

Comprehensive transcriptomic analysis of Papillary Thyroid Cancer: potential biomarkers associated with tumor progression

Nazanin Hosseinkhan

Institute of Endocrinology and Metabolism, Iran University of Medical Sciences

Maryam Honardoost

Institute of Endocrinology and Metabolism, Iran University of Medical Sciences

Kevin Blighe

Biomedical Sciences Research Institute, Ulster University

C.B. Tara Moore

Biomedical Sciences Research Institute, Ulster University

Mohammad Ebrahim Khamseh (✉ khamseh.m@iums.ac.ir)

Institute of Endocrinology and Metabolism, Iran University of Medical Sciences <https://orcid.org/0000-0003-4313-8440>

Research article

Keywords: Papillary Thyroid Cancer, tumor progression, patient stratification, biomarkers

Posted Date: October 4th, 2019

DOI: <https://doi.org/10.21203/rs.2.15631/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at Journal of Endocrinological Investigation on January 21st, 2020. See the published version at <https://doi.org/10.1007/s40618-019-01175-7>.

Abstract

Background Identification of stage-specific prognostic/predictive biomarkers could lead to the more efficient clinical management in papillary thyroid carcinoma (PTC). The main objective of this study was to characterize the stage-specific deregulations in gene and miRNA expression in PTC and also to identify potential prognostic biomarkers.

Methods 495 RNASeq and 499 miRNASeq PTC samples (stage I-IV) as well as, respectively, 56 and 57 normal samples were retrieved from The Cancer Genome Atlas (TCGA). DESeq2 was used to identify deregulation of genes and miRNAs between sequential stages of PTC. To identify the minority of patients who progress from stage I to higher stages, we additionally performed a clustering analysis focused on the stage I RNASeq data. Moreover a validation study was done on an independent PTC RNASeq study.

Results Large amount of heterogeneity in gene expression patterns was observed in stage I PTC patients. LTF and PLA2R1 were identified as two promising biomarkers that exhibited down-regulation in a small subgroup of stage I (both in TCGA and in the validation data set) and in the majority of stage IV patients (in TCGA data set). hsa-miR-205, hsa-miR-509-2, hsa-miR-514-1 and hsa-miR-514-2 were also identified as up-regulated miRNAs in both PTC patients with stage I and stage III.

Conclusion Common gene expression alterations in early and advanced stages of PTC, could be used for individualized risk stratification and personalized treatment approach.

Background

The incidence of thyroid cancer has been increased worldwide during the past decades. According to Globocan 2018, thyroid cancer is in ninth place for incidence with 567,000 cases worldwide. It is 3 times more prevalent in women with the incidence rate of 10.2 per 100,000 ([Bray et al., 2018](#)). Four major types of thyroid cancer, including Papillary (PTC), Follicular (FTC), Medullary (MTC) and Anaplastic Thyroid Carcinoma (ATC) have already been characterized, among which PTC is the most frequent histotype, comprising 85–90 % of all thyroid cancer cases. Despite good prognosis, lymph node or distant metastasis is observed in approximately 10% of PTC patients. Although it is still controversial, it has been suggested that there are specific PTC subtypes related to the risk of loco regional recurrence or distant metastasis ([Katoh et al., 2015](#)).

Thyroid cancer staging provides prognostic information related to disease surveillance, risk stratification and therapeutic strategies. Accurate initial staging requires detailed data obtained from preoperative work up, intra operative findings and postoperative follow-up. Moreover risk stratification based on clinical and pathology risk factors or molecular profiling can be efficiently used to guide follow-up management decisions. In addition identification of at-risk patients is fundamental to suggest additional therapies or select a more conservative management approach.

According to the American Joint Committee on Cancer (AJCC) TNM staging system, based on the size of tumor (T), spread to nearby lymph nodes (N), and metastasis to distant sites (M), four stages for thyroid cancer (I-IV) can be specified. Moreover the staging system uses a combination of risk factors such as age at diagnosis, size of the primary tumor, specific tumor histology, direct extension of the tumor to the outside of thyroid gland, loco-regional metastases, and/or distant metastases to stratify patients into various risk categories with differing risks of death from thyroid cancer. PTC has already been subjected to the large-scale genetic characterization studies through which the genetic alterations associated with its tumorigenesis were partly uncovered (Agrawal et al., 2014, Haraldsdottir and Shah, 2014, Liu et al., 2018). It has been indicated that activation of mitogen-activated protein kinase (MAPK) and phosphoinositide 3-kinase AKT (PI3K-AKT) pathways through gene fusions or somatic mutations could be considered as primary alterations in PTC (Liu et al., 2018). Recently Tengs and colleagues, in a comprehensive transcriptomics study, characterized a signature of deregulated genes and lncRNAs associated with extra-thyroidal extension. Moreover they suggested several biomarkers with the potential use in the early detection of thyroid cancer (Teng et al., 2018)

In the present study we aimed to investigate the transcriptomic alterations between different PTC stages at both RNA and miRNA levels to identify gene specific signatures of different stage transitions. Moreover, since most PTC patients do not progress to the higher stages, we were interested to identify the subgroup of PTC patients in stage I with the potential to progress into the higher stages. Having this knowledge about different PTC high-risk subgroups could help to stratify patients based on the molecular signatures of tumor. This may lead to the adoption of a personalized approach for treating PTC patients and to reduce the financial burden of unnecessary treatment modalities.

Method

RNASeq and miRNASeq open access data on Thyroid Cancer Adenocarcinoma (PTC type) was retrieved from TCGA (<http://portal.gdc.cancer.gov/>). Tumor samples were categorized into four stages ("stage I-IV") according to the clinical data provided by TCGA (Table 1 and 2).

Differential Expression Analysis on RNASeq and miRNASeq data

495 Illumina Genome Analyzer RNASeq row counts (htseq-count) files of PTC in which the expression value of 60,483 transcripts had been measured, and 499 Illumina HighSeq raw counts covering 1,046 miRNAs, beside respectively 56 and 57 RNASeq and miRNASeq htseq-count files of normal thyroid tissues were included in this study.. Figure 1 shows the overall work flow of this study. Normalization steps and differential expression analysis between different stages, including "normal to stage I", "stage I to II", "stage II to III" and "stage III to IV" was performed separately using DeSeq2 package (Love et al., 2014) in R 3.5.0. Transcripts and miRNAs with row counts less than 10 were removed from all samples before normalization by DeSeq2. The list of the resulted differential expressed genes (DEGs) and miRNAs was then filtered for ≤ 0.05 adjusted p-value and $|\log_2 FC| \geq 1$.

