The role of skin-specific genes in skin cancer

Project Proposal

Group 3: Pinar Cavus, Binnur Özay, Christofer Richard, Denisa Neacsu

Supervisors: Maria Dinkelacker, Nils Mechtel
Molecular Biotechnology B.Sc, Ruprecht Karls University of Heidelberg

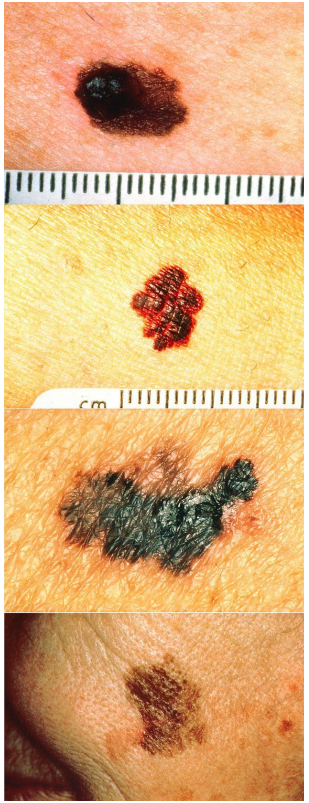
Background

Skin cancer -> Most common type of cancer in humans

- **Melanoma** (Only 2% of all skin cancers, but most deaths!)
- **Non-melanoma**

Risk factors:

- ❏ Moles
- ❏ UV exposure
- ❏ Genetic background



https://en.wikipedia.org/wiki/Melanoma

Background

Problem : Cancer escaping immune surveillance

How?

- **Tissue restricted antigens (TRA)** are upregulated in cancer cells

What can be done:

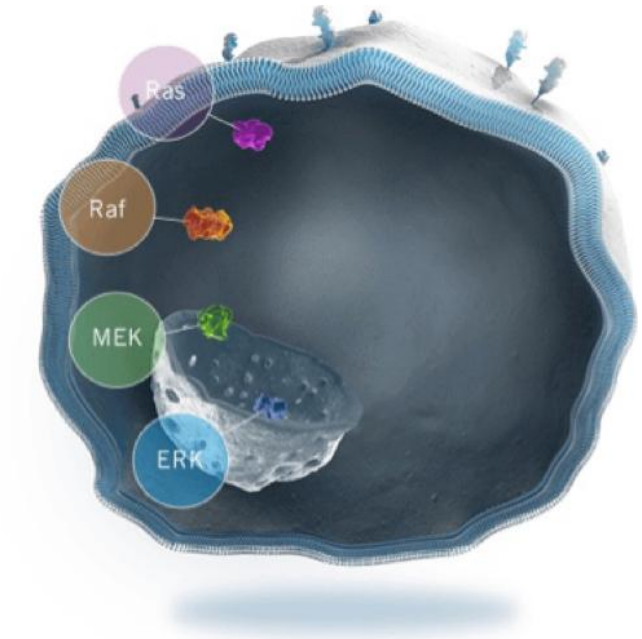
- TRA as drug targets **(Cancer immunotherapy!)**

Background

The significance of MAPK-pathway in melanoma development:

- Mutated BRAF can lead to abnormal MAPK-signaling

  → 50% of all melanoma cases



https://www.genentechoncology.com/pathways/cancer-tumor-targets/mapk.html#:~:text=The%20mitogen%2Dactivated%20protein%20kinase,%2C%20cellular%20growth%2C%20and%20survival.&text=Abnormal%20MAPK%20signaling%20may%20lead,proliferation%20and%20resistance%20to%20apoptosis.&text=Research%20into%20the%20MAPK%20pathway,be%20important%20in%20some%20cancers.

What are we working with?

WM266-4 human melanoma cell line with the genotype BRAF V600D

Treatment with:

- Trametinib: MEK-Inhibitor
- ERK1/2-Inhibitor
- doxycycline-shERK1: silencing ERK1

Extracted mRNA from cells ->->-> labeled fragmented cRNA, hybridized on microarrays

Data Description

NUMERICAL DATA: Expression Dataframe

Microarrays, Skin Cancer and Breast Cancer GSE27830

Quantification of the gene expression levels

CATEGORICAL DATA: TRA Vector

Nominal values

-> skin specific TRA gene names

```
 [1] "SCGB2A2"     "SCGB1D2"     "KRT1"       "KRT10"      "LOR"
[11] "ALS2CL"      "AP001267.1"  "APCDD1"     "ARHGEF19"   "ARHGEF37"
[21] "CAMSAP3"     "CD44"        "CDC42BPG"   "CENPP"      "CXCL14"
[31] "FAM110A"     "FAM57A"      "FAM83H"     "FBXO27"     "FBXW7"
```

Our Main Objective

Identify **upregulated genes in skin cancer**, compare with TRA data to find out if any **TRA genes match** with our results and has the potential to be a drug target.

