

The role of skin-specific genes in skin cancer

Project Proposal

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

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

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.

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

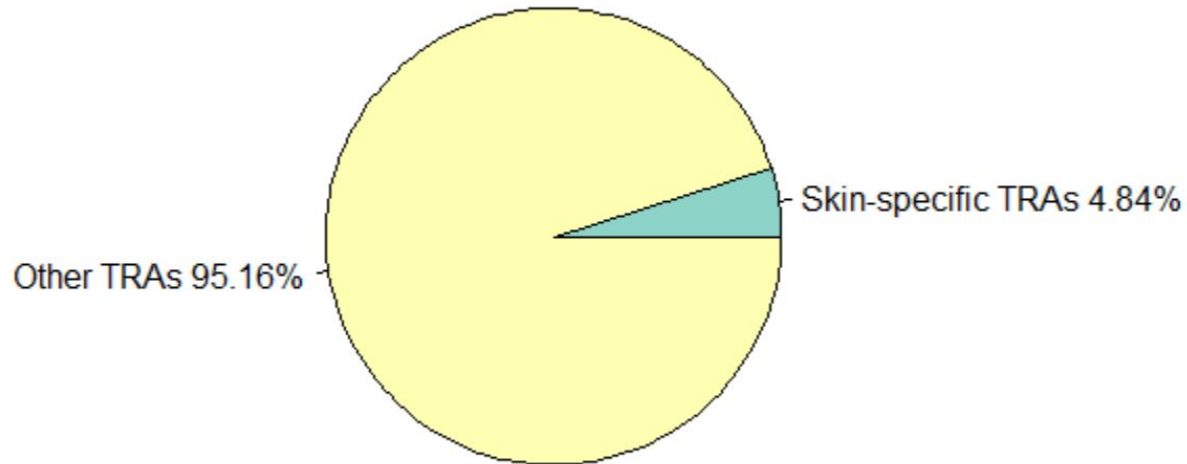
What can be done:

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



Not many skin specific TRA genes

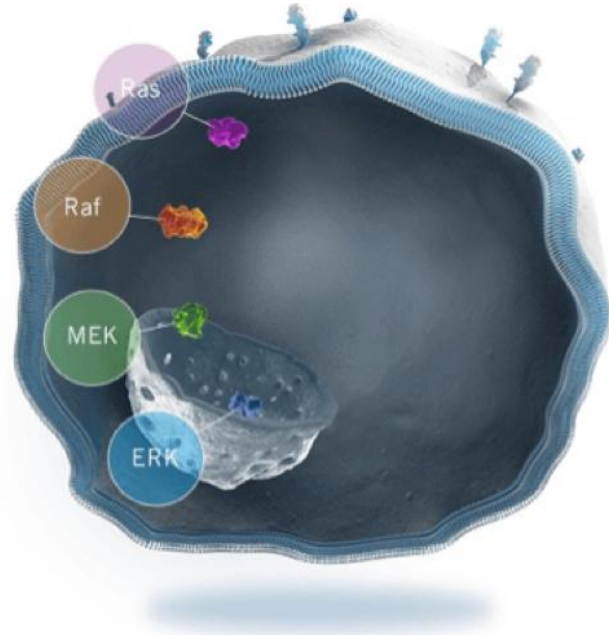
Pie Chart skin specific TRAs



Background

The significance of MAPK-pathway in melanoma development:

- mutated BRAF can lead to abnormal MAPK-signaling
→ 50% of all melanoma cases



<https://www.genentechonology.com/pathways/cancer-tumor-targets/mapk.html>.



