The Role of Skin-Specific TRAs on Skin Cancer

Binnur Özay, Christofer Richard, Denisa Neacsu, Pinar Cavus

Data Analysis Project SS21

ABSTRACT

Tissue restricted antigen genes have been researched on because of their ability to open ways to new personalized and specific cancer treatments which is not harmful for the rest of the organism, except for the concerned tissue. In this project, certain tissue restricted antigen genes in melanoma show upregulated expressions after being treated with inhibitors of the MAPK pathway. As they are analyzed further, genes that are expressed under the influence of the sun and genes related to melanosome biogenesis are found. The dependency of the upregulated genes on the genes representing the rest of the expressions with highest variance was questioned. For this analysis, methods such as principal component analysis, box plots, log fold change and linear regression models have been used.

INTRODUCTION

Skin cancer makes up most of the malignancies all around the world, and consists of two main different types; melanoma and non-melanoma (Linares *et al.*, 2015). Both are mainly caused by exposure to UV lights, but non-melanoma cancer develops in the outer layers of the skin (Leiter & Garbe, 2008), whereas melanoma develops from the melanocytes (Smith *et al.*, 2003). Melanomas only make up a small amount of reported skin cancer diagnoses, however they still account for the majority of skin cancer deaths, making them extremely dangerous (Smith *et al.*, 2003). Melanomas have the highest rate of metastatic capacity of all tumors when they have delayed diagnoses (Potrony *et al.*, 2015).

More than half of all melanoma cases contain a mutation of the BRAF gene (Alqathama, 2020). The B-Raf protein is a signaling molecule in the Mitogen-activated protein kinase (MAPK) pathway, activating MEK and leading to the phosphorylation of ERK (Alqathama, 2020). This kinase cascade is known to direct important cell processes such as proliferation, differentiation, and survival (Alqathama, 2020).

In this analysis, data from the melanoma cell line WM266-4 is used. This cell line has a BRAF-V600D mutation, meaning that the valine at codon 600 is changed into aspartic acid. This mutation leads to an overexpression of BRAF, thus an uncontrolled increase in oncogenic signaling (Alqathama, 2020). The cell line is treated with Trametinib, which is an inhibitor of MEK 1 and MEK2 (Zeiser, 2014), ERK1/2-inhibitor, and doxycycline-shERK1 for different periods of time; 3 and 24 hours, 3 and 7 days. Later, the mRNA from these cells are extracted and processed into labeled cRNA fragments. These fragments are hybridized on microarrays and scanned, resulting in expression level quantifications of large amounts of genes. It is to mention that there are no units in this data, because it is simply light intensity scanned.

The focus in this microarray analysis is a mechanism cancer cells have been shown to use; the upregulation of tissue restricted antigen (TRA) genes. TRAs are highly expressed selectively in their tissues, and T-cells reacting to these self-antigens result in autoimmune diseases (Kyewski & Derbinski, 2004). In order for immature T-cells to gauge the self-reactivity of their antigen receptors, they go through a selection in the thymus, where T-cells reacting to TRAs undergo apoptosis (Klein *et al.*, 2014). TRA genes are mostly upregulated in cancer cells for undiscovered reasons, making them good potential drug targets for therapy (Rosenberg, 1999). Current treatments for melanomas include chemotherapy, radiation, immunotherapy, and surgical excision (Gruber & Zito, 2020). Developing a new therapy, where the upregulated skin specific TRA

genes are attacked, therefore other tissues don't get as damaged, would be a significant discovery in the medical field.

Therefore, the main aim of this project is to find upregulated skin specific genes in skin cancer. By looking into the gene expression data of TRA genes in a melanoma cell line, and comparing the variety of the expression levels between different treatments and treatment times, it is possible to gain insight into how the genes affect the survival of the cancer cells. The genes which present significant changes in their expression would be potential drug targets to further investigate, as they would be helping the cancer cell survive under the treatment.

METHODS AND RESULTS

Quality Control and Normalization of the Data

The programming language R Studio Version 1.4.1106 (R Core Team, 2021) is used throughout this entire project. First, all of the data to be analyzed is downloaded; these are the numerical microarray data with gene expression intensities of the melanoma cell line from the GEO website series GSE57721 and the nominal TRA data (Dinkelacker, 2007;2019) with names of genes originating from different tissues. The TRA data has other information such as chromosome numbers of the given genes which will be put to use later.

TRA data is made up of genes from different tissues, different cell lines of humans and mice. All skin-specific genes are extracted and combined in a cleaned vector with all NA values erased. As the skin data available on mice is very little, only the expression of human genes is analyzed. One set of TRA data has the information of sun exposure, as this information might be valuable further into the analysis, these genes are separated into two vectors; exposed to the sun and not exposed to the sun. A pie chart is finalized to visualize how much of the TRA data comes from skin. It is found out that skin specific genes make up only 2.98% of all TRA data.

Before the analysis, the data must be investigated to see if there are any problems that would alter the results, such as broken microarray chips. Single chips are controlled by visualizing raw microarray data, and apparently there are no abnormalities visible in all of the fifteen chips. Thus, all of the chips are usable.

Gene expression profiling results are highly influenced by RNA quality and degradation (Opitz *et al.*, 2010), therefore RNA degradation plots are made to investigate. All of the lines, each representing a chip, showed similar slope values, and there is not any deviation visible. For this reason, all chips are to be included in the analysis as they are all of quality.

The microarray data is normalized using the package vsn (v3.58.0; Hüber, NA) to remove any systematic variation that can affect the analysis. The pre-normalization and normalized data, which is glog2 transformed, are compared with boxplots and histograms, in order to see how much systematic variation existed. The boxplots didn't show a significant difference, even before normalization the median lines are almost in a linear line, indicating that the chips did not have much variation to begin with. In both pre- and after normalization histograms, the chip lines follow the same path, and none of the chips stand out. After the normalization, a meanSdPlot is made, and from this plot, it is seen that the mean values of the chips are almost linear, once again proving that they are high quality chips.

