Thyroid specific antigens

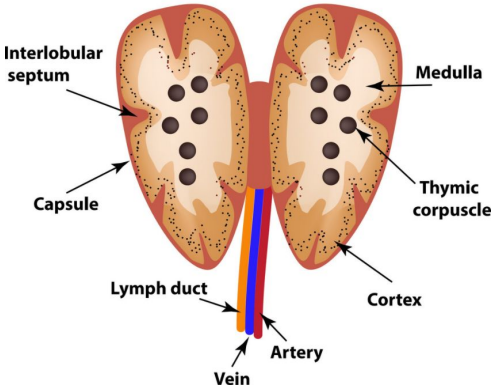
and their gene expression in thyroid cancer

A data science project performed by Camila Vacas, Line Weiß and Carina Keßler
Supervisor: Dr. Maria Dinkelacker, Dr. Carl Herrmann
Tutor: Nils Mechtel

General Informations

TRAs and thyroid cancer

- Tumor Associated Antigens are expressed in thymus
- Radiation & thyroid cancer
  - sensitive to long-term effects
  - 80% papillary carcinoma
  - activation of MAPK signal cascade



structure of the thymus gland

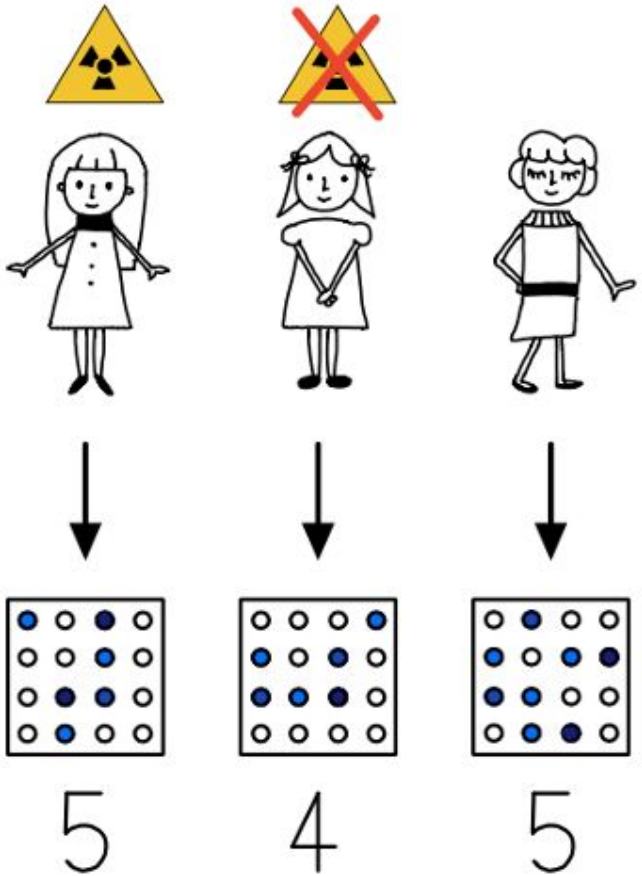
https://www.vectorstock.com/royalty-free-vector/structure-thymus-gland-infographics-vector-12904537