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



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

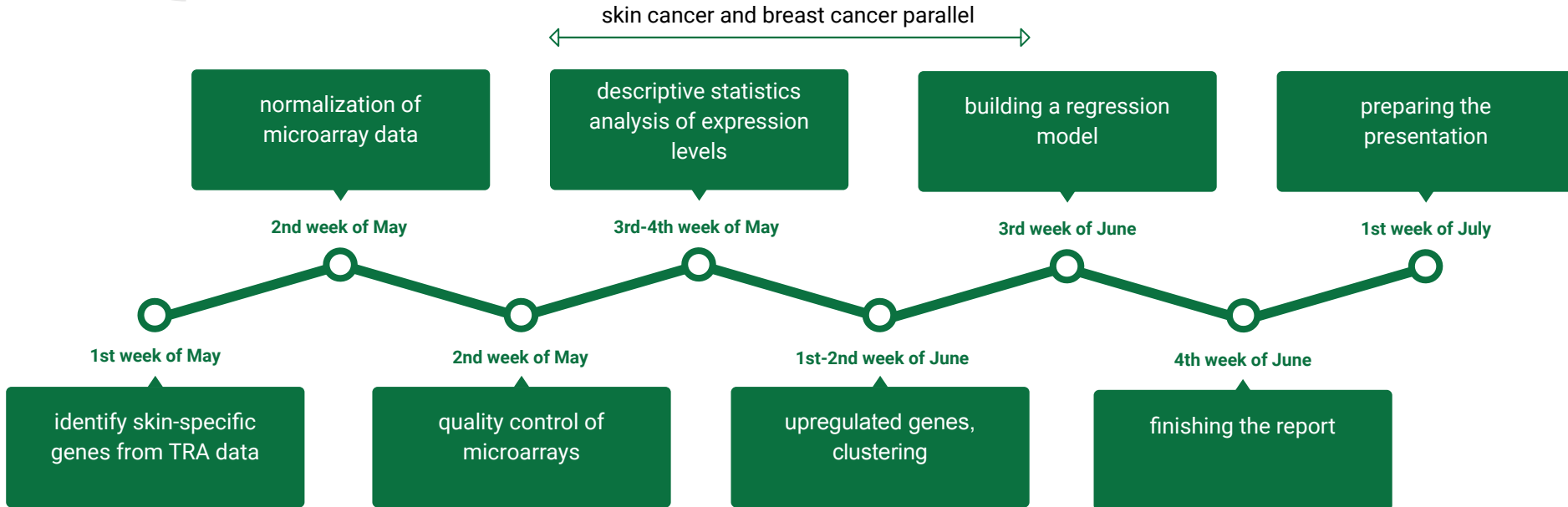
After identifying upregulated genes:

- Are our identified genes **mostly from the sun exposure data**?
- In **which chromosome** are the identified upregulated genes mostly localized?
- 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