In order to validate the final results of our RNASeq analysis, we also analyzed another independent PTC RNASeq data set (bio project PRJEB11591) which was the only available data set including 77 PTC (49 stage I, 18 stage III and 2 stage IV) beside 31 adjacent normal tissue samples. Trimming of paired end Fastq files was performed using Trimmomatic algorithm. Trimmed Fastq files were then mapped on human genome (homo sapiens) (b37): hg19 using BWA algorithm. The resulted bam files were then introduced in to htseq-count algorithm. DESeq2 program was then used to get the list of DEGs between normal vs. stage I and also between stage I vs. the integration of stage III and stage IV htseq-count files, due to the insufficient number of stage IV samples. Similarly, as a result of the small number of stage IV samples, differential expression analysis could not be carried out between stage III and stage IV.

Contrasting and comparing deregulated genes and miRNAs across individual tumor stages

In order to identify differences and similarities in the expression patterns of deregulated genes and miRNAs between different stages (including comparison of changes in direction, i.e., up to down-regulated, or down to up-regulated), the lists of the resulting differentially expressed genes and miRNAs were contrasted and compared between the results of the sequential stage comparisons.

To identify interactions between deregulated miRNAs and their corresponding targets, the lists of deregulated miRNAs in four group studies were separately introduced into miRDB.V6 (<http://mirdb.org>) and their corresponding target genes were identified. These target genes were then searched in the list of deregulated genes with the opposite pattern of expression (for example target genes of down-regulated miRNAs in each group were searched in the list of up-regulated genes in that group).

“Stage I” sample clustering of RNASeq and miRNASeq data

We clustered all genes in “stage I” samples (278 RNASeq and 282 miRNASeq data sets) normalized htseq-count files, using hierarchical clustering by using “Euclidean distance” measure in GENE-E [program\(gene-e@broadinstitute.org\)](http://gene-e@broadinstitute.org). Visually separated clusters of patients with distances ≤ 0.4 were then identified. The expressional changes of each cluster of stage I tumor samples were compared with normal samples and the lists of deregulated genes and miRNAs in each “stage I” clusters were identified using DESeq2. Finally the lists of down and up-regulated genes and miRNAs in each cluster were compared with the lists of deregulated genes in different stage transitions (from normal to stage IV).

Similarly the clustering of RNASeq stage I samples was carried out for 49 stage I samples of the validation data set.

Gene enrichment analysis of identified deregulated genes

To find deregulated pathways in each stage comparisons we performed pathway enrichment using Enrichr ([Chen et al., 2013](#)) and pathways with Benjamini-Hochberg adjusted p-value ≤ 0.05 was reported.

Results

mRNAs and miRNAs deregulations

The list of deregulated genes and miRNAs in different sequential stage progressions are respectively represented in Table 3–4 and Table 5. Deregulated genes and miRNAs were ranked based on their \log_2 FC of expression values. The results of the RNASeq validation study (normal to stage I and stage I to the integration of stage III-IV) showed the statistical significant results (adjusted p-value ≤ 0.05) for only normal to stage I comparison.

mRNA

Figure 2 and 3 show KEGG signaling pathways identified for up and down-regulated genes, respectively. Simultaneous down-regulation of RAC2, VAV1 and DOCK2 genes was observed in stage I to II progression. These proteins are involved in leukocyte transendothelial migration pathway. ITGAL and ITGB2 were also two down-regulated integrins ([Kanehisa et al., 2016, 2016](#)) (in stage I to II) that are present at the leukocyte membrane and involved in leukocyte binding to the endothelial membrane which in turn leads to the direct activation of VAV1 and indirect activation of Rac2. Concerted up-regulation of several extra cellular matrix (ECM) proteins, including THBS1, THBS2, COL6A3, SPP1, TNC, FN1, LAMB3 and COL1A2, along with ITGA11 (an ECM associated receptor) was observed in stage II to III transition. The up-regulation of these proteins results in the considerable activation of PI3K-AKT pathway and therefore leads to the continuous cell proliferation ([Kanehisa et al., 2016](#)). Moreover the activation of TGF-beta signaling pathway which contributes to the epithelial-mesenchymal transition (EMT), has previously been indicated following the up-regulation of THBS1 (up-regulated in stage II to III progression) ([Seliger et al., 2013](#)).

The up-regulation of three MHC class II members, including HLA-DOA, HLA-DQA1 and HLADRB1, was detected in stage II to III progression. The higher frequency of HLA-DRB1 in PTC patients compared to controls has been already demonstrated ([Amoli et al., 2010](#)). Similar results were demonstrated by JO and colleagues. They also indicated that the expression of MHC class II antigens (HLADQ and HLA-DR) is inversely correlated with the recurrence of PTC ([Jo et al., 2008](#)).

Comparing the list of up and down-regulated genes between pairs of continuous stages resulted in the identification of 3 down-regulated genes that were also detected as down-regulated in higher stages. These include FBLN1, identified as down-regulated in the “normal to stage I”, and also in “stage I to II” progression. Similarly, LTF and PLA2R1 were identified as down-regulated in both “normal to stage I” and “stage III to IV” progressions. No common up-regulated gene was identified in more than one stage comparison.

In order to identify any gene expression inversion between two continuous stages (e.g. down/up-regulated in “normal to stage I”, but up/down-regulated in stage I to II), the list of DEGs between each pair of continuous stages was compared with its immediate next stages pair. This led to the identification of a few genes with inverse pattern of expression (Table 6). Notably, 28 down-regulated genes in “stage I to II” progression were identified as up-regulated in “stage II to III”. The corresponding KEGG signaling pathways for these 28 genes has been shown in Figure 4. Several numbers of these genes have a direct or indirect role in the activation of apoptosis. These include HLA-DOA, HLA-DQA, BIRC3, COL1A1 and IL2RG. Down-regulation of these genes in stage II is expected in the primary stages of tumor development. However their over-expression, which leads to the increase of apoptotic activity, has also been previously confirmed in some malignancies including breast, endometrial and thyroid cancers (Soini et al., 1998). This increase in the apoptotic activity has previously been shown to be due to the presence of hypoxic conditions in the core regions of solid tumors and indicating the large area occupied by tumor cells (Gottfried et al., 2004). Similarly, changing the pattern of expression from down to up-regulated was observed for RAC2 which along with DOCK2 and VAV1 (identified as down-regulated in stage I to II progression), are associated with the regulation of actin cytoskeleton, production of reactive oxygen species (ROS) and metastasis (Kanehisa et al., 2016).

miRNA

There is large discrepancy among the results of studies on miRNA expression in thyroid cancer [20]. However the deregulation of several identified miRNAs in this study has previously been well confirmed in PTC (bold fonts in Table 5). Among previously reported deregulated miRNAs, the up-regulation of hsa-miR-146b and hsa-miR-222 and the down-regulation of hsa-miR-144 have been also reported in the blood of PTC patients (Poulios et al., 2018, Yu et al., 2012). A similar comparison to that of mRNAs was also carried out for differential expressed miRNAs to identify further deregulation of their expression in higher stages. hsa-miR-205, hsa-miR-509-2, hsa-miR-514-1, and hsa-miR-514-2 were four miRNAs whose up-regulation was detected in both “normal to stage I” and in “stage II to III” progressions. Similarly, hsa-miR-9-1 and hsa-miR-9-2 were identified as down-regulated in “normal to stage I” and “stage I to II” progressions. Moreover, gene expression pattern inversion was observed for hsa-miR-1247, which was detected as down-regulated in “stage I to II”, and as up-regulated in “stage II to III” progression.