The structure of our dataset

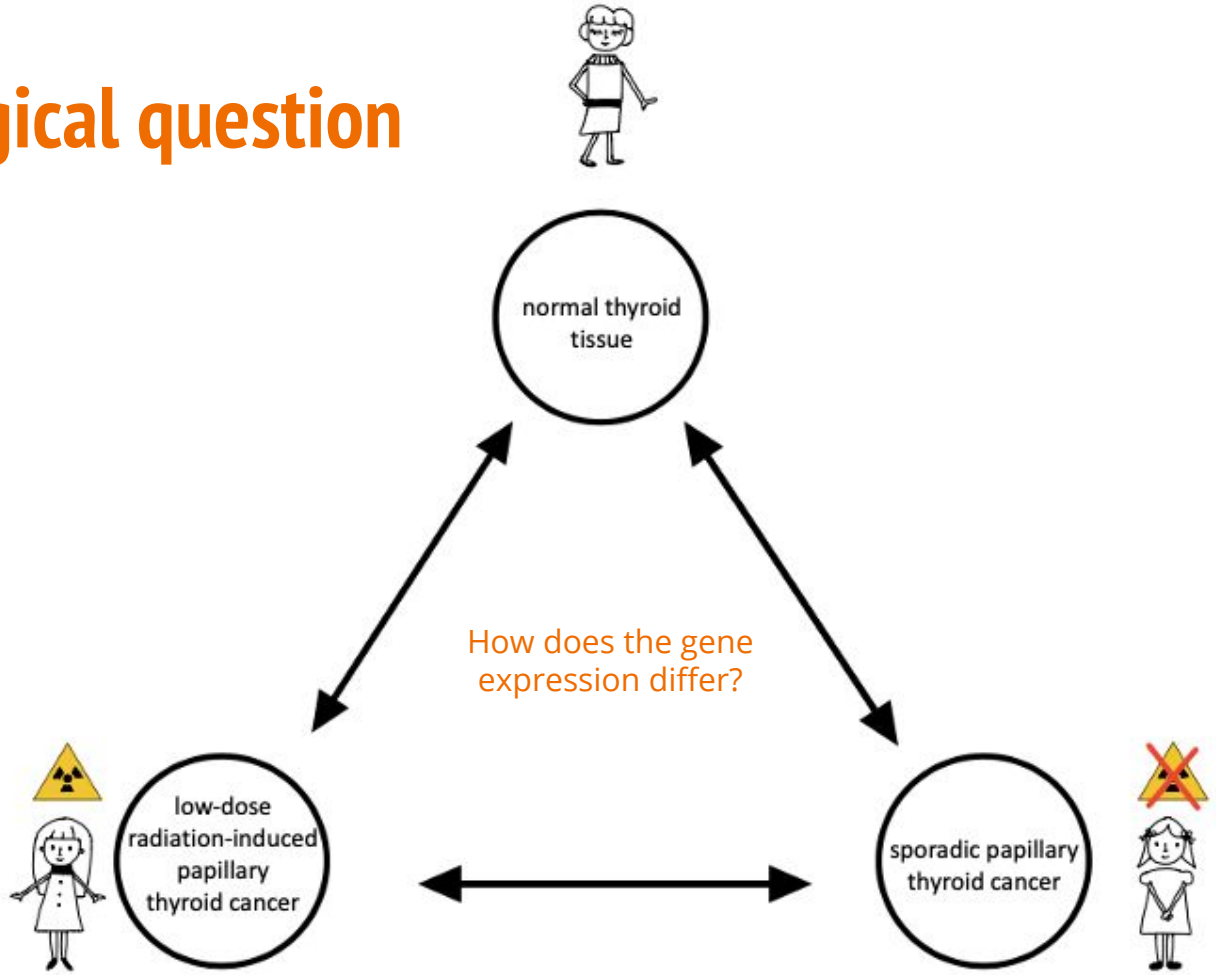
5 arrays of patients with papillary thyroid cancer (PTC), who were exposed to radiation