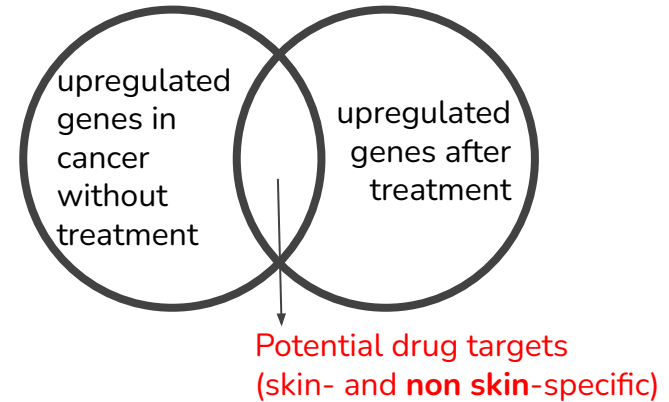
Methods

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





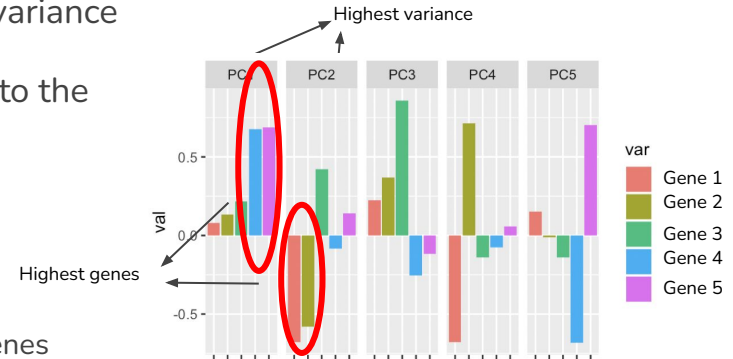
Methods

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

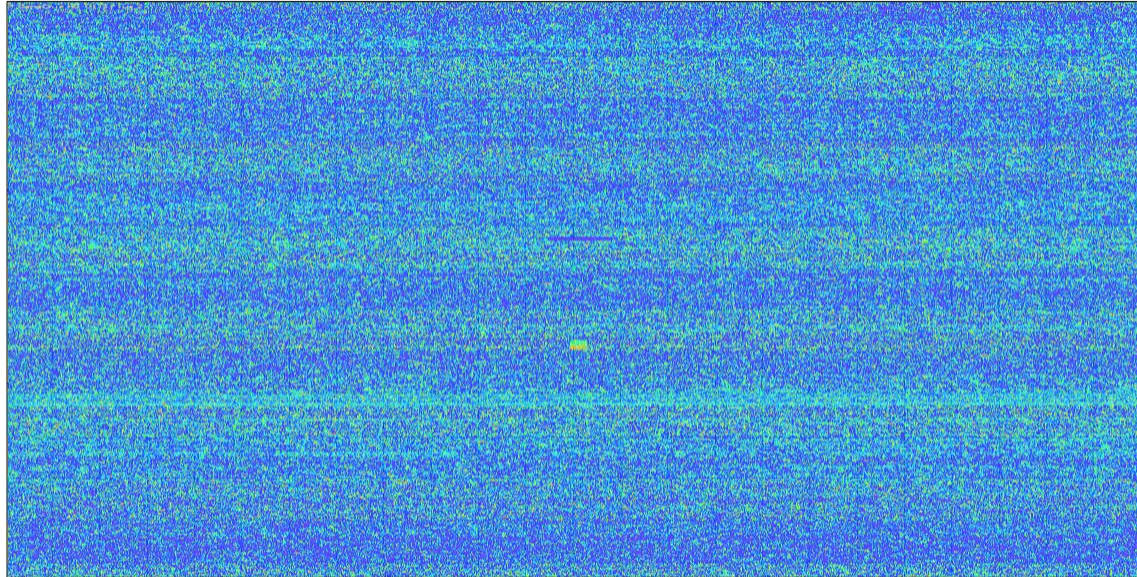


	^ GSM1387513_sherk1- D3a ^	^ GSM1387514_sherk1- D3b ^	^ GSM1387515_sherk1- D3c ^	^ GSM1387525_sherk1- D7a ^	^ GSM1387527_sherk1- D7c ^
A2ML1	5.712456	5.840425	5.822129	5.624177	5.757851
AACS	9.700679	9.577292	9.493689	9.584939	9.575073
AADACL2	5.121151	5.200550	5.171339	5.215717	5.208109
ABCA12	5.758000	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