Results of stage I tumor sample clustering

Clustering of stage I subjects revealed 6 clusters for RNASeq, and 5 clusters for miRNASeq (Table 7) (Figures 5 and 6). Stage I RNAseq sample clustering revealed much more heterogeneity than that of miRNASeq. Therefore 6 identified clusters (totally comprising 265 out of 278 tumor samples, after removing outliers) of stage I RNASeq samples were subjected to further analyses. Comparing the lists of down and up-regulated genes in each “stage I” cluster (Supplementary table 1) with the lists of deregulated genes in higher stages led to the identification of few deregulated genes (Table 8).

LTF which its down-regulation was also identified in “stage III to IV” progression, belongs to the cluster 3 of “stage I”, which includes 102 samples. Since we did not expect tumor progression to the highest stage (stage IV) in a large number of PTC patients, we further sub-clustered these 102 samples in order to find a specific sub-cluster with the most contribution in the observed down-regulation of LTF. A total of 7 sub-clusters were identified in “cluster 3” of “stage I”. Running DeSeq2 showed that in only one sub-cluster (containing 21 samples), LTF was significantly down-regulated (\log_2 FC = -3). This level of down-regulation (8 fold reduction in the expression) appeared to influence LTF average expression in the whole cluster 3 (Table 8).

No deregulated genes were detected in “cluster 4” of stage I compared with the higher stages. Therefore it seems that patients in clusters “1, 2, 5 and 6” of stage I, as well as 21 patients in “cluster 3” (totally comprising 71 subjects out of 278 stage I tumor samples) which have shown further deregulation in higher stages, would probably be in the higher risk for tumor progression. No statistically significant difference between the potential high and low risk (71 vs. 207) groups of stage I patients was detected based on their age, age of onset, race and sex.

The changes in the expressional pattern of two genes i.e. LTF and PLA2R1 was also checked in different stage I clusters’ of the validation data set. In total 3 clusters of samples (including C1 = 12, C2 = 10 and C3 = 6 samples) were identified in stage I. While the expression of LTF and PLA2R1 showed mild down-regulation in C1 and C2, their significant down-regulation was observed in C3 which includes only 6 samples (\log_2 FC = -3.78 and -3.57 for LTF and PLA2R1 respectively, equals to approximately 13.5 fold expressional reduction). This indicates almost similar percentage of stage I patients with significant down-regulation of LTF and PLA2R1 (for LTF, 21 out of 262 samples (8%) in TCGA data set and 6 out of 49 (12%) in the validation data set and for PLA2R1, 42 out of 262 samples (16%) in TCGA data set and 6 out of 49 (12%) in the validation data set).

Results of comparing down and up-regulated miRNAs in each miRNAs clusters with the list of deregulated miRNAs obtained from stage comparisons was resulted in to the identification of hsa-miR-1247 whose down-regulation was observed in cluster 4 of “stage I”, and also in “stage I to II” progression. However it was also detected as up-regulated in “stage II to III” progression.

Discussion

Despite good prognosis of PTC, it has been indicated that a sub-group of PTC patients show an aggressive phenotype characterized by local recurrence and/or distant metastasis (Blanco et al., 2012). Characterizing DNA level molecular alterations of PTC including mutations and structural variations has already been the subject of several studies (Katoh et al., 2015, Acquaviva et al., 2018). Recently, a few transcriptomic investigations have also been conducted to generate a comprehensive view of PTC molecular alterations (Teng et al., 2018, Liyanarachchi et al., 2016). Results of these studies have suggested several biomarkers that may have application in the early detection of PTC or as novel therapeutic targets. However no systematic study has been conducted to unravel the stage specific

transcription alterations occurring in PTC. The identified stage-specific transcriptional alterations in our results can significantly improve our understanding about the gene expression signature of different PTC stages and suggests potential biomarkers which could potentially be implemented to classify subgroups of patients who might progress to the more advanced stages.

Biomarkers with prognostic potential in PTC

The deregulation of a total of 12 genes in different sub-clusters of stage I were further identified in higher stages (Table 7). As we expected, for the majority of stage I PTC samples (207), no further deregulation was observed in the higher stages.

LTF and PLA2R1 were two down-regulated tumor suppressors in “stage I” that, interestingly, exhibited further down-regulation in “stage III to IV” progression. The association of down-regulation of LTF with the poor prognosis of PTC has recently been indicated ([Brennan et al., 2016](#)). The role of PLA2R1 in the activation of signaling pathways related to apoptosis and senescence and its down-regulation in some cancers including thyroid, renal, mammary and leukemia has been also previously established ([Menschikowski et al., 2015](#)). As stated earlier, the significant down-regulation of LTF and PLA2R1 was also confirmed in respectively 12% and 16% of stage I samples of the validation data set. However as a result of the small number of stage IV samples (only 2 samples), we could not check their expressional changes in stage III to IV progression. Nevertheless it seems that these two genes can be considered as the potential biomarkers associated with the poor prognosis in PTC.

TRIB3 and KIF5C are the other two down-regulated genes in stage I which were also detected as down-regulated in stage II to III progression. TRIB3 is a tumor suppressor which its loss has been shown to lead to the enhanced phosphorylation of AKT and in turn to the hyperphosphorylation of FOXO3 which results to its inactivation ([Salazar et al., 2015](#)). There are contradictory reports about KIF5C pattern of deregulation in thyroid cancer (both the down and up-regulation of KIF5C in thyroid tumors has already been reported) ([Lu et al., 2011](#), [Rodriguez et al., 2012](#)).

There is still no report about the deregulation of SNX20 in thyroid cancer. However in a very recent case report study, somatic mutations of this gene have been indicated in both tissue DNA and in circulating cell free DNA (cfDNA) of a follicular thyroid carcinoma patient with lung and bone metastasis ([Song and Yang, 2018](#)).