4 arrays of patients with PTC, who were NOT exposed to radiation

5 arrays of patients with healthy thyroid tissue

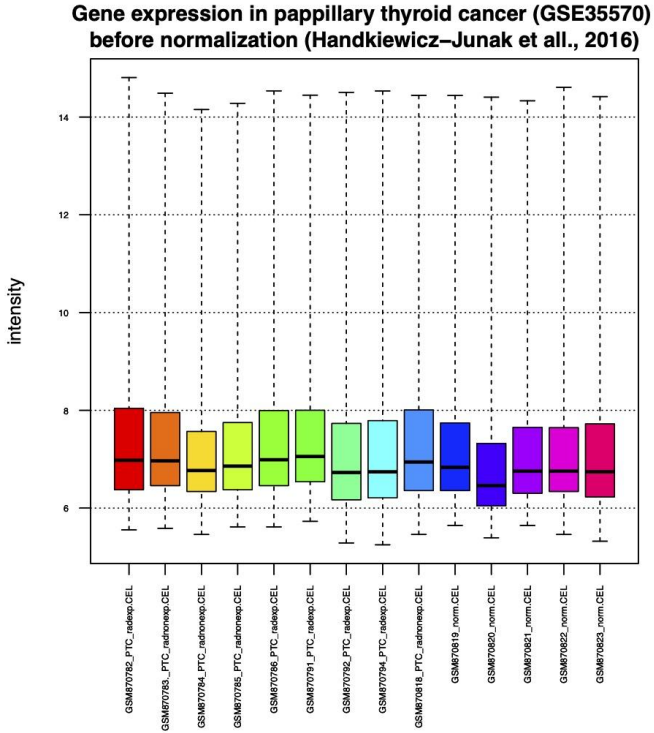


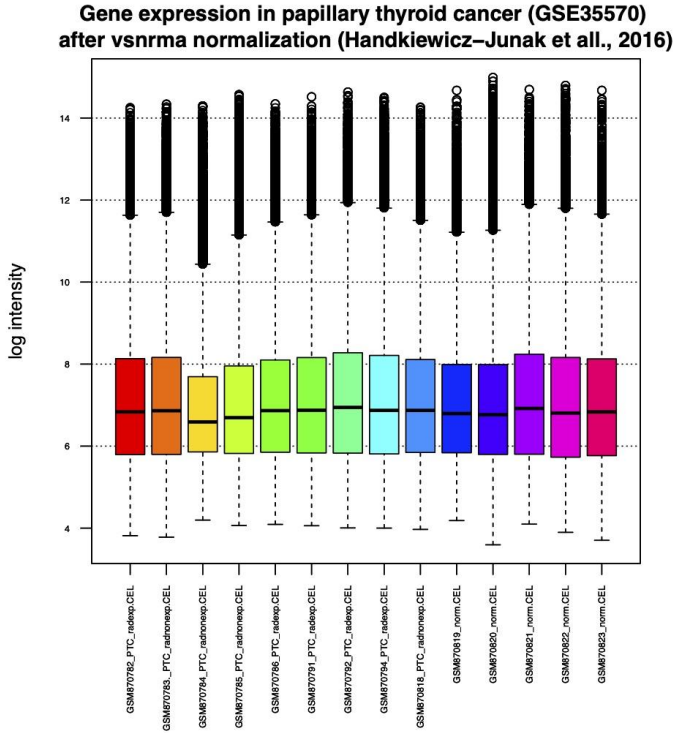
The biological question



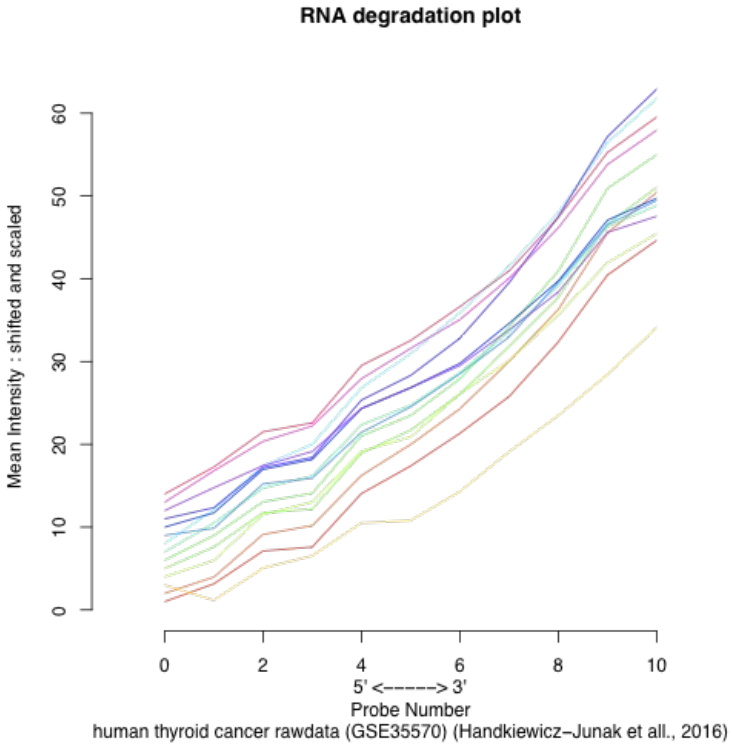
Healthy, radiation exposed, non-exposed tissue