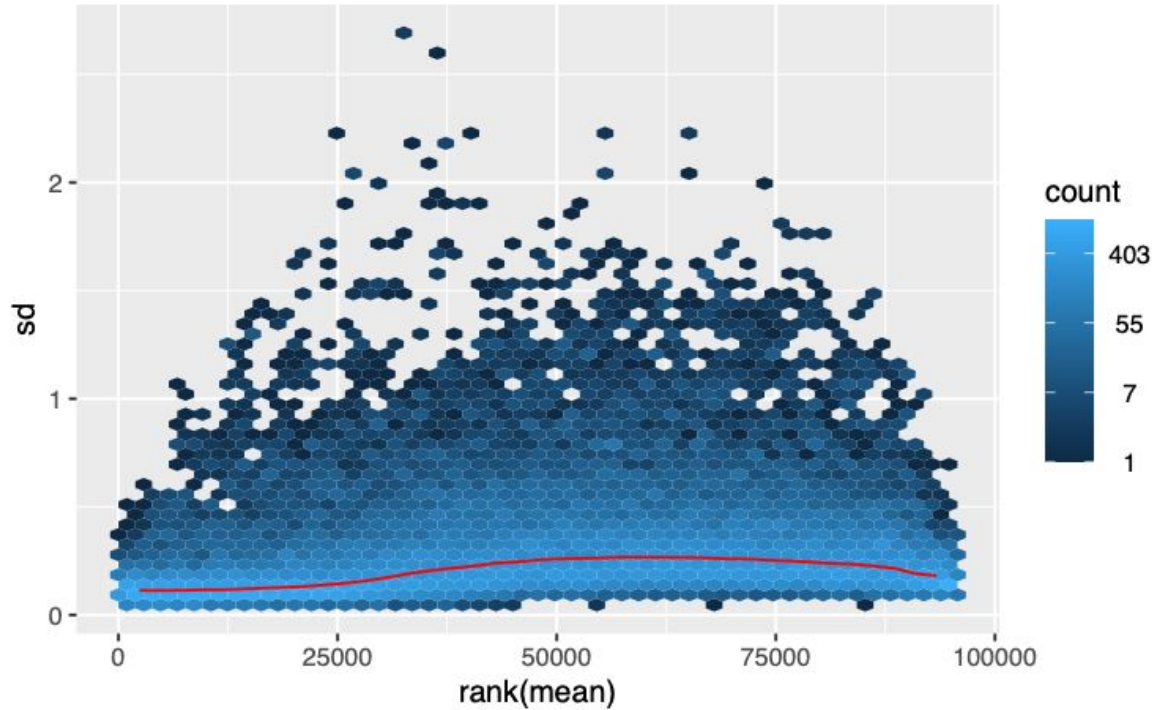


Quality control of chips: Single chip control

GSM1387537_ERK1_2 inh-3ha.cel



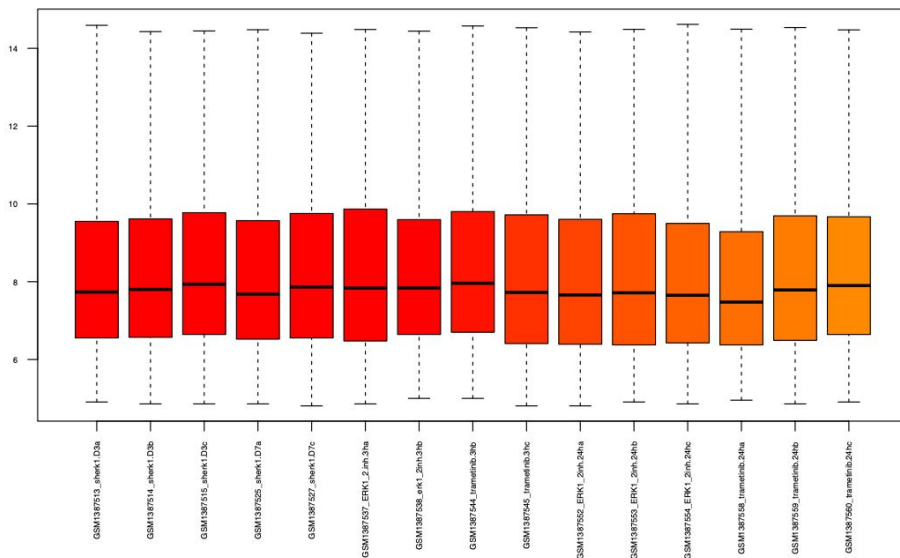
Quality control of chips: meanSd-plot



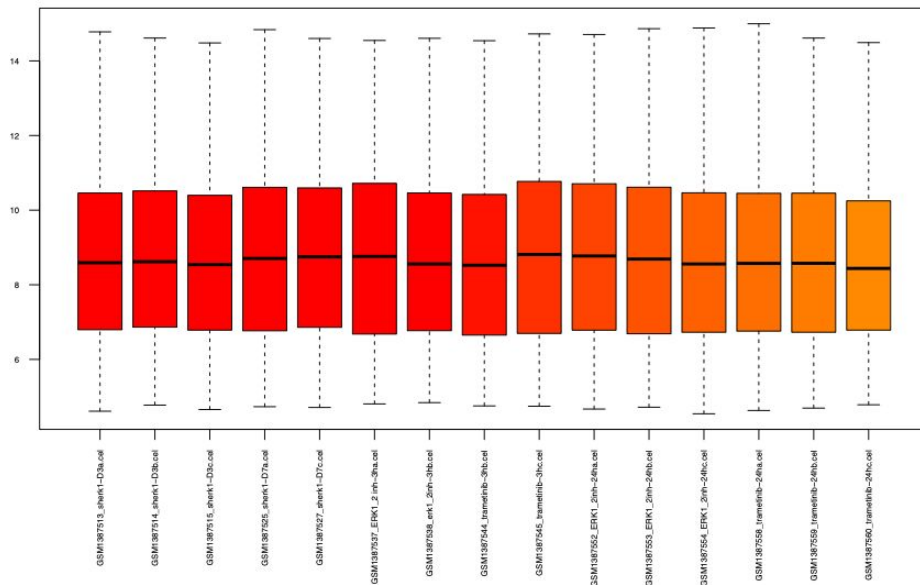