The further up-regulation of all seven up-regulated genes in different sub-clusters of stage I were identified in stage II to III progression (Table 8). Among these, CST7 and FN1 were two genes that have previously been suggested as biomarkers associated with poor prognosis of PTC ([Teng et al., 2018](#), [Cerutti et al., 2007](#)). Similarly the over-expression of TNFSF13B in tumor cells has been shown to be related to the poor prognosis in non small cell lung cancer ([Qian et al., 2014](#)). The up-regulation of SCE1 in PTC has previously been demonstrated by Huang and colleagues ([Huang et al., 2001](#)). The genetic polymorphism of IL1RN has been previously associated with the poor prognosis in cancer ([Niedźwiecki et](#)

al., 2008). Moreover IL1RN has been already suggested as a biomarker involved in tumor differentiation of thyroid carcinoma (Blanco et al., 2012).

It has been confirmed by several studies that the hypermethylation of Runx3 is in association with poor prognosis in thyroid cancer (Joung and Shong, 2012, dos Reis et al., 2016, Chen et al., 2013). However we identified RUNX3 as a up-regulated gene in both a very small sub-cluster of “stage I” (including 3 samples) and also in stage “I to II” progression. Interestingly the results of a recent comprehensive meta-analysis conducted on three different cancers (including Breast, Colon and Lung), provide further interesting evidence about the direct relationships between the level of hypermethylation of genes and the levels of genes expression. They indicated that except for the genes that are lowly expressed in normal tissues, the hyper or hypomethylation of genes in cancer tissue does not necessarily result in respectively their down or up-regulation (Moarii et al., 2015).

No significant difference was identified between the age, age of onset, race and sex of potential 71 high risk and 207 low risk stage I patients. This indicates the importance of using transcriptional biomarkers for classifying PTC patients and selecting the efficient personalized treatment.

Prognostic miRNAs for PTC tumor progression

As stated earlier, the results of miRNA expression studies vary significantly due to the several factors including the heterogeneity of tumor samples, the variations in sample size and in the quantitative method implemented for expression measurements. The problem becomes more complex considering the fact that there is a direct correlation between the deregulation of some miRNAs with some oncogenic mutations. For instance, the up-regulation of hsa-miR-146b, hsa-miR-221, hsa-miR-222, hsa-miR-21-5p, hsa-miR-31-5p, hsa-miR-34a-5p, hsa-miR-181b-5p and hsa-miR-551b-3p have been confirmed in BRAF V600E samples (Saiselet et al., 2016).

In spite of the large number of analyzed tumor samples in the present study, due to the imbalance between the number of tumor vs. normal samples, the results of this study are not comparable to other studies performed on miRNA expression alterations in PTC. However as it is shown in Table 5, the deregulation of a considerable number of identified miRNAs have already been confirmed by several studies. As stated in the results section, the down-regulation of hsa-miR-205, hsa-miR-509-2, hsa-miR-514-1, and hsa-miR-514-2 were detected in both stage I and stage III and therefore after further confirmation they can be implemented as new predictors of tumor progression. There are contradictory reports about the expression pattern of hsa-miR-1247 (down-regulated in stage I and up-regulated in stage III). Down-regulation of hsa-miR-1247 has previously been reported in thyroid tumors with follicular pattern of growth (Mancikova et al., 2015). Similarly the down-regulation of hsa-miR-1247 has been confirmed in the lymph node metastasis in breast cancer (Zhang et al., 2018). However the association between the up-regulation of this miRNA with lung metastasis has been shown in liver cancer. It has been indicated that tumor derived exosomal hsa-miR-1247 affects cancer associated fibroblasts and activates lung metastasis (Fang et al., 2018).

Expression measurements of other identified deregulated miRNAs with no previous reports on their association with PTC, in a large number of PTC patients, will reveal if they can be truly considered as the novel prognostic biomarkers in PTC. The result of searching for the targets of deregulated miRNAs among the list of deregulated mRNAs led to the identification of only one interaction between HSPA1B (down-regulated in normal to stage I) and hsa-miR-34a (up-regulated in normal to stage I). This is however in accordance with studies demonstrating that the most alterations in mRNA expression are short term and independent of miRNAs (Saiselet et al., 2016).

We believe that the strength of this study lies in its stage-wise approach, which to the best of our knowledge, has not already been performed in PTC. Moreover the large sample size used in this study increases the confidence of our results. The main limitations of this study include the small number of normal samples compared to the large number of tumor samples and also the lack of complete demographic data for the PTC patients included in this study.

Conclusions

We found large amount of heterogeneity in the gene expressional pattern of stage I PTC patients. Stratifying these patients in to high and low risk subgroups could help us to select more personalized cancer management in these patients. Stage-specific biomarkers identified in stage I PTC patients could efficiently help to distinguish the majority of patients with low risk of cancer progression from those who might progress to higher stages in PTC.

Declarations

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Availability of data and material

Data are available at (<http://portal.gdc.cancer.gov/>)

Competing interests

Authors declare no competing interests.

Funding

Financial support was provided by Iran University of Medical Sciences, (Grant number 33444).

Authors' contribution

M.K designed and supervised the project. N.H performed the data analyses and wrote the manuscript. K.B, M.K and M.H commented on the manuscript. T.M reviewed the paper.

Acknowledgements

The authors gratefully acknowledge the supports provided by Iran University of Medical Sciences, (Grant number 33444).

References

2016. UniProt: the universal protein knowledgebase. *Nucleic acids research*, 45, D158-D169.

ACQUAVIVA, G., VISANI, M., REPACI, A., RHODEN, K. J., DE BIASE, D., PESSION, A. & GIOVANNI, T. 2018. Molecular pathology of thyroid tumours of follicular cells: a review of genetic alterations and their clinicopathological relevance. *Histopathology*, 72, 6–31.

AGRAWAL, N., AKBANI, R., AKSOY, B. A., ALLY, A., ARACHCHI, H., ASA, S. L., AUMAN, J. T., BALASUNDARAM, M., BALU, S. & BAYLIN, S. B. 2014. Integrated genomic characterization of papillary thyroid carcinoma. *Cell*, 159, 676–690.

AMOLI, M. M., YAZDANI, N., AMIRI, P., SAYAHZADEH, F., HAGHPANAH, V., TAVANGAR, S. M., AMIRZARGAR, A., GHAFARI, H., NIKBIN, B. & LARIJANI, B. 2010. HLA-DR association in papillary thyroid carcinoma. *Disease markers*, 28, 49–53.

