2): 859-866, DOI: 10.3892/or.2018.6502

27. Braicu, C.; Zimta, A.A.; Harangus, A.; Iurca, I.; Irimie, A.; Coza, O.; Berindan-Neagoe, I. The Function of Non-Coding RNAs in Lung Cancer Tumorigenesis. *Cancers* (Basel). **2019**, 11(5). pii: E605. doi: 10.3390/cancers11050605. Review.
28. Braicu, C.; Gulei, D.; Cojocneanu, R.; Raduly, L.; Jurj, A.; Knutsen, E.; et al. miR-181a/b therapy in lung cancer: reality or myth? *Mol Oncol* **2019**; 13(1):9-25. doi: 10.1002/1878-0261.12420.
29. Iacobas, D.A.; Iacobas, S.; Haddad, G.G. Heart rhythm genomic fabric in hypoxia. *Biochem. Biophys. Res. Commun* **2010**, 391(4):1769-1774. doi: 10.1016/j.bbrc.2009.12.151
30. Iacobas, D.A. The Genomic Fabric Perspective on the transcriptome between universal quantifiers and personalized genomic medicine. *Biol Theory* **2016**, 11(3): 123-137. DOI 10.1007/s13752-016-0245-3.
31. Gallagher, R.; Collins, S.; Trujillo, J.; McCredie, K.; Ahearn, M.; Tsai, S.; et al. Characterization of the continuous, differentiating myeloid cell line (HL-60) from a patient with acute promyelocytic leukemia. *Blood* **1979**, 54(3): 713-33.
32. Corso, C.; Ulucan, H.; Parry, E.M.; Parry, J.M. Comparative analysis of two thyroid tumor cell lines by fluorescence in situ hybridization and comparative genomic hybridization. *Cancer Genet Cytogenet* **2002**, 137(2):108-18.
33. Ito, T.; Seyama, T.; Hayashi, Y.; Hayashi, T.; Dohi, K.; Mizuno, T.; et al. Establishment of 2 human thyroid-carcinoma cell-lines (8305c, 8505c) bearing p53 gene-mutations. *Int J Oncol* **1994**, 4(3):583-6. https://doi.org/10.3892/ijo.4.3.583.
34. Horoszewicz, J.S.; Leong, S.S.; Chu, T.M.; Wajsman, Z.L.; Friedman, M.; Papsidero, L; et al. The LNCaP cell line--a new model for studies on human prostatic carcinoma. *Prog Clin Biol Res* **1980**, 37:115-32.
35. Stone, K.R.; Mickey, D.D.; Wunderli, H.; Mickey, G.H.; Paulson, D.F. Isolation of a human prostate carcinoma cell line (DU 145). *Int. J. Cancer* **1978**, 21 (3): 274–81. doi:10.1002/ijc.2910210305.
36. Kravchick, D.O.; Hrdinka, M.; Iacobas, S.; Iacobas, D.A.; Kreutz, M.R.; Jordan, B.A. Synaptonuclear messenger PRR7 inhibits c-Jun ubiquitination and regulates NMDA mediated excitotoxicity. *EMBOJ* **2016**, 35(17):1923-34. doi: 10.15252/embj.201593070.
37. Iacobas, D.A.; Iacobas, S.; Urban-Maldonado, M.; Spray, D.C. Sensitivity of the brain transcriptome to connexin ablation, *Biochim Biophys Acta* **2005**, 1711: 183-196. Review. DOI: 10.1016/j.bbamem.2004.12.002.
38. Iacobas, D.A.; Iacobas, S.; Tanowitz, H.B.; deCarvalho, A.C.; Spray, D.C. Functional genomic fabrics are remodeled in a mouse model of Chagasic cardiomyopathy and restored following cell therapy. *Microbes Infect* **2018**, 20(3), 185-195. doi: 10.1016/j.micinf.2017.11.003
39. Iacobas, D.A.; Iacobas, S.; Nebieridze, N.; Velisek, L.; Veliskova, J. Estrogen protects neurotransmission transcriptome during status epilepticus. *Front Neurosci* **2018**, 12:232. doi: 10.3389/fnins.2018.00332.
40. Iacobas, D.A.; Iacobas, S.; Spray, D.C. Connexin43 and the brain transcriptome of the newborn mice. *Genomics* **2007,** 89(1), 113-123. DOI:10.1016/j.ygeno.2006.09.007.