Cleaning Data and Creating a Data Frame

In the data from microarray chips the genes are saved as digits; to be able to view and analyse the data easier, these digits are replaced with their gene names using the ensembl IDs that were downloaded from https://www.ensembl.org/biomart . Later, all expression values are extracted from the microarray data and put into a data frame with chip IDs as column names and gene symbols as row names.

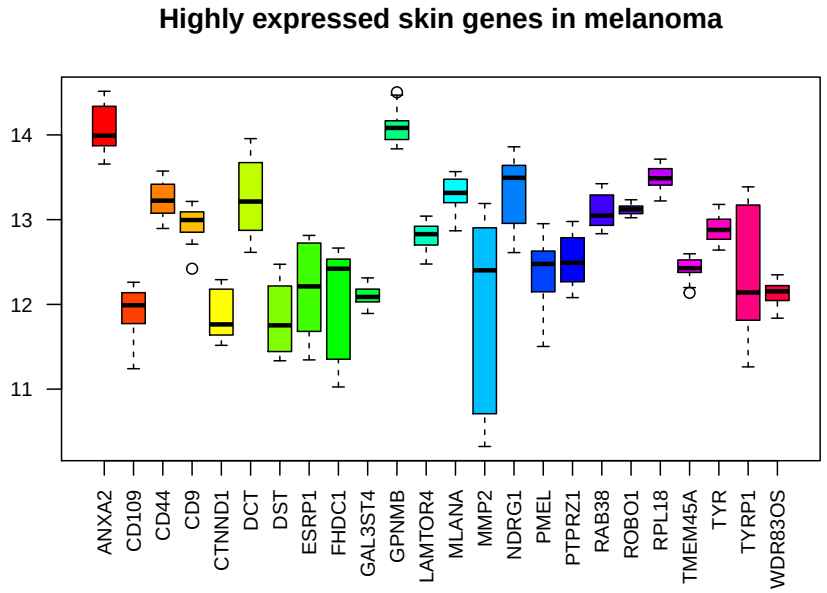
After empty spaces and NA values are removed from the data, skin specific antigens are extracted to create the main data frame of skin TRA expression levels in skin cancer to analyze.

Before getting to the gene expression analysis, a bar plot is created to find out which chromosome the skin-specific TRAs are mostly located at. According to the literature, most skin related genes are located on chromosome 1 (Volz *et al.*, 1993), which is also the case in this data.

Gene Expression Analysis

To begin the analysis a heat map is created, showing the expression levels of skin specific TRA genes in all chips. This heat map is useful in recognizing expression patterns to focus on, and it gives a starting point for further analysis. It is observed that the expression levels of some genes increase as the treatments goes longer. These are the genes that the cancer cells need to upregulate to survive against the different treatments.

To begin the analysis, the gene expression is visualized with box plots. There are too many genes overall, therefore the boxes look like overlapping lines, which does not give much information. Instead, only the highly expressed genes are chosen to be visualized. A highly expressed gene does not have to be an upregulated gene, and an upregulated gene has the possibility of not being a highly expressed gene. A threshold of twelve is set according to the overall expression levels to define the high expression margin. Aside from the two constantly highest expressed genes GPNMB and ANXA2, while the genes NDRG1 and DCT show varying expression levels between chips, with values above the threshold in some chips and below in others.

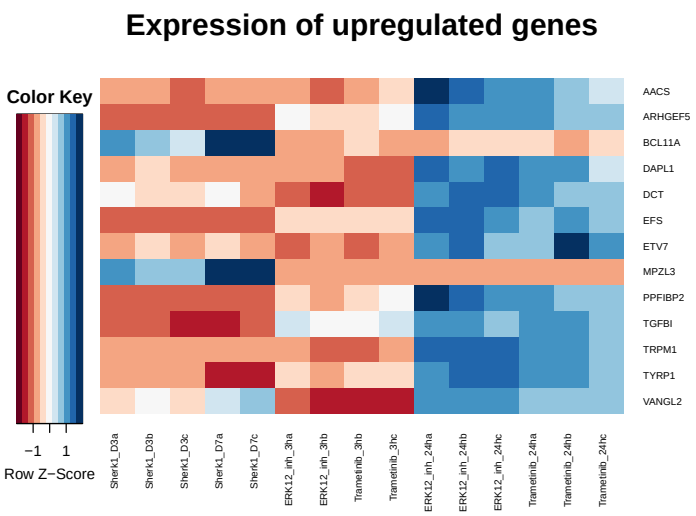


Upregulated Genes

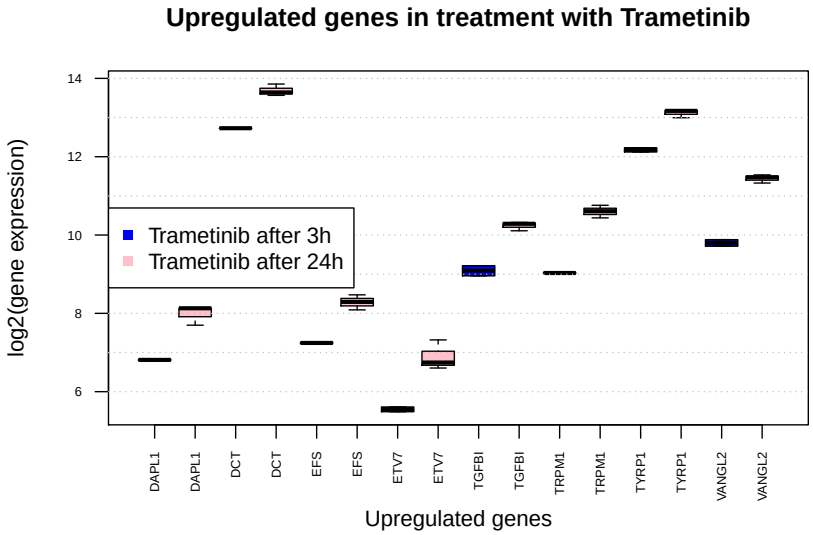
In order to determine the mostly upregulated genes, the expression data has to be transformed by using a log fold change. There is no data available from cells without any treatment, therefore for chips that have been treated with ERK1/2-inhibitor or Trametinib, the gene expression between the three hour and twenty four hour time points is compared to see which genes showed a higher expression after twenty one hours of treatment. The zero point for dox-shERK1 chips is taken as three days, as there are no earlier data available.