BLANCO, C. G., MATUTE, E. M. & DE LEIVA HIDALGO, A. 2012. Molecular biomarkers involved in the tumor dedifferentiation process of thyroid carcinoma of epithelial origin: perspectives. *Endocrinología y Nutrición (English Edition)*, 59, 452–458.

BRAY, F., FERLAY, J., SOERJOMATARAM, I., SIEGEL, R. L., TORRE, L. A. & JEMAL, A. 2018. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: a cancer journal for clinicians*, 68, 394–424.

BRENNAN, K., HOLSINGER, C., DOSIOU, C., SUNWOO, J. B., AKATSU, H., HAILE, R. & GEVAERT, O. 2016. Development of prognostic signatures for intermediate-risk papillary thyroid cancer. *BMC cancer*, 16, 736.

CERUTTI, J. M., OLER, G., MICHALUART, P., DELCELO, R., BEATY, R. M., SHOEMAKER, J. & RIGGINS, G. J. 2007. Molecular profiling of matched samples identifies biomarkers of papillary thyroid carcinoma lymph

node metastasis. *Cancer research*, 67, 7885–7892.

CHEN, E. Y., TAN, C. M., KOU, Y., DUAN, Q., WANG, Z., MEIRELLES, G. V., CLARK, N. R. & MA'AYAN, A. 2013. Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC bioinformatics*, 14, 128.

DOS REIS, M., BELTRAMI, C., KOWASLKI, L. & ROGATTO, S. 2016. Epigenetic alterations in well-differentiated thyroid cancer. *J Clin Epigenet*, 1, 1.

FANG, T., LV, H., LV, G., LI, T., WANG, C., HAN, Q., YU, L., SU, B., GUO, L. & HUANG, S. 2018. Tumor-derived exosomal miR-1247-3p induces cancer-associated fibroblast activation to foster lung metastasis of liver cancer. *Nature communications*, 9, 191.

GOTTFRIED, Y., VOLDAVSKY, E., YODKO, L., SABO, E., BEN-ITZHAK, O. & LARISCH, S. 2004. Expression of the pro-apoptotic protein ARTS in astrocytic tumors: Correlation with malignancy grade and survival rate. *Cancer: Interdisciplinary International Journal of the American Cancer Society*, 101, 2614–2621.

HARALDSDOTTIR, S. & SHAH, M. H. 2014. New era for treatment in differentiated thyroid cancer. *The Lancet*, 384, 286–288.

HUANG, Y., PRASAD, M., LEMON, W. J., HAMPEL, H., WRIGHT, F. A., KORNACKER, K., LIVOLSI, V., FRANKEL, W., KLOOS, R. T. & ENG, C. 2001. Gene expression in papillary thyroid carcinoma reveals highly consistent profiles. *Proceedings of the National Academy of Sciences*, 98, 15044–15049.

JO, Y. S., LEE, J. C., LI, S., CHOI, Y. S., BAI, Y. S., KIM, Y. J., LEE, I. S., RHA, S. Y., RO, H. K. & KIM, J. M. 2008. Significance of the expression of major histocompatibility complex class II antigen, HLA-DR and-DQ, with recurrence of papillary thyroid cancer. *International journal of cancer*, 122, 785–790.

JOUNG, K. H. & SHONG, M. 2012. Epigenetic regulation of RUNX3 in thyroid carcinoma. *The Korean journal of internal medicine*, 27, 391.

KANEHISA, M., FURUMICHI, M., TANABE, M., SATO, Y. & MORISHIMA, K. 2016. KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic acids research*, 45, D353-D361.

KATOH, H., YAMASHITA, K., ENOMOTO, T. & WATANABE, M. 2015. Classification and general considerations of thyroid cancer. *Ann Clin Pathol*, 3, 1045.

LIU, R., ZHANG, T., ZHU, G. & XING, M. 2018. Regulation of mutant TERT by BRAF V600E/MAP kinase pathway through FOS/GABP in human cancer. *Nature communications*, 9, 579.

LIYANARACHCHI, S., LI, W., YAN, P., BUNDSCHUH, R., BROCK, P., SENTER, L., RINGEL, M. D., DE LA CHAPELLE, A. & HE, H. 2016. Genome-wide expression screening discloses long noncoding RNAs involved in thyroid carcinogenesis. *The Journal of Clinical Endocrinology & Metabolism*, 101, 4005–4013.

- LOVE, M., ANDERS, S. & HUBER, W. 2014. Differential analysis of count data–the DESeq2 package. *Genome Biol*, 15, 10.1186.
- LU, C., MISHRA, A., ZHU, Y. J., MELTZER, P. & CHENG, S.-Y. 2011. Genomic profiling of genes contributing to metastasis in a mouse model of thyroid follicular carcinoma. *American journal of cancer research*, 1, 1.
- MANCIKOVA, V., CASTELBLANCO, E., PINEIRO-YANEZ, E., PERALES-PATON, J., DE CUBAS, A. A., INGLADA-PEREZ, L., MATIAS-GUIU, X., CAPEL, I., BELLA, M. & LERMA, E. 2015. MicroRNA deep-sequencing reveals master regulators of follicular and papillary thyroid tumors. *Modern Pathology*, 28, 748.
- MENSCHIKOWSKI, M., HAGELGANS, A., NACKE, B., JANDECK, C., SUKOCHEVA, O. & SIEGERT, G. 2015. Epigenetic control of phospholipase A 2 receptor expression in mammary cancer cells. *BMC cancer*, 15, 971.
- MOARII, M., BOEVA, V., VERT, J.-P. & REYAL, F. 2015. Changes in correlation between promoter methylation and gene expression in cancer. *BMC genomics*, 16, 873.
- NIEDŹWIECKI, S., STĘPIEŃ, T., KUZDAK, K., STĘPIEŃ, H., KRUPIŃSKI, R., SEEHOFER, D., RAYES, N. & ULRICH, F. 2008. Serum levels of interleukin–1 receptor antagonist (IL–1ra) in thyroid cancer patients. *Langenbeck's archives of surgery*, 393, 275–280.
- POULIOS, E., ZOGRAFOS, G., BALAFOUTA, S. & LINOS, D. 2018. Identification of a Novel. *Diagnostic MicroRNA Signature in Papillary Thyroid*.
- QIAN, Z., QINGSHAN, C., CHUN, J., HUIJUN, Z., FENG, L., QIANG, W., QIANG, X. & MIN, Z. 2014. High Expression of TNFSF13 in Tumor Cells and Fibroblasts Is Associated With Poor Prognosis in Non–Small Cell Lung Cancer. *American journal of clinical pathology*, 141, 226–233.
- RODRIGUEZ, M. D.C. P., FREIRE, M. D.C. C., CAMESELLE-TEIJEIRO, J. M., PUENTE, F. D. & FIGUEROA, A. V. 2012. Methods for diagnosing follicular thyroid cancer. Google Patents.
- SAISELET, M., PITA, J. M., AUGENLICHT, A., DOM, G., TARABICHI, M., FIMERELI, D., DUMONT, J. E., DETOURS, V. & MAENHAUT, C. 2016. miRNA expression and function in thyroid carcinomas: a comparative and critical analysis and a model for other cancers. *Oncotarget*, 7, 52475.
- SALAZAR, M., LORENTE, M., GARCÍA-TABOADA, E., GOMEZ, E. P., DÁVILA, D., ZÚÑIGA-GARCÍA, P., FLORES, J. M., RODRÍGUEZ, A., HEGEDUS, Z. & MOSEN-ANSORENA, D. 2015. Loss of Tribbles pseudokinase–3 promotes Akt-driven tumorigenesis via FOXO inactivation. *Cell death and differentiation*, 22, 131.
- SELIGER, C., LEUKEL, P., MOECKEL, S., JACHNIK, B., LOTTAZ, C., KREUTZ, M., BRAWANSKI, A., PROESCHOLDT, M., BOGDAHN, U. & BOSSERHOFF, A.-K. 2013. Lactate-modulated induction of THBS–1 activates transforming growth factor (TGF)–beta2 and migration of glioma cells in vitro. *PloS one*, 8, e78935.