Quality control of chips: Boxplots

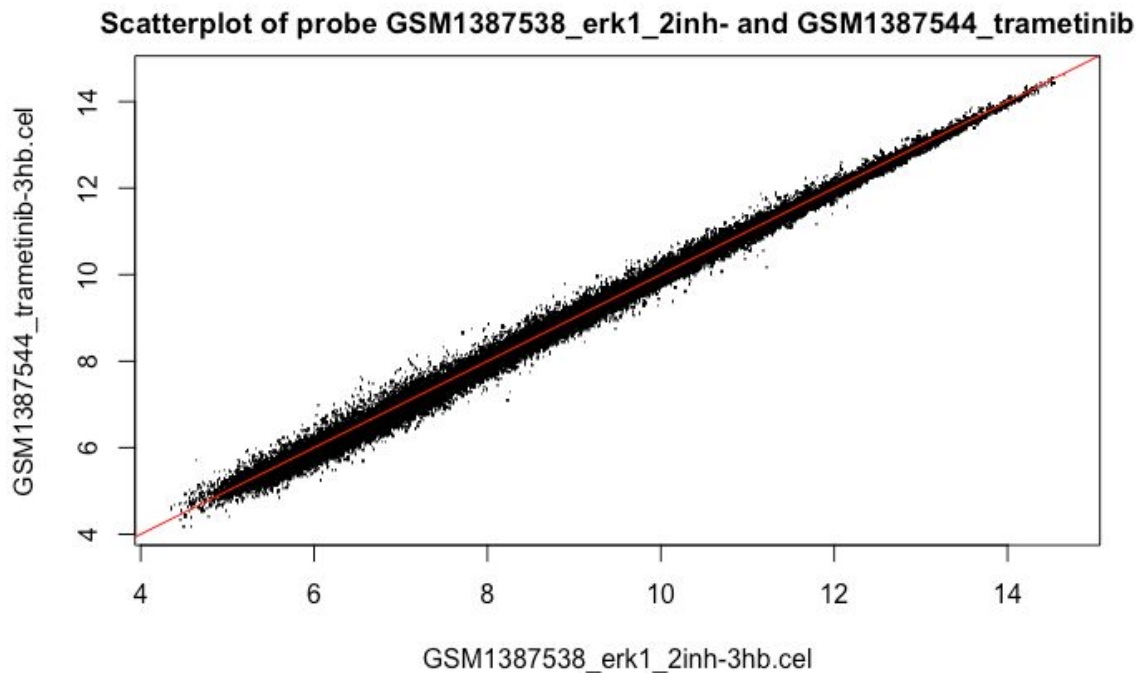
Gene expression in skin cancer before normalization



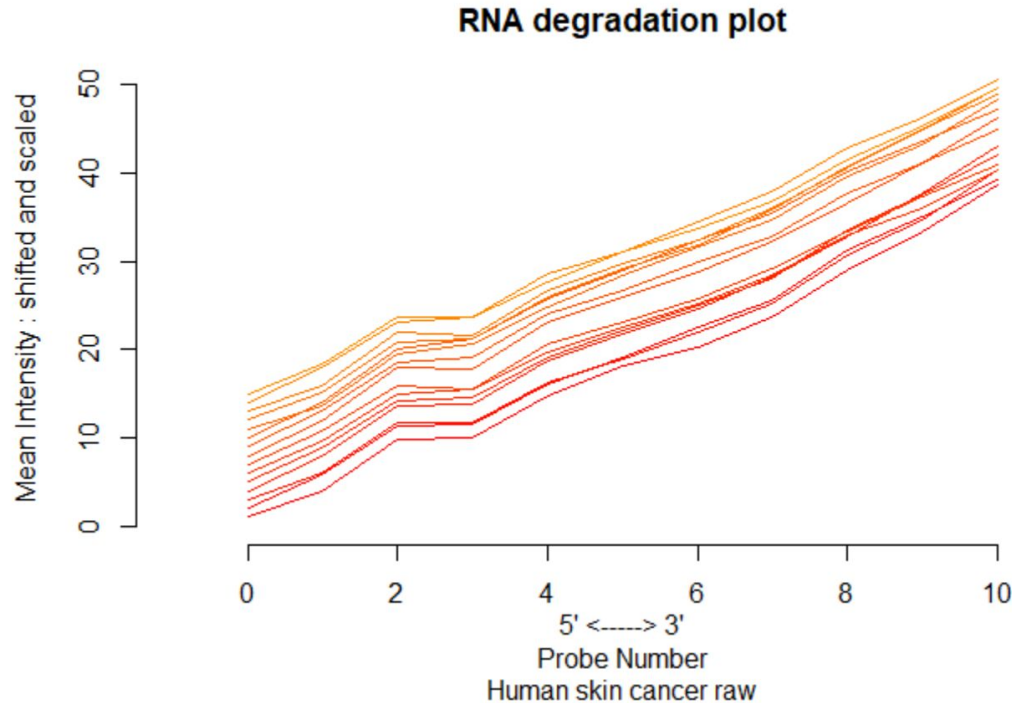
Gene expression in skin cancer after normalization