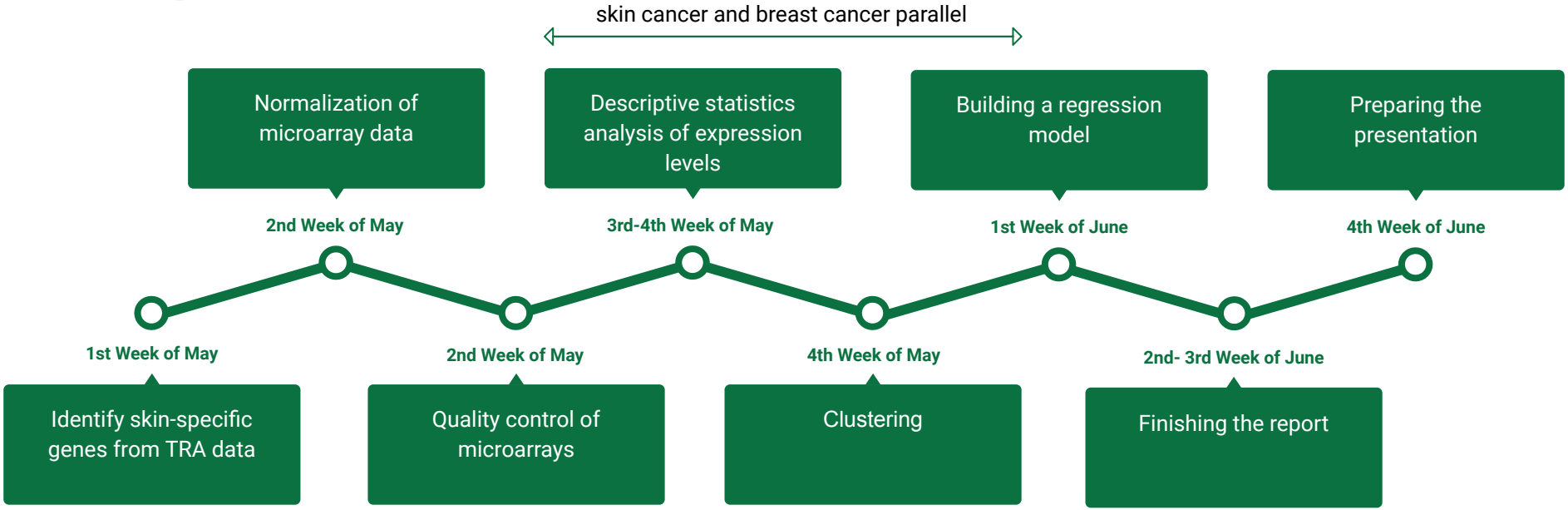
Further questions to answer along the way

After identifying upregulated genes:

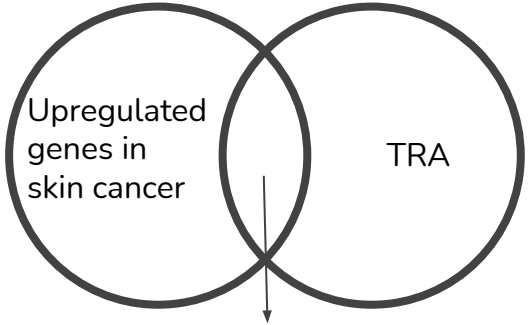
- Are our identified genes **mostly from the sun exposure data**?
- In **which chromosome** are the identified upregulated genes mostly localised?
- Are there any **non-skin specific genes** that are upregulated in skin cancer?

Do the **efficacy of the drugs** vary? Comparing the expression intensity from our dataset, connecting our information with the duration and type of therapy used in each chip.