SOINI, Y., PÄÄKKÖ, P. & LEHTO, V. 1998. Histopathological evaluation of apoptosis in cancer. *The American journal of pathology*, 153, 1041–1053.

SONG, J. & YANG, Z. 2018. Case report: Whole exome sequencing of circulating cell-free tumor DNA in a follicular thyroid carcinoma patient with lung and bone metastases. *Journal of circulating biomarkers*, 7, 1849454418763725.

TENG, H., MAO, F., LIANG, J., XUE, M., WEI, W., LI, X., ZHANG, K., FENG, D., LIU, B. & SUN, Z. 2018. Transcriptomic signature associated with carcinogenesis and aggressiveness of papillary thyroid carcinoma. *Theranostics*, 8, 4345.

YU, S., LIU, Y., WANG, J., GUO, Z., ZHANG, Q., YU, F., ZHANG, Y., HUANG, K., LI, Y. & SONG, E. 2012. Circulating microRNA profiles as potential biomarkers for diagnosis of papillary thyroid carcinoma. *The Journal of Clinical Endocrinology & Metabolism*, 97, 2084–2092.

ZHANG, P., FAN, C., DU, J., MO, X. & ZHAO, Q. 2018. Association of miR-1247-5p expression with clinicopathological parameters and prognosis in breast cancer. *International journal of experimental pathology*, 99, 199–205.

Tables

Due to technical limitations, tables are only available as a download in the supplemental files section

Figures

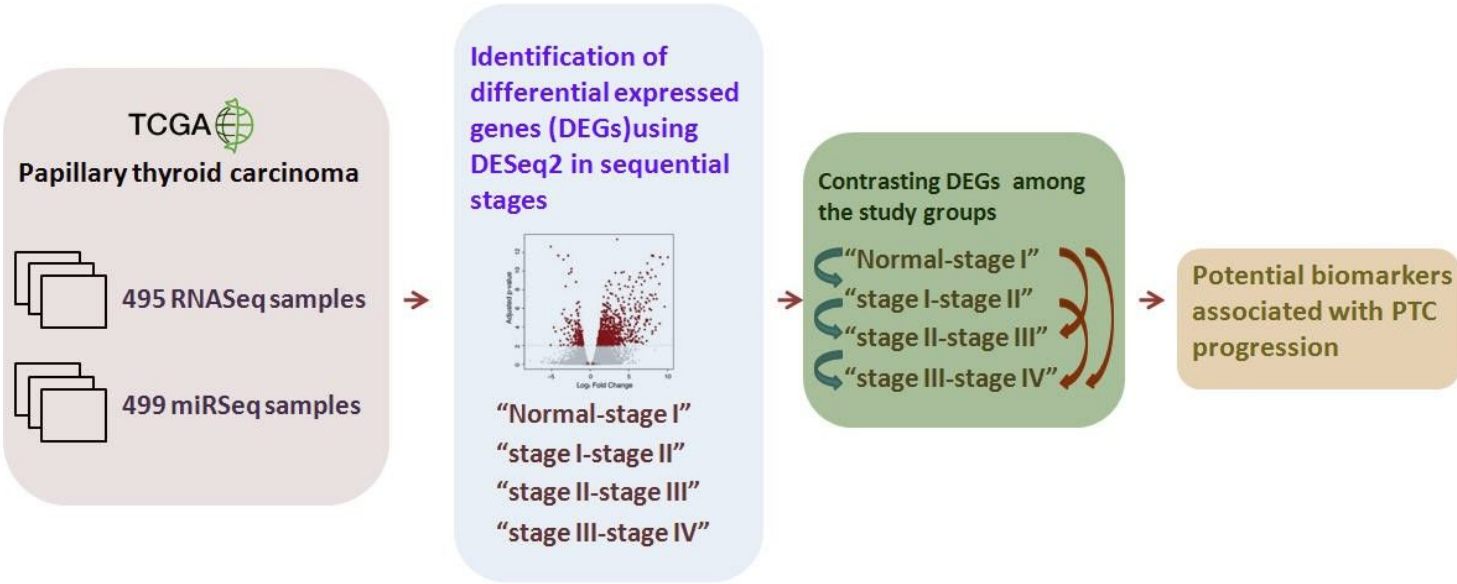


Figure 1

Overall work flow of the study.