Quality control of chips: Scatter plot



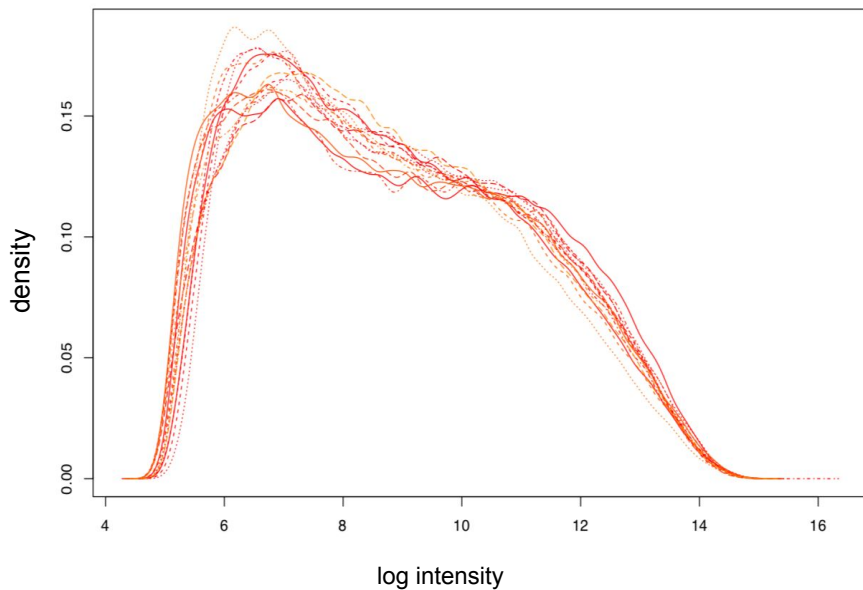
Quality control of chips: RNA degradation