Not all genes encountered at the end of this analysis are significantly upregulated to become drug targets. Therefore, ten genes with the highest log-ratios from different treatments are chosen to be filtered through a Welsh t-test. Here, the null hypothesis states that the genes are not differently expressed between the two time points after treatment. Throughout the chips of dox-shERK1, only MPZL3 and BCL1A genes were detected as significantly upregulated genes. From ERK1/2-inhibitor data, the genes TRPM1, EFS, ARHGEF5, TYRP1, ETV7, AACS, PPFBIBP2, TGFBI and from the Trametinib data, the genes VANGL2, TRPM1, ETV7, DAPL1, EFS, DCT, TYRP1 are identified as significantly upregulated genes.

Another heatmap is used to visualize the upregulation of certain skin TRAs. It is clearly visible that the skin-specific genes have a considerable increase in expression on chips from a later time period after treatment, especially on those treated with ERK1/2 inhibitor or Trametinib.

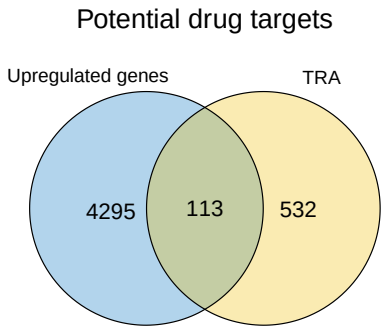


Box plots were made to visualize the expressions of the significantly upregulated genes for each treatment..



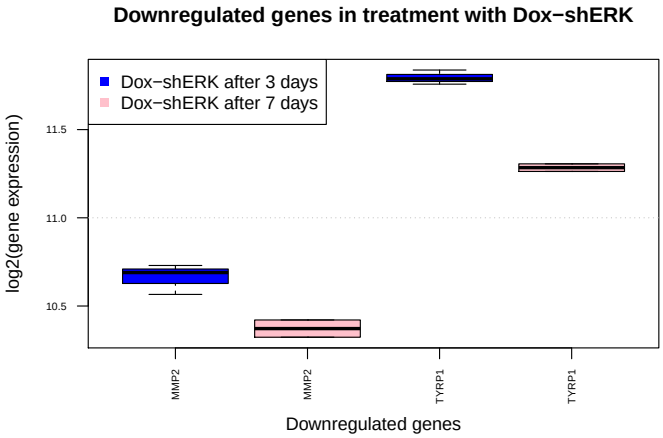
The package limma (v4.1, Ritchie *et al.*, 2015) is used as a model of comparison to the upregulated genes that were found manually using the log fold change. Limma takes the normalized data and performs statistical tests to discover any quantitative changes in the expression (Ritchie *et al.*, 2015). The results show mostly identical genes to the already found upregulated genes.

Genes that are upregulated in skin cancer and are at the same time TRAs are especially interesting, as they pose as potential drug targets. To visualize these upregulated TRAs, a venn diagram was made. Apparently there are no non-skin specific TRAs that are upregulated in skin cancer.



Downregulated Genes

The existence of significantly downregulated genes indicates that the treatment has been successful, and that these genes were inducing tumors before the treatment. Therefore, ten genes from different treatments with the lowest log-ratio are filtered out, evaluated with t-tests and visualized via box plots. The genes MMP2, TYRP1 from dox-shERK1 chips, CMSS1, KRT27, MPP4, CHTF18, POLA1, WEE1, REEP4, MFSD2A, CA12, C3of52 from ERK1/2 inhibitor chips, and RHNO1, MPP4, IRX3, MFSD2A, WEE1, CA12, C3orf52, POLA1 from Trametinib chips proved to be significantly downregulated.

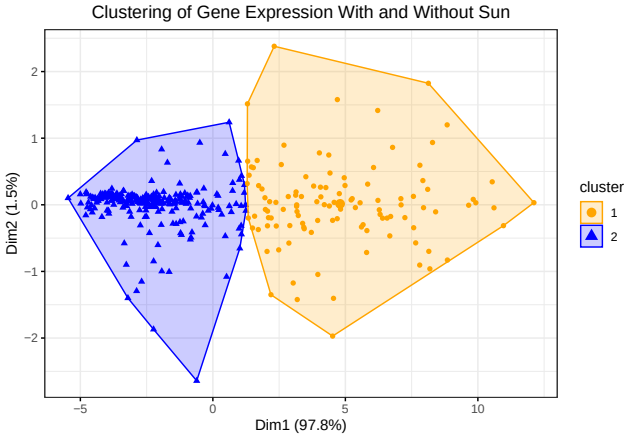


It should be noted that TYRP1 is significantly upregulated under Trametinib treatment, but significantly downregulated under dox-shERK1.