Organisation



Methods



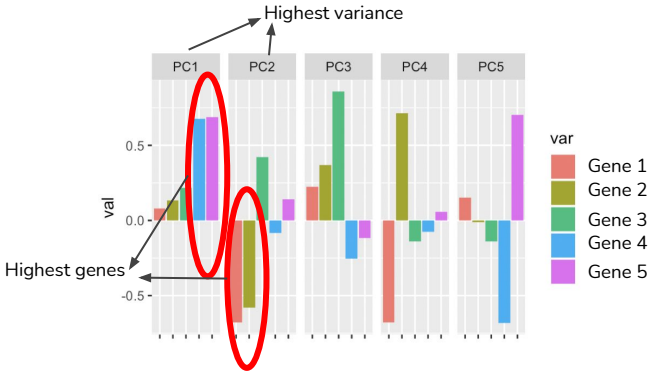
1) **Analysis with descriptive statistics:**

- Heatmaps (Visualization of skin-specific gene expression)
- Barplot (Distribution in chromosomes)
- Venn diagram
- Box plots (Expression of skin-specific genes in skin cancer):
  - Grouped by different treatments & on different time periods
  - Sun VS no sun exposure

2) **Clustering: k-means**

- Elbow method, silhouette → Optimal amount of clusters
- Observe patterns

Methods



3) **Dimension reduction: PCA**

- Find the principal components that explain most of the variance of our skin cancer dataset
- Identify genes from these PCs that contribute the most to the variance

4) **Linear modeling:**

- Regression Model
  - Predict relationship of identified genes with drug target genes (ERK, etc.)
  - Example: Expression of non-skin-specific gene used to predict ERK expression
- F-test → Best prediction model

What did we do until now?