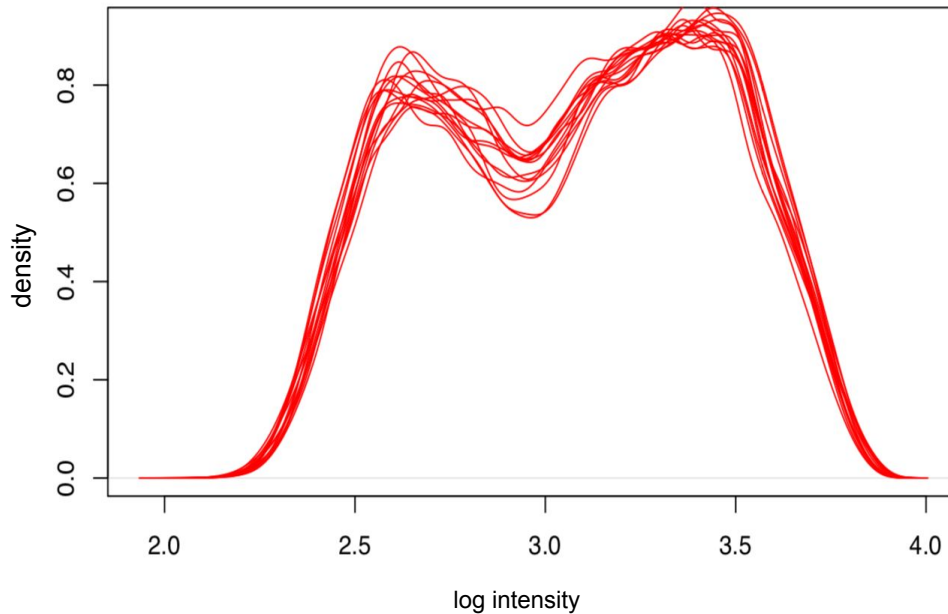


Quality control: Histograms

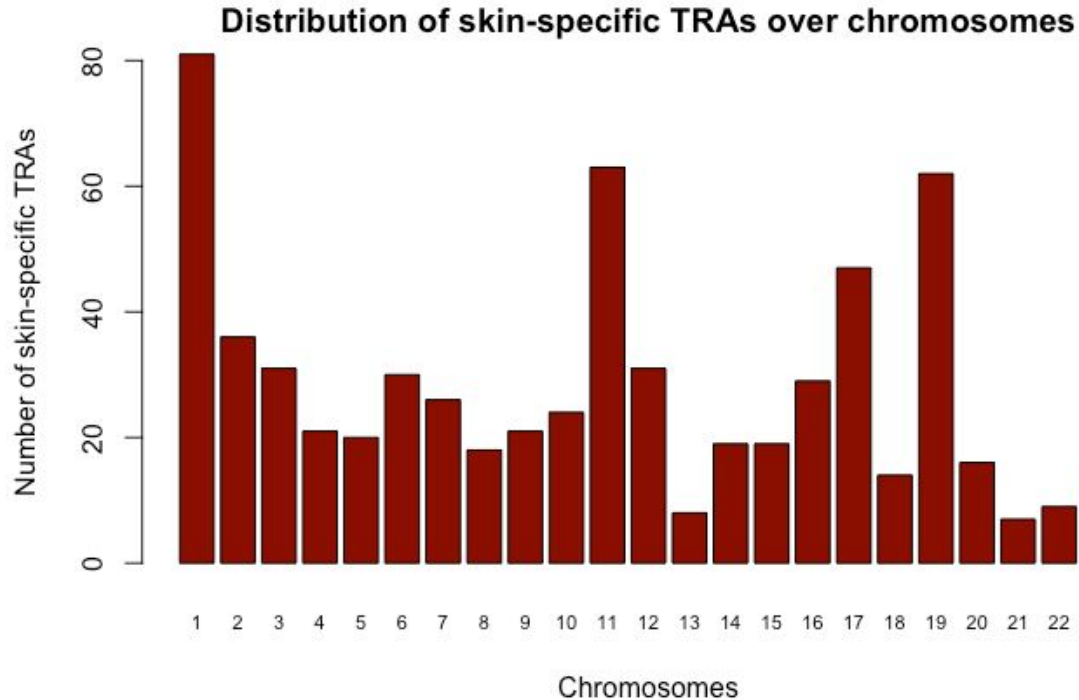
data pre-normalization



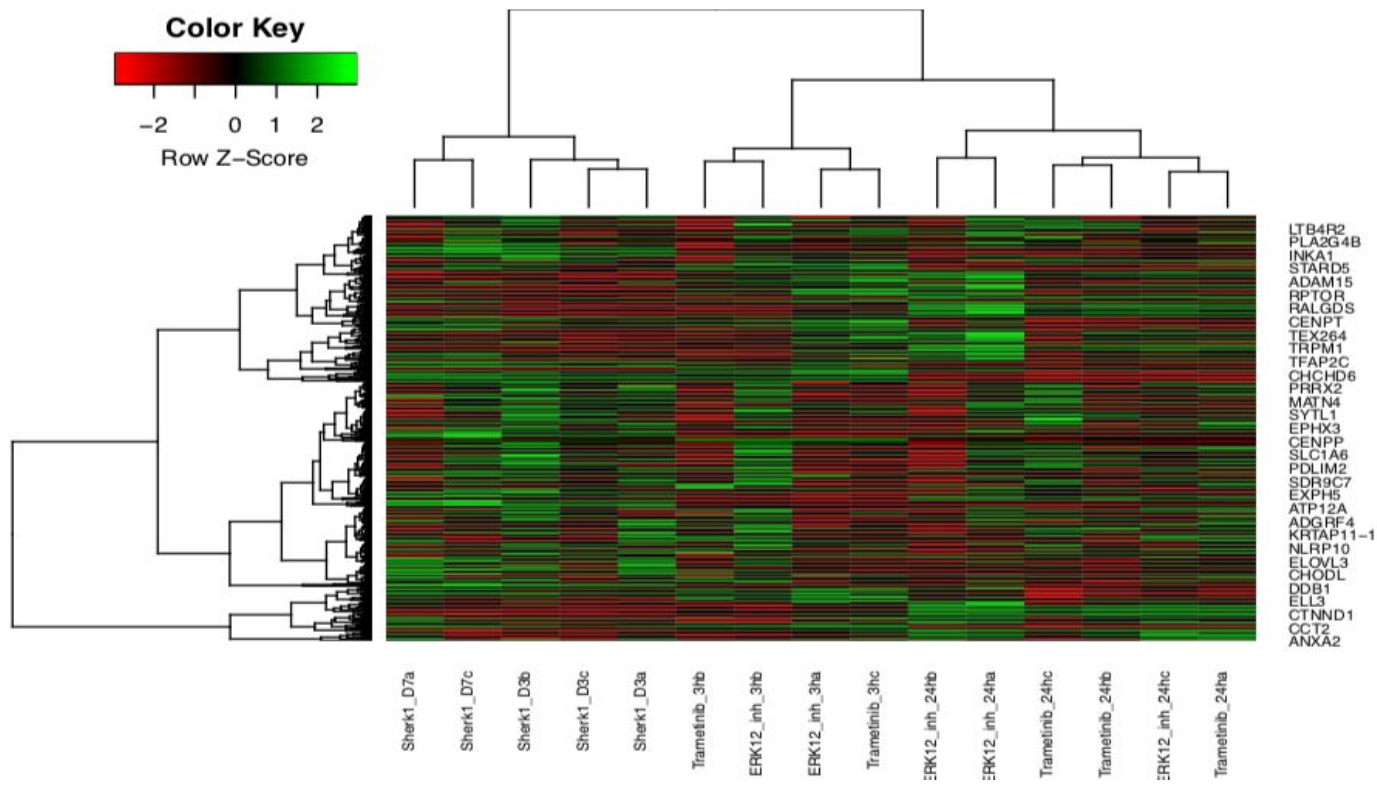