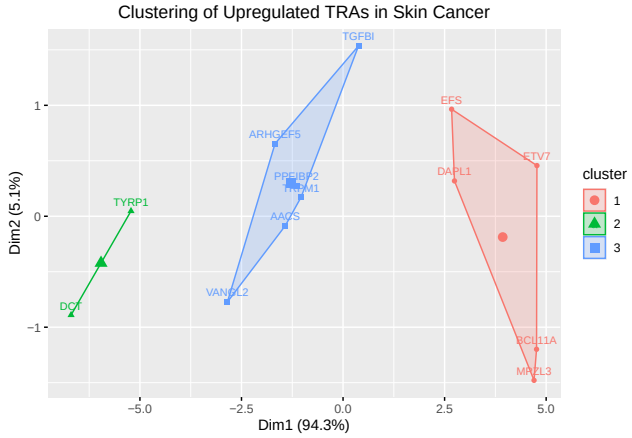
After finding the significantly upregulated and downregulated TRAs, a further category of gene expression is analysed; expression of genes with or without sun exposure. A Welsh t-test was run for each chip and a significant difference between the expression levels of the two gene groups was observed. Genes that are expressed both with and without sun exposure are removed from the data in order to acquire the genes that were only expressed under sun exposure. Four of these sun exposure related genes are also upregulated genes: ETV7, TGFBI, DAPL1 and ARHGEF5.

Clustering

The expression difference of genes exposed to sun and genes that were not exposed to sun is visualized by k-means. After running the elbow, silhouette, and one other method provided by the package NbClust (v3.0; Charrad, 2015), it is decided that two clusters are the optimal amount for this data. After clustering and plotting with the package factoextra (v1.0.3; Kassambara, 2020), two clusters are defined with an average silhouette width of 0.65, indicating that there is indeed a clusterable difference between the expressions of these two gene groups.



Same process is repeated for upregulated genes to see if there are any certain sub-groups of genes inbetween them. Three distant clusters are viewed in k-means, with an average silhouette of 0.61.



PCA and Linear Regression

Principal component analysis (PCA) is carried out on the 24h and 7d chips to find the genes that explain most of the variance in the data and to reduce the amount of dimensions. PC1 and PC2 represent 99.7% of all the variance of the melanoma dataset. Five genes with the top variance from each of these components are chosen to continue the analysis with: ANXA2, DCT, NDRG1, TYRP1, PTPRZ1 from PC1, and DSP, GJA1, PPP1R14C, TGFBI, UCN2 from PC2.

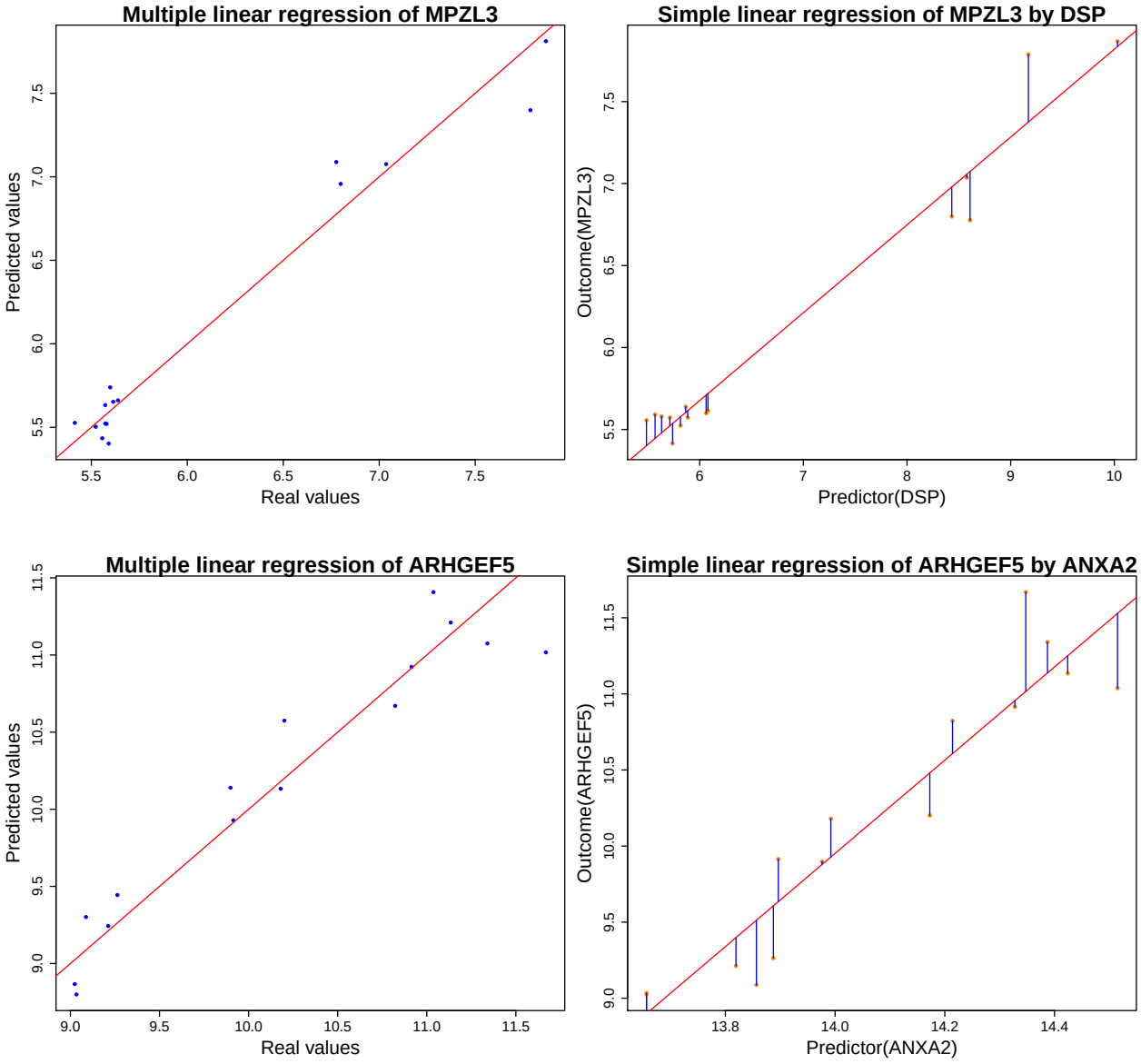
Linear regression models are used to see if there are any significant dependencies between the PC genes and the upregulated genes that would justify making a prediction of the expressions. The ability to modulate the upregulation genes by targeting the genes from PCA would open up good opportunities in developing treatments with new drug targets. Before the modeling, Pearson correlation is used to see if the upregulated genes and the PC genes have any correlations within their group. The package corrplot (v0.90; Wei, Levy & Simko, 2021) is used to visualize and calculate the correlations. All of the upregulated genes turn out to be highly correlated with each other except for MPZL3. Since the rest are highly correlated, a representative gene, ARHGEF5 is chosen to do the modeling with. Even though not as much as the upregulated genes, PC genes are found to be somewhat correlated and 3 genes as representatives are chosen. The upregulated genes MPZL3 and ARHGEF5 are going to be predicted by DSP, GJA1 and ANXA2 in the linear models.