Filter TRA tables -> Human skin-specific TRAs

+

Normalization of melanoma microarray data set

+

Gene Expression melanoma data frame

↓

Filtering skin-specific TRAs in the gene expression melanoma data frame

Melanoma microarray chip names →

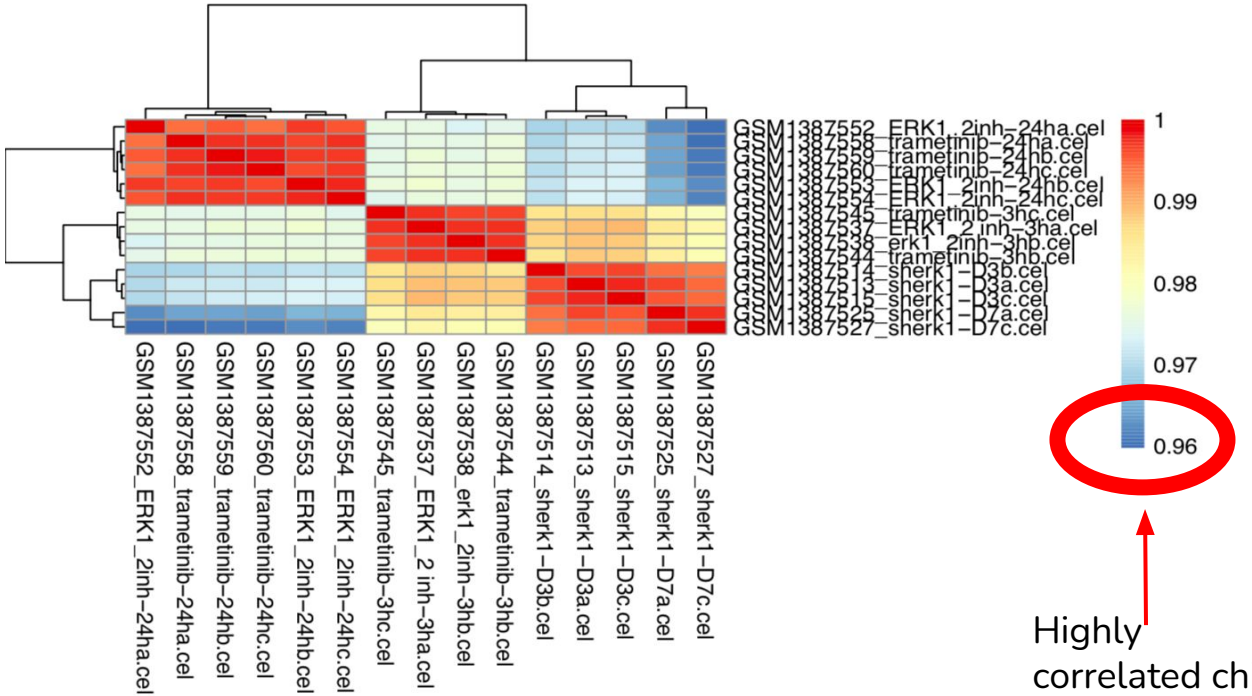
Gene names ↓

	GSM1387513_sherk1-D3a	GSM1387514_sherk1-D3b	GSM1387515_sherk1-D3c	GSM1387525_sherk1-D7a	GSM1387527_sherk1-D7c
A2ML1	5.712456	5.840425	5.822129	5.624177	5.757851
AACS	9.700679	9.577292	9.493689	9.584939	9.575073
AADACL2	5.121151	5.200550	5.171339	5.215717	5.208109
ABCA12	5.758000	5.651394	5.800670	5.742912	5.720462
ABHD12B	5.500152	5.595327	5.463102	5.655431	5.693763
ABHD5	8.863773	8.782844	8.585597	9.031559	8.810857
ACAP3	7.086765	7.164285	7.056074	7.039174	7.282637
ACER1	6.375048	6.443568	6.299403	6.277664	6.364174
ACVR2A	9.783693	10.078800	9.801101	9.907579	9.795192
ADAM15	8.243245	8.344145	8.270667	8.240065	8.204795
ADGRF2	5.942704	5.845375	5.860404	5.860404	5.829077
ADGRF4	5.806937	5.718587	5.629339	5.829695	5.662445
ADRB2	6.913961	7.224369	6.972564	7.592311	7.701687
AHNAK	8.234221	8.214370	8.047431	8.293025	8.129481
AJUBA	8.668198	8.668334	8.826362	8.897877	8.956261

Heatmap melanoma microarrays

Quality control ✅

- MeanSdPlot
- RNA degradation
- Boxplots
- Heatmap to show correlation



Highly correlated chips => qualitative microarrays

Breast cancer data set

Question: Are there any up-regulated skin specific genes in breast cancer?

➔ Possible drug target

Plan:

- Normalize data
- Skin specific TRAs in breast cancer vs melanoma
    - Examine overall (di)similarity of gene expression patterns
        - Box plots
        - Heatmaps
        - Distributions
        - Clustering etc.

Literature

Dinkelacker, 2007. A database of genes that are expressed in a tissue-restricted manner to analyse promiscous 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.

GEO: https://www.ncbi.nlm.nih.gov/geo/geo2r/?acc=GSE57721

GTEX 2013, The Genotype-Tissue Expression (GTEx) project. GTEx Consortium. Nat Genet. 2013 Jun;45(6):580-5. https://pubmed.ncbi.nlm.nih.gov/23715323/

GTEX 2015, Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. GTEx Consortium. Science. 2015 May 8;348(6235):648-60. https://pubmed.ncbi.nlm.nih.gov/25954001/

Herrmann, C. 2020. Finding structure in the data. Moodle course for M3: Bioinfo, University of Heidelberg, Germany.

Kidger A.M., Sipthorp J., Cook S.J., 2018 ERK1/2 inhibitors: New weapons to inhibit the RAS-regulated RAF-MEK1/2-ERK1/2 pathway, Pharmacology & Therapeutics, Volume 187, 2018, Pages 45-60, ISSN 0163-7258, https://doi.org/10.1016/j.pharmthera.2018.02.007.

Kyewski et al. 2004, Self-representation in the thymus: an extended view, Nat. Rev. Immunol. 2004 Sep, 4 (9) 699-698. https://pubmed.ncbi.nlm.nih.gov/15343368/

Literature

Lattin et al. 2008, Expression analysis of G Protein-Coupled Receptors in mouse macrophages. Immunome Res. 2008 Apr 29;4:5. https://pubmed.ncbi.nlm.nih.gov/18442421/

Linares et al. 2015. Skin Cancer. Prim Care. 2015 Dec;42(4):645-59. https://pubmed.ncbi.nlm.nih.gov/26612377/

Rosenberg et al. 1999. A new era for cancer immunotherapy based on the genes that encode cancer antigens. Immunity. 1999 Mar;10(3):281-7. https://pubmed.ncbi.nlm.nih.gov/10204484/

Roth et al. 2006, Gene expression analyses reveal molecular relationships among 20 regions of the human CNS. Neurogenetics. 2006 May;7(2):67-80. https://pubmed.ncbi.nlm.nih.gov/16572319/

Su et al. 2002, Large-scale analysis of the human and mouse transcriptomes. Proc Natl Acad Sci U S A. 2002 Apr 2;99(7):4465-70. https://pubmed.ncbi.nlm.nih.gov/11904358/

Su et al. 2004, A gene atlas of the mouse and human protein-encoding transcriptomes. Proc Natl Acad Sci U S A. 2004 Apr 20;101(16):6062-7. https://pubmed.ncbi.nlm.nih.gov/15075390/

Uhlén et al. 2015, Proteomics. Tissue-based map of the human proteome. Science. 2015 Jan 23;347(6220):1260419. https://pubmed.ncbi.nlm.nih.gov/25613900/