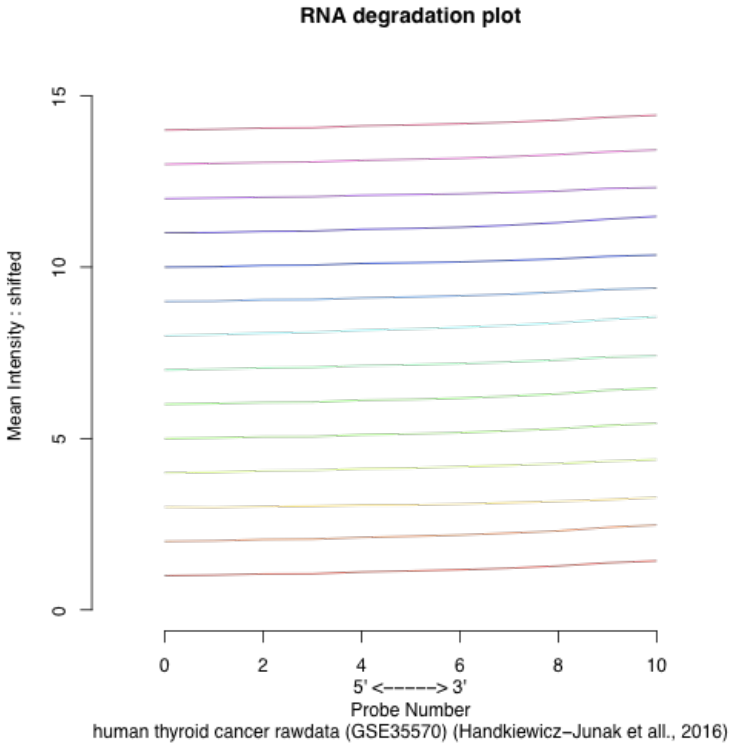
Box-plots of the gene expression in PTC



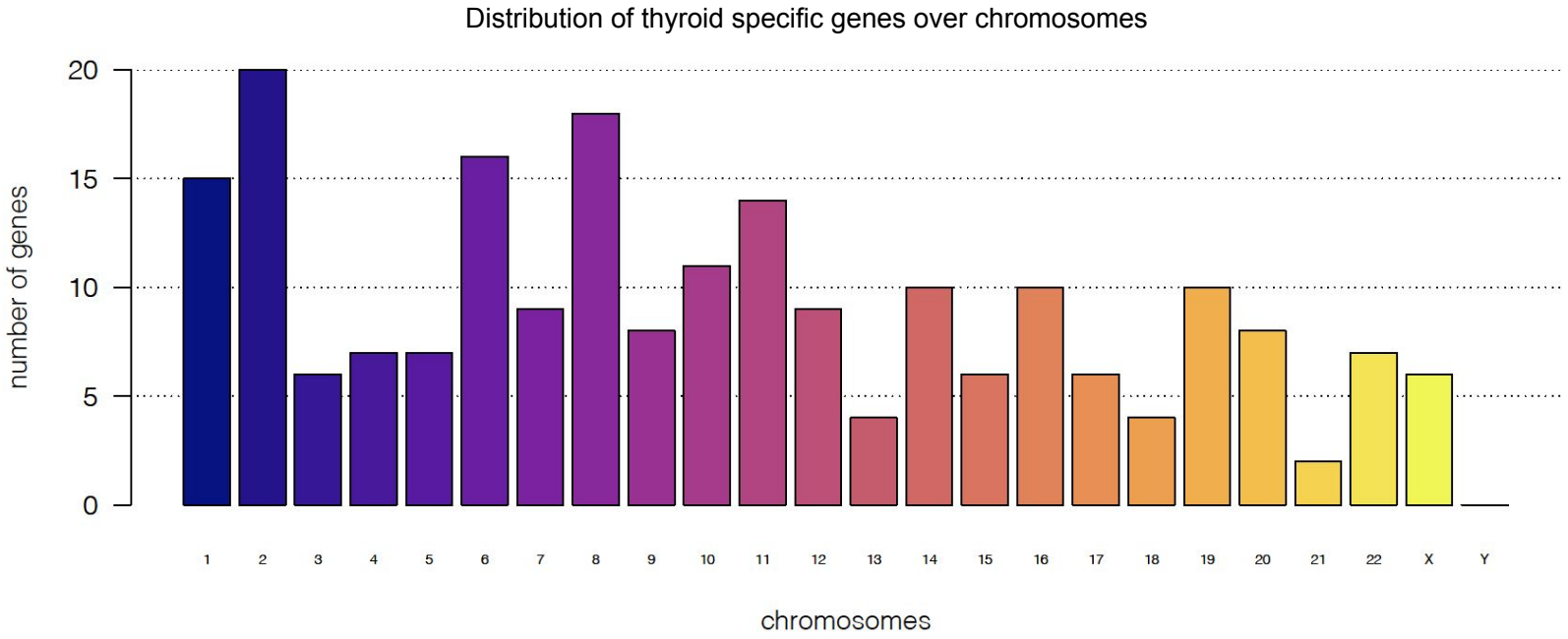