Next, the normality of the distribution of the residuals of the upregulated genes are confirmed with a histogram and a Shapiro-Wilk test. Furthermore, the residuals of the linear models will have to be analysed in order to veerify the correctness of the linear models. It is confirmed that the residuals have a mean of approximately zero, and that they do not correlate with any of the explanatory PC genes.

Multiple linear regression is carried out between MPZL3 and the aforementioned three PC genes, and a highly significant result is observed. With an adjusted R squared value of 0.9557 and a low p-value, it is right to say that MPZL3 expression can definitely be predicted by these three genes' expression levels. DSP appears to be able to predict the expression of MPZL3 the best out of all. This has been decided with the help from the package QuantPsyc (v1.5; Fletcher, 2012).

The linear relationship of the PC genes and ARHGEF5, with an adjusted R squared value of 0.8897, and a low p-value also show a good dependency. The expression of ARHGEF5 is best represented by ANXA2 among all three PC genes.

To confirm the accuracy of the predictions made by the PC genes on the expression of the two representative upregulated genes ARHGEF5 and MPZL3, the predicted values are plotted against the real values of expression. They seem to be nearly identical. Since the gene ARHGEF5 is highly correlated with the rest of the upregulated genes, it can be concluded that the expression of almost all of the upregulated genes can be predicted by the expression profiles of the PC genes.



In conclusion, in both multiple regression models the p-value is low enough to overthrow the H0 hypothesis. The linear models show that the dependency between genes is significant enough to predict upregulated genes.

DISCUSSION

Regardless of treatment type and time variation between different chips, GPNMB and ANXA2 proved to be highest expressed in all of them. ANXA2 is known as a biomarker in cancers, it is already seen as an ideal target for cancer therapy because of its crucial functions in cancer cell proliferation, invasion and metastasis (Wang *et al.*, 2019). GPNMB has been formerly identified as an aggressive pro-metastatic protein that is highly expressed in many tumor types, including melanoma (Taya & Hammes, 2018). The fact that brings these two together is that they are both responsible for the genesis of melanosomes (Setaluri & Jayanthy, 2013; Delevoye *et al.*, 2015). Melanoma develops from melanocytes, which normally assemble melanosomes, organelles in mammals responsible for melanin production (Setaluri & Jayanthy, 2013). It makes sense that these two genes related to the assembly and maturation of melanosomes are found to be highly expressed in melanoma.

The two other important genes TYRP1 and DCT found in this analysis, which are also involved in the development of melanosomes (Setaluri & Jayanthy, 2013), are both upregulated and principal component genes at the same time. DCT, also known as TYRP2, encodes proteins that play roles in the melanin synthesis, just as TYRP1 (Setaluri & Jayanthy, 2013). The upregulation and high expression of these melanosome related genes indicates that a therapy focusing to direct melanosome functions would be effective against melanoma.

Microphthalmia-associated transcription factor (MITF) is a key signal regulator with the subset target genes GPNMB and TYRP1 (Rose *et al.*, 2016). MITF plays important roles in melanosomal function and pigmentation, but it is also required for sustaining the BRAF-mutant melanoma and supporting the melanoma cell proliferation (Rose *et al.*, 2016). It has been observed before that MITF related genes are upregulated under certain BRAF and MEK inhibitors, such as Trametinib (Rose *et al.*, 2016). This would explain the higher expression results of TYRP1 after treatment. It is upregulated not only because it is needed for the maintenance of the melanoma cells, but also due to the Trametinib treatment, which is a MEK1/2 inhibitor. The other two treatments are inhibitors of ERK, which partially successfully seem to have worked, because treatment with Doxycyclin-shERK managed to downregulate TYRP1.

When the upregulated genes are clustered, three groups are to be seen. The genes in the first cluster seem to be mostly prognostic markers (Zheng *et al.*, 2021; Zhou *et al.*, 2020; Neumann *et al.*,). MPZL3 is the most important one between them, as it did not correlate with any of the other upregulated genes in this data. An overexpression of this gene is usually associated with increased CD8+ immune infiltration and microsatellite instability, leading to better prognosis (Zheng *et al.*, 2021). The second cluster (green) is made up of two genes that are closely related each other; TYRP1 and DCT (TYRP2) code for the same family of proteins and take part in the melanosome biosynthesis as mentioned before (Kobayashi *et al.*, 1998).

In the third cluster (blue) a couple of genes are found to be related to each other. The TGFBI protein regulates cyclin-dependent kinases in melanomas, which are known to be in charge of the cell cycle, and allows a high metastatic potential to the melanoma (Lauden *et al*, 2014). It is induced by TGFß, proven to be playing a role in the epithelial-mesenchymal transition (EMT) of cells (Pérez *et al.*, 2017). EMT is how epithelial tumor cells usually obtain malignant properties (Komiya *et al.*, 2016). ARHGEF5 is also activated by TGFß and helps execute malignancy through EMT (Komiya *et al.*, 2016). However, a melanoma cell does not go through EMT as it is not derived from epithelial cells (Noguchi *et al.*, 2017). The proliferative cells go into their invasive stage by a phenotype switching when the MITF is lowly expressed, meaning that an EMT-like process happens when the high levels of MITF in the cell drop (Noguchi *et al.*, 2017).