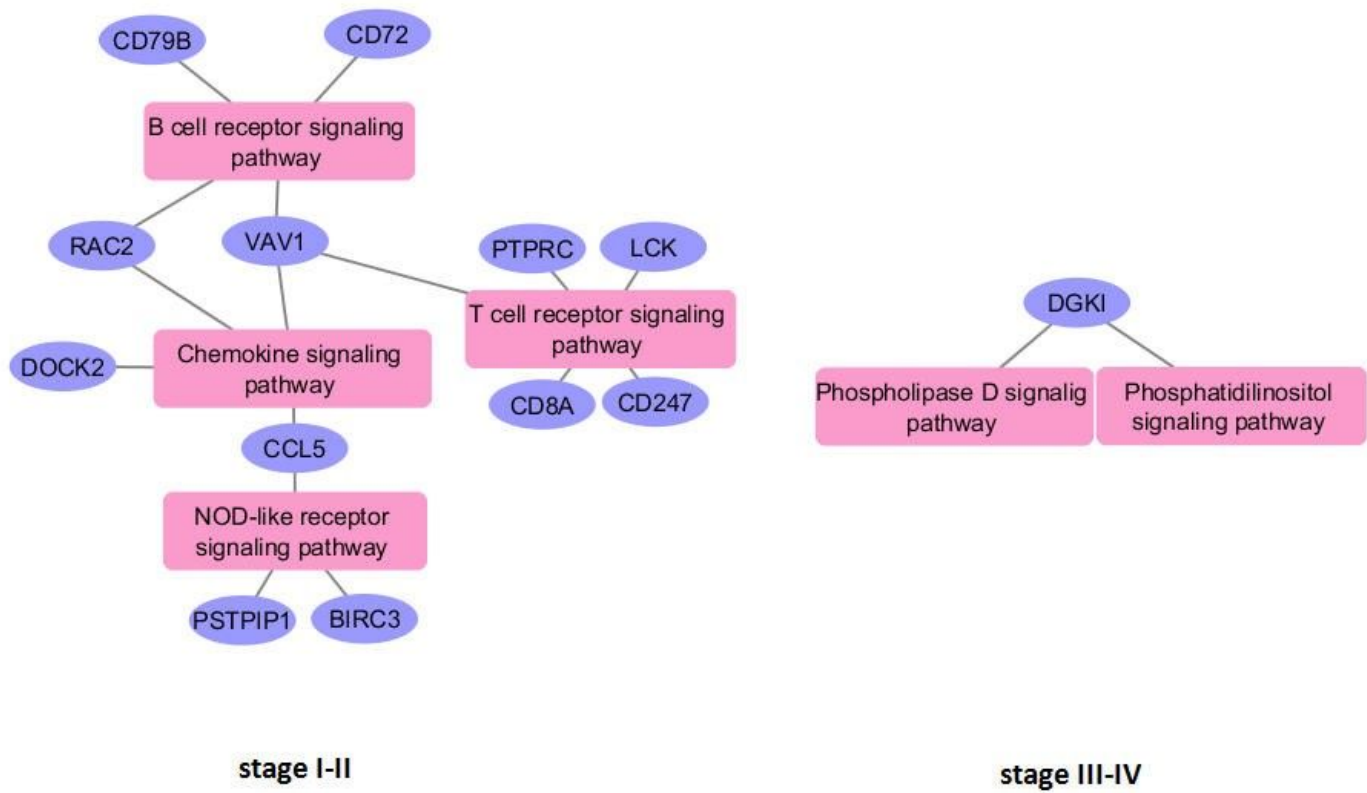


Figure 2

Down-regulated genes in PTC (blue) involved in signaling pathways (pink).

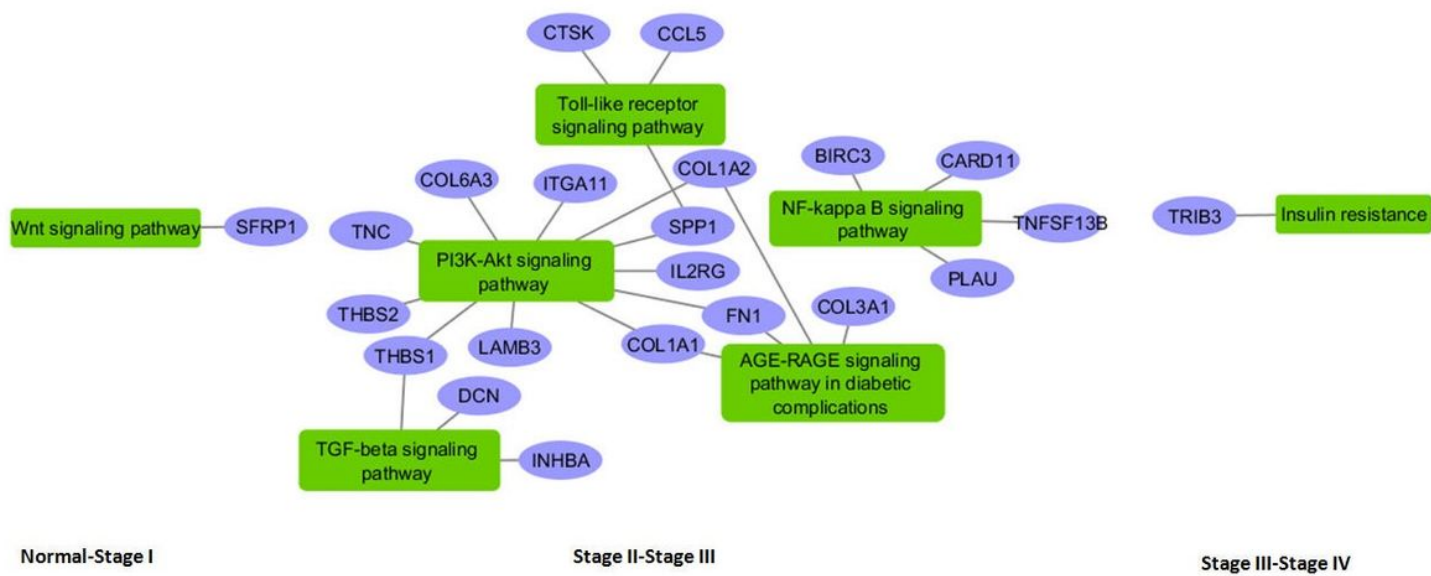


Figure 3

Up-regulated genes in PTC (blue) involved in signaling pathways (green).