RNA degradation plot





Chromosome distribution

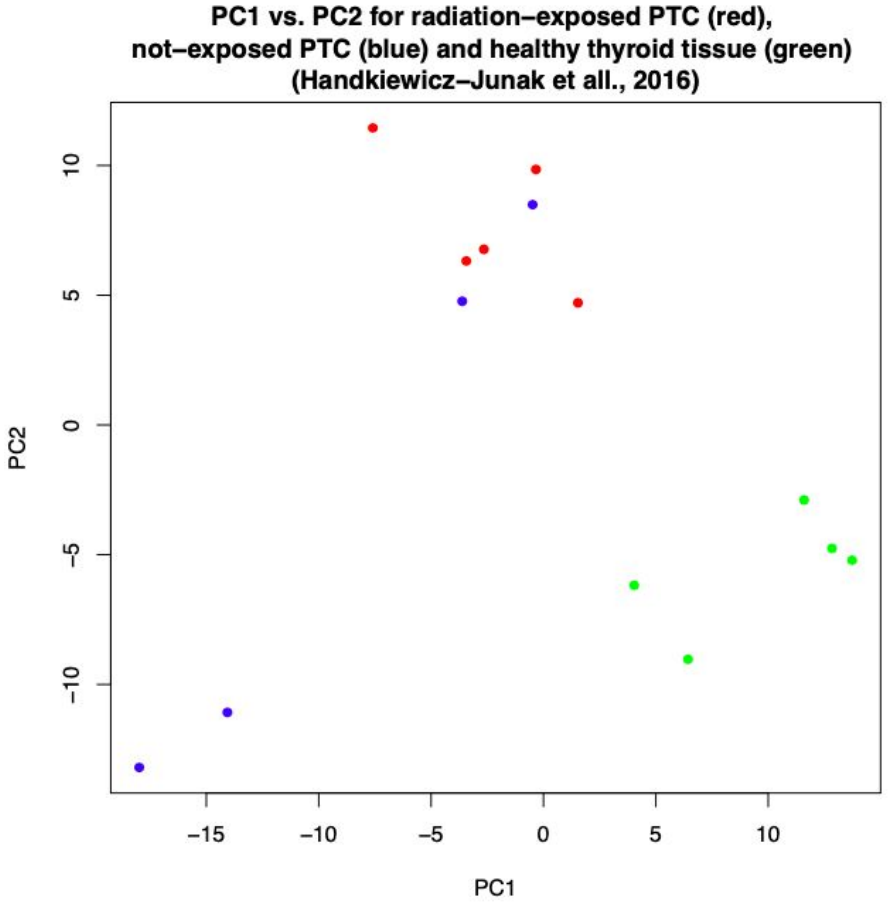


PCA

Further questions:

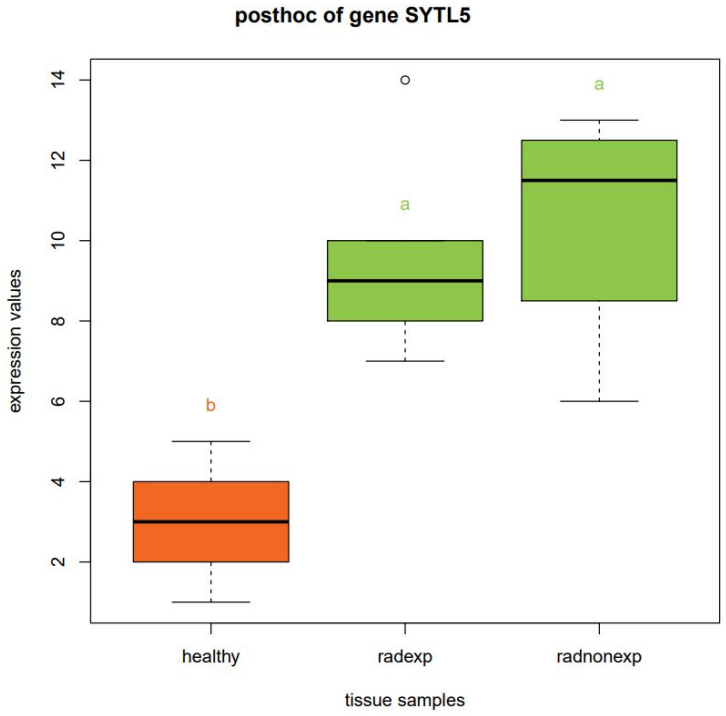
Are always the healthy samples separated from patients who received radiation?

→ Test of hypothesis by adding more chips



ANOVA and post hoc test

=> SYTL5 is upregulated
in thyroid cancer

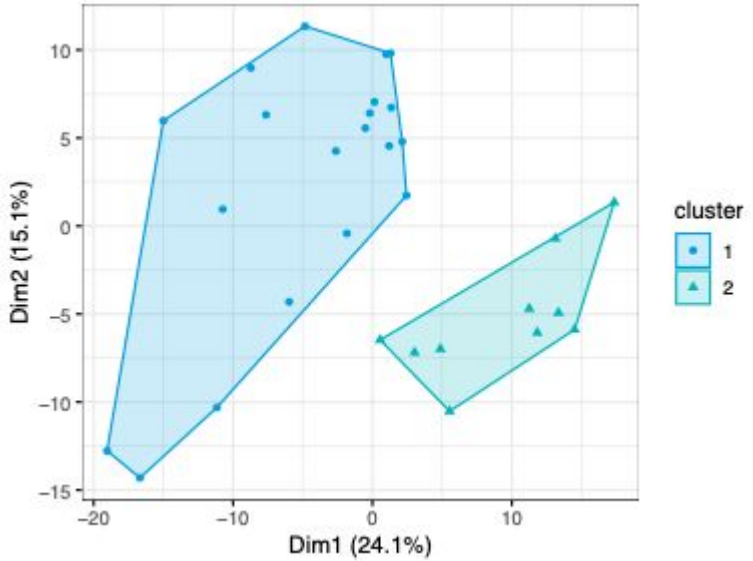


Cancer subtypes

But were those just outliers?