Data in this project on the other hand shows that MITF related genes and EMT related genes have been both upregulated and are even under the third cluster, contradicting the former information. Two other genes, TRPM1 and PPFIBP2, from the same cluster also show a relation to the already known melanoma marker MITF (Mohammadi *et al.*, 2018). Former studies have revealed that TRPM1 is an indicator of melanocytic differentiation and is regulated by MITF (Miller *et al.*, 2004), and that PPFIBP2 is a part of the MITF sub-network (Mohammadi *et al.*, 2018). This makes the fact that they are clustered together more of a wonder.

They are all related because MITF has been found to indirectly induce the expression of TGFBI in melanocytes (Lavelle *et al.*, 2020). It has been shown in a former study that the forced high expression of MITF in melanocytes resulted in some defects in the TGFß signaling pathway, making it upregulated (Lavelle *et al.*, 2020). This causes the upregulation of the EMT related gene TGFBI (Lavelle *et al.*, 2020), and naturally ARHGEF5. All four of the upregulated genes in the third cluster (TRPM1, PPFIBP2, TGFBI, ARHGEF5) are upregulated under the ERK1/2 inhibitor and Trametinib, which has inducing effects on the expression of MITF related genes as mentioned before (Rose *et al.*, 2016). High ERK activity lowers the MITF protein levels (Wellbrock & Arozarena, 2015), meaning an inhibitor of ERK also has an enhancing effect on the upregulation of the MITF related genes. The existence of ERK1/2-inhibitor and Trametinib push on the already high expression of MITF in melanoma cells, resulting in activation of EMT genes by TGFß, TGFBI and ARHGEF5, are upregulated. As a concluding hypothesis it can be stated that the high expression of MITF not only gives the melanoma cells proliferative, but also invasive characteristics, which makes it an even more interesting target for melanoma to research on.

There is not enough related information found on the other two genes AACS and VANGL2 from the same cluster to build up any theories about their relationship. However, they both seem to be related to breast cancer (Smith *et al.*, 2019; Chou *et al.*, 2013).

Another significant result is the discovery of the four upregulated genes that are only expressed under the sun. DAPL1, TGFBI are genes upregulated by the aforementioned MITF (Ma *et al.*, 2018; Lavelle *et al.*, 2020) and ARHGEF5 aids in the invasive potential of cancer cells (Kuroiwa *et al.*, 2011). The knowledge that such genes, supporting the development of melanoma, have only been upregulated under sun exposure confirms the already commonly known fact that the exposure to sun is a trigger for skin cancer.

The fourth upregulated gene under sun exposure, ETV7, codes for a transcription factor, which is usually seen to be downregulated in melanoma patients and correlates with a poor prognosis (Qu *et al.*, 2020). However, ETV7 is seen to be upregulated in our data under Trametinib and ERK1/2 inhibitor, suggesting that it might have a role in attacking the tumor microenvironment, leading to a better prognosis. The fact that this gene is expressed under the sun suggests that it might be a natural response to the tumorigenic genes, upregulated under sun exposure. The treatments help upregulate it further to fight tumor proliferation. The relationship between ETV7 and the three genes mentioned above should be looked into with more detail, as a treatment upregulating its expression seems to be possibly effective against melanoma.

It is important to pay attention to the outcome that almost all of the upregulated genes are either regulated by MITF or are from the data that was exposed to the sun. This indicates that the high expression of the genes that are related to MITF are as effective as exposure to sun when it comes to helping create or supporting the melanoma. This means that MITF related genes make extremely good drug targets for melanoma. Since all of the upregulated genes analysed on are human skin specific TRA genes, a new treatment where the expression of the genes mentioned above are targeted, would not harm healthy tissues. This opens countless opportunities for future research and treatment of melanomas, especially because melanoma is highly metastatic.

Through the opposite expression of TYRP1, and varying up- and downregulated genes between chips of different treatments, the different effects of MEK and ERK inhibitors on melanoma have been shown. As it is observed that MITF related genes are upregulated under treatment with MAPK pathway inhibitors, the usage of ERK/MEK inhibitors on melanoma should be questioned. Furthermore, WEE1, an important G2-M checkpoint regulator in the cell cycle, is downregulated under these treatments, which results in an increase in tumor aggressiveness (Bhattacharya *et al.*, 2012). Therefore, it should be further researched and discussed on if these inhibitors should be used on melanoma. If yes, a combination of other drugs are needed as treatment to inhibit the accelerating effects of MAPK inhibitors on tumoregenic or melanosome related genes.

The strong linear dependency found between the expression of the gene DSP and MPZL3 also raises new questions requiring additional research. To be able to predict MPZL3 is important for characterizing the cancer; identifying it would help to see if the patient has a good prognosis or not. DSP seems to be fit for this prediction. It codes for a desmosomal protein, which helps human melanoma metastases to escape the immune system (Salerno *et al.*, 2016). Some desmosomal proteins are used to identify melanomas, even when there are not any immune signatures seen (Salerno *et al.*, 2016).

The significant linear relationship between the genes ARHGEF5 and ANXA2 mentioned above might also be useful for treatment development as ANXA2 is an already widely known cancer biomarker (Wang *et al.*, 2019). Considering ARHGEF5 strongly correlates with and represents almost all of the upregulated genes, all of them except for MPZL3 can be predicted by this already popular drug target. Besides the prediction of upregulated genes, PC genes can also be considered as potential drug targets, since they have shown a correlation between each other.