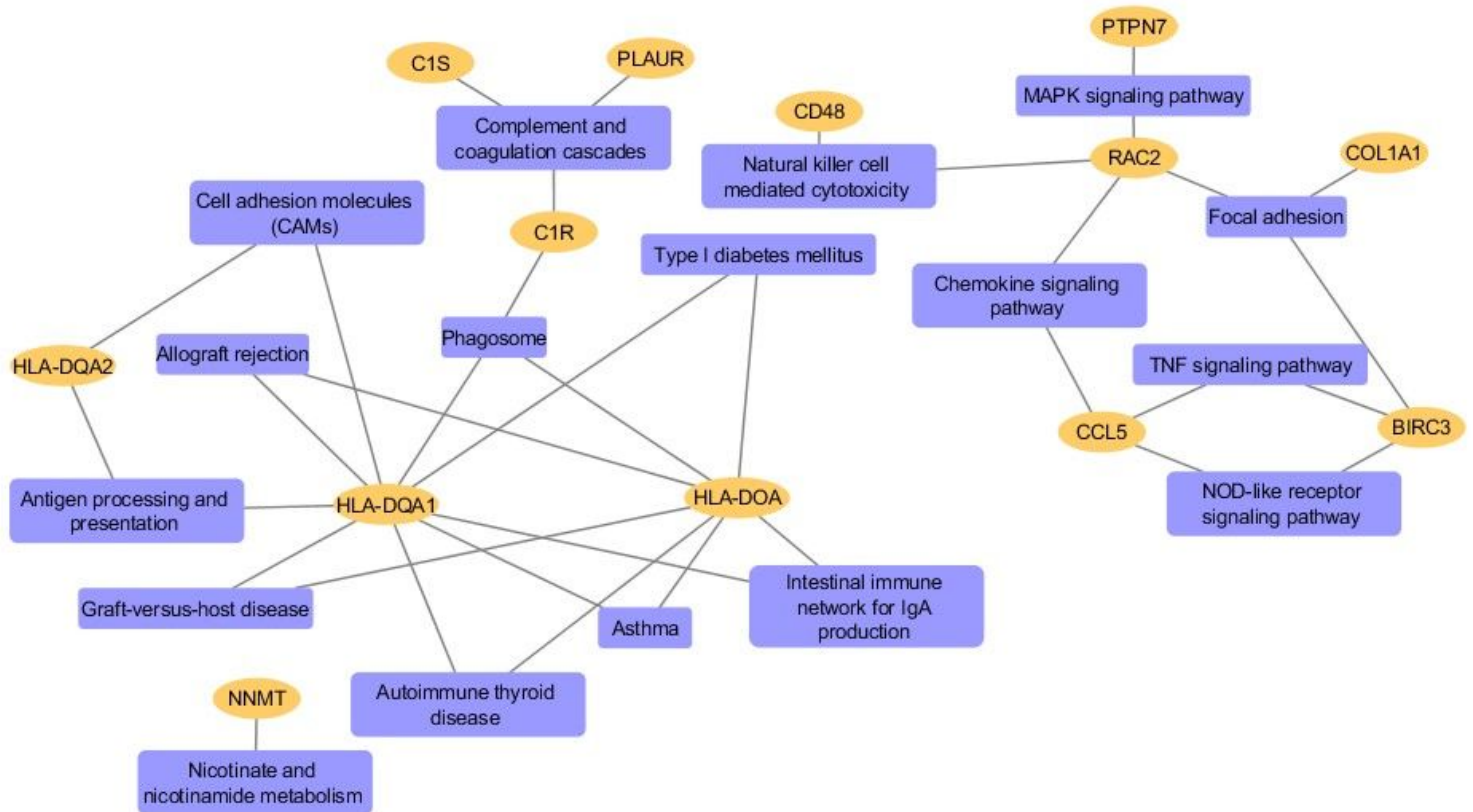


Figure 4

Pathways (light blue) associated with deregulated genes (orange) with the altered pattern of expression in stage I to II and stage II to III progressions.

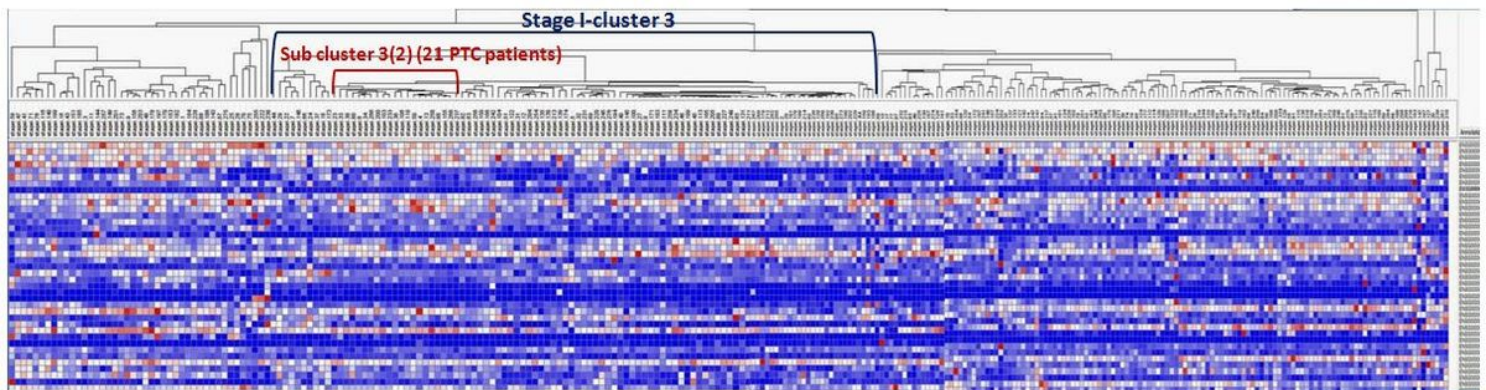


Figure 5

Clustering of stage I patients using GeneE program.

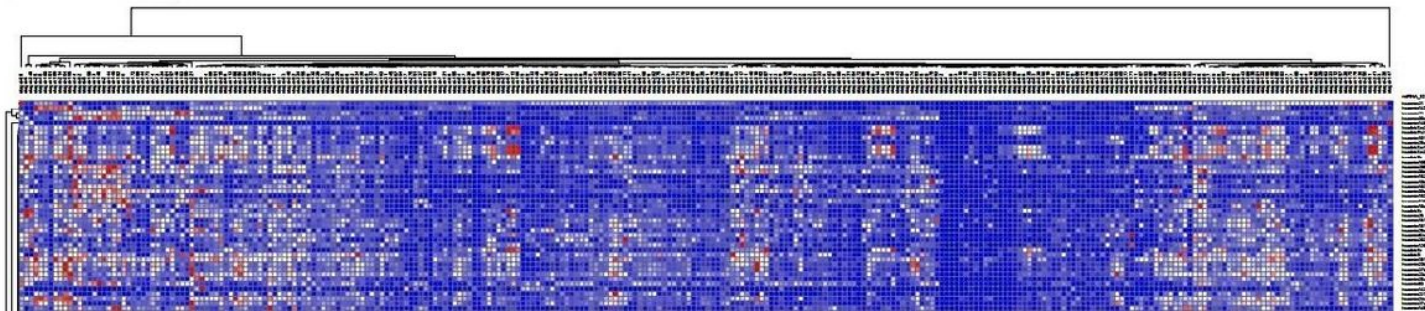


Figure 6

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Table8.xls](#)
- [Table7.xls](#)
- [Table5.xls](#)
- [Table2.xls](#)
- [Table1.xls](#)
- [Table3.xls](#)
- [Table4.xls](#)
- [Supplementarytable1.xls](#)
- [Table6.xls](#)