41. Ashburner, M.; Ball, C.A.; Blake, J.A.; Botstein, D.; Butler, H.; Cherry, J.M.; et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* **2000**, 25(1):25-9.
42. The Gene Ontology Consortium. Expansion of the Gene Ontology knowledgebase and resources. *Nucleic Acids Res* **2017**, 45(D1):D331-D338.
43. Kanehisa, M.; Furumichi, M.; Tanabe, M.; Sato, Y.; Morishima, K. KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res* **2017**, 45, D353–D361
44. Huynh, T.; Xu, S. Gene Annotation Easy Viewer (GAEV): Integrating KEGG's Gene Function Annotations and Associated Molecular Pathways **2018** F1000Research, 7:416. doi: 10.12688/f1000research.14012.2.
45. Dou, R.; Zhang, L.; Lu, T.; Liu, D.; Mei, F.; Huang, J.; et al. Identification of a novel HRAS variant and its association with papillary thyroid carcinoma. *Oncol Lett* **2018**, 15(4):4511-4516. doi: 10.3892/ol.2018.7818.
46. Cipriani, N.A.; Agarwal, S.; Dias-Santagata, D.; Faquin, W.C.; Sadow, P.M. Clear Cell Change in Thyroid Carcinoma: A Clinicopathologic and Molecular Study with Identification of Variable Genetic Anomalies. *Thyroid* **2017**, 27(6):819-824. doi: 10.1089/thy.2016.0631.
47. Liu, Y.; Tong, C.; Cao, J.; Xiong, M. NEMP1 Promotes Tamoxifen Resistance in Breast Cancer Cells. *Biochem Genet* **2019,** doi: 10.1007/s10528-019-09926-0. [Epub ahead of print]

48. De Antonellis, P.; Carotenuto, M.; Vandenbussche, J.; De Vita, G.; Ferrucci, V.; Medaglia, C.; et al. Early targets of miR-34a in neuroblastoma. *Mol Cell Proteomics* **2014**, 13(8):2114-31. doi: 10.1074/mcp.M113.035808.
49. Pereira, M.S.; de Almeida, G.C.; Pinto, F.; Viana-Pereira, M.; Reis, R.M. SPINT2 Deregulation in Prostate Carcinoma, *J Histochem Cytochem* **2016**, 64(1):32-41. doi: 10.1369/0022155415612874
50. Liu, F.; Cox, C.D.; Chowdhury, R.;, Dovek, L.; Nguyen, H.; Li, T.; et al. SPINT2 is hypermethylated in both IDH1 mutated and wild-type glioblastomas, and exerts tumor suppression via reduction of c-Met activation. *J Neurooncol* **2019**, 142(3):423-434. doi: 10.1007/s11060-019-03126-x.
51. Wu, W.F.; Maneix, L.; Insunza, J.; Nalvarte, I.; Antonson, P.; Kere, J.; et al. Estrogen receptor β, a regulator of androgen receptor signaling in the mouse ventral prostate. *Proc Natl Acad Sci U S A*. **2017**, 114(19):E3816-E3822. doi: 10.1073/pnas.1702211114.
52. Wang, X.F.; Chen, J.; Gong, Y.B.; Qin, Y.C.; Wang, L.; Li, N.C. Long non-coding RNA DUXAP10 promotes the proliferation, migration, and inhibits apoptosis of prostate cancer cells. *Eur Rev Med Pharmacol Sci* **2019**, 23(9):3699-3708. doi: 10.26355/eurrev_201905_17793.
53. Bian, Z, Yu, Y.; Quan, C.; Guan, R.; Jin, Y.; Wu, J.; et al. RPL13A as a reference gene for normalizing mRNA transcription of ovarian cancer cells with paclitaxel and 10-hydroxycamptothecin treatments. *Mol Med Rep* **2015**, 11(4):3188-94. doi: 10.3892/mmr.2014.3108.
54. de Thé H, Lavau C, Marchio A, Chomienne C, Degos L, Dejean A. The PML-RAR alpha fusion mRNA generated by the t(15;17) translocation in acute promyelocytic leukemia encodes a functionally altered RAR. Cell. 1991; 66(4):675-84.
55. Zelent A, Guidez F, Melnick A, Waxman S, Licht JD. Translocations of the RARalpha gene in acute promyelocytic leukemia. Oncogene 2001; 20(49):7186-203. DOI: 10.1038/si.onc.1204766.