Not all of the upregulated genes found in this dataset were informative, as there has not been much research done on them before. More research is needed for further analysis. However,the findings that melanosome and MITF related genes, and genes highly expressed under sun exposure strongly influence the melanoma cells, and that there is a strong linear relationship between the upregulated and the high variance genes set a significant base for future research in new treatments for melanoma.

LITERATURE

Alqathama A. (2020). BRAF in malignant melanoma progression and metastasis: potentials and challenges.

American journal of cancer research, 10(4), 1103–1114.

Bhattacharya, A., Schmitz, U., Wolkenhauer, O. *et al.*, (2013). Regulation of cell cycle checkpoint kinase WEE1 by miR-195 in malignant melanoma. Oncogene 32, 3175-3183. https://doi.org/10.1038/onc.2012.324

Chou, HL., Yao, CT., Su, SL. *et al.* (2013).Gene expression profiling of breast cancer survivability by pooled cDNA microarray analysis using logistic regression, artificial neural networks and decision trees. BMC Bioinformatics 14, 100. https://doi.org/10.1186/1471-2105-14-100

Delevoye, C., Heiligenstein, X., Ripoll, L., Gilles-Marsens, F., Dennis, M. K., Linares, R. A., Derman, L., Gokhale, A., Morel, E., Faundez, V., Marks, M.S., Raposo, G. (2015). BLOC-1 Brings Together the Actin and Microtubule Cytoskeletons to Generate Recycling Endosomes, Current Biology, Vol 26(1), 1-13, https://doi.org/10.1016/j.cub.2015.11.020

Dinkelacker, 2007. A database of genes that are expressed in a tissue-restricted manner to analyse promiscuous gene expression in medullary thymic epithelial cells. Diplomarbeit, Albert-Ludwigs-Universitaet, Freiburg, Germany.

Dinkelacker, 2019. Chromosomal clustering of tissue restricted antigens, Dissertation, University Heidelberg, Germany.

Gruber, P., Zito, P. M. (2021). Jan, Skin Cancer, Treasure Island (FL): StatPearls Publishing; Available from: https://www.ncbi.nlm.nih.gov/books/NBK441949/

Klein, L., Kyewski, B., Allen, P. *et al.* (2014). Positive and negative selection of the T cell repertoire: what thymocytes see (and don't see). Nat Rev Immunol 14, 377–391. https://doi.org/10.1038/nri3667

Kobayashi, T., Imokawa, G., Bennett, D.C., and Hearing, V.J. (1998). Tyrosinase Stabilization by Tyrp1 (the brown Locus Protein). Journal of Biological Chemistry 273, 31801-31805. https://doi.org/10.1074/jbc.273.48.31801.

Komiya, Y., Onodera, Y., Kuroiwa, M., Nomimura, S., Kubo, Y., Nam, J. M., Kajiwara, K., Nada, S., Oneyama, C., Sabe, H., & Okada, M. (2016). The Rho guanine nucleotide exchange factor ARHGEF5 promotes tumor malignancy via epithelial-mesenchymal transition. Oncogenesis, 5(9), e258. https://doi.org/10.1038/oncsis.2016.59

Kuroiwa, M., Oneyama, C., Nada, S., & Okada, M. (2011). The guanine nucleotide exchange factor Arhgef5 plays crucial roles in Src-induced podosome formation. Journal of cell science, 124(Pt 10), 1726–1738. https://doi.org/10.1242/jcs.080291

Kyewski, B., Derbinski, J. (2004) Self-representation in the thymus: an extended view. Nat Rev Immunol 4, 688–698. https://doi.org/10.1038/nri1436

Lauden, L., Siewiera, J., Boukouaci, W., Ramgolam, K., Mourah, S., Lebbe, C., Charron, D., Aoudjit, F., Jabrane-Ferrat, N., Al-Daccak, R. (2014). TGF-ß-Induced (TGFBI) Protein in Melanoma: A Signature of High Metastatic Potential, Journal of Investigative Dermatology, Vol:134(6), 1675-1685, ISSN 0022-202X, https://doi.org/10.1038/jid.2014.20.

Lavelle, T. J., Alver, T. N., Heintz, K. M., Wernhoff, P., Nygaard, V., Nakken, S., Øy, G. F., Bøe, S. L., Urbanucci, A., & Hovig, E. (2020). Dysregulation of MITF Leads to Transformation in MC1R-Defective Melanocytes. Cancers, 12(7), 1719. https://doi.org/10.3390/cancers12071719

Leiter, U., & Garbe, C. (2008). Epidemiology of melanoma and nonmelanoma skin cancer–the role of sunlight. Advances in experimental medicine and biology, 624, 89–103. https://doi.org/10.1007/978-0-387-77574-6_8

Linares, M. A., Zakaria, A., & Nizran, P. (2015). Skin Cancer. Primary care, 42(4), 645–659. https://doi.org/10.1016/j.pop.2015.07.006

Ma, X., Hua, J., Zheng, G., Li, F., Rao, C., Li, H., Wang, J., Pan, L., & Hou, L. (2018). Regulation of cell proliferation in the retinal pigment epithelium: Differential regulation of the death-associated protein like-1 DAPL1 by alternative MITF splice forms. Pigment cell & melanoma research, 31(3), 411–422. https://doi.org/10.1111/pcmr.12676

Miller, A. J., Du, J., Rowan, S., Hershey, C. L., Widlund, H. R., & Fisher, D. E. (2004). Transcriptional regulation of the melanoma prognostic marker melastatin (TRPM1) by MITF in melanocytes and melanoma. Cancer research, 64(2), 509–516. https://doi.org/10.1158/0008-5472.can-03-2440