data normalized



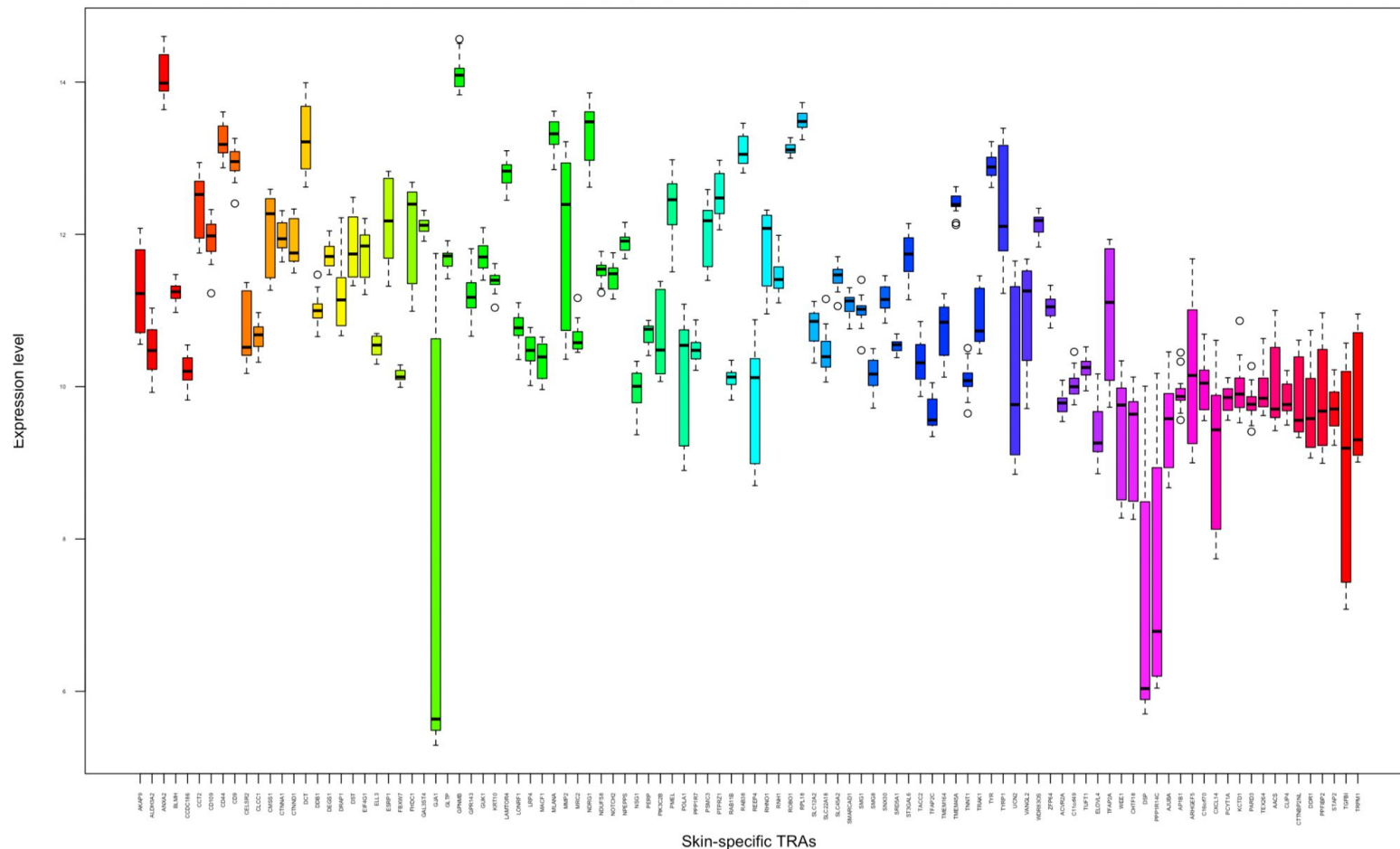
Chromosome localization



Heatmap



Highly expressed skin-specific TRAs



Thank you for listening!

Any questions?