Mohammadi, S., Ravindra, V., Gleich, D. F., & Grama, A. (2018). A geometric approach to characterize the functional identity of single cells. Nature communications, 9(1), 1516. https://doi.org/10.1038/s41467-018-03933-2

Neumann, L. C., et al. 2011. EFS shows biallelic methylation in uveal melanoma with poor prognosis as well as tissue-specific methylation. BMC Cancer 11:380. doi:10.1186/1471-2407-11-380

Noguchi, K., Dalton, A. C., Howley, B. V., McCall, B. J., Yoshida, A., Diehl, J. A., & Howe, P. H. (2017). Interleukin-like EMT inducer regulates partial phenotype switching in MITF-low melanoma cell lines. PloS one, 12(5), e0177830. https://doi.org/10.1371/journal.pone.0177830

Opitz, L., Salinas-Riester, G., Grade, M., Jung, K., Jo, P., Emons, G., Ghadimi, B. M., Beissbarth, T., & Gaedcke, J. (2010). Impact of RNA degradation on gene expression profiling. BMC medical genomics, 3, 36. https://doi.org/10.1186/1755-8794-3-36

Pérez, L. *et al* (2017). Endothelial-to-mesenchymal transition: Cytokine-mediated pathways that determine endothelial fibrosis under inflammatory conditions. Cytokine & Growth Factor Reviews 33, 41-54. https://doi.org/10.1016/j.cytogfr.2016.09.002.

Potrony, M., Badenas, C., Aguilera, P., Puig-Butille, J. A., Carrera, C., Malvehy, J., Puig, S. (2015). Update in genetic susceptibility in melanoma. Annals of translational medicine, 3(15), 210. https://doi.org/10.3978/j.issn.2305-5839.2015.08.11

Qu, H., Zhao, H., Zhang, X., Liu, Y., Li, F., Sun, L., & Song, Z. (2020). Integrated Analysis of the ETS Family in Melanoma Reveals a Regulatory Role of ETV7 in the Immune Microenvironment. Frontiers in immunology, 11, 612784. https://doi.org/10.3389/fimmu.2020.612784

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., & Smyth, G. K. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic acids research, 43(7), e47. https://doi.org/10.1093/nar/gkv007

Rose, A. A., Annis, M. G., Frederick, D. T., Biondini, M., Dong, Z., Kwong, L., Chin, L., Keler, T., Hawthorne, T., Watson, I. R., Flaherty, K. T., Siegel, P. M. (2016). MAPK Pathway Inhibitors Sensitize BRAF-Mutant Melanoma to an Antibody-Drug Conjugate Targeting GPNMB, DOI: 10.1158/1078-0432.CCR-16-1192

Rosenberg S. A. (1999). A new era for cancer immunotherapy based on the genes that encode cancer antigens. Immunity, 10(3), 281–287. https://doi.org/10.1016/s1074-7613(00)80028-x

Salerno, E. P., Bedognetti, D., Mauldin, I. S., Deacon, D. H., Shea, S. M., Pinczewski, J., Obeid, J. M., Coukos, G., Wang, E., Gajewski, T. F., Marincola, F. M., & Slingluff, C. L., Jr (2016). Human melanomas and ovarian cancers overexpressing mechanical barrier molecule genes lack immune signatures and have increased patient mortality risk. Oncoimmunology, 5(12), e1240857. https://doi.org/10.1080/2162402X.2016.1240857

Setaluri, V., Jayanthy, A. (2013). Coat Color Mutations, Animals, Brenner's Encyclopedia of Genetics, Second Edition, Academic Press, 58-60, ISBN 9780080961569, https://doi.org/10.1016/B978-0-12-374984-0.00275-8.

Smith, P., Godde, N., Rubio, S. *et al.* (2019). VANGL2 regulates luminal epithelial organization and cell turnover in the mammary gland. Sci Rep 9, 7079. https://doi.org/10.1038/s41598-019-43444-8

Smith, R. A., Mettlin, C. J., Eyre, H. (2003). Melanoma and Nonmelanoma Skin Cancer, Holland-Frei Cancer Medicine. 6th edition. Hamilton (ON): BC Decker; https://www.ncbi.nlm.nih.gov/books/NBK13764/

Taya, M., Hammes, S. R. (2018). Glycoprotein Non-Metastatic Melanoma Protein B (GPNMB) and Cancer: A Novel Potential Therapeutic Target. Steroids, 133, 102–107. https://doi.org/10.1016/j.steroids.2017.10.013

Volz, A., Korge, B. P., Compton, J. G., Ziegler, A., Steinert, P. M., & Mischke, D. (1993). Physical mapping of a functional cluster of epidermal differentiation genes on chromosome 1q21. Genomics, 18(1), 92–99. https://doi.org/10.1006/geno.1993.1430

Wang, T., Wang, Z., Niu, R., & Wang, L. (2019). Crucial role of Anxa2 in cancer progression: highlights on its novel regulatory mechanism. Cancer biology & medicine, 16(4), 671–687. https://doi.org/10.20892/j.issn.2095-3941.2019.0228

Wellbrock, C., & Arozarena, I. (2015). Microphthalmia-associated transcription factor in melanoma development and MAP-kinase pathway targeted therapy. Pigment cell & melanoma research, 28(4), 390–406. https://doi.org/10.1111/pcmr.12370

Zeiser, R. (2014). Trametinib. Recent results in cancer research. Fortschritte der Krebsforschung. Progrès dans les recherches sur le cancer, 201, 241–248. https://doi.org/10.1007/978-3-642-54490-3_15

Zheng, J., et al. 2021. Biomarkers Associated with CD8+ T Cell Infiltration in Childhood AML. https://doi.org/10.21203/rs.3.rs-659529/v1

Zhou, J., et al. 2020. BCL11A Promotes the Progression of Laryngeal Squamous Cell Carcinoma. Frontiers in oncology 10:375. doi:10.3389/fonc.2020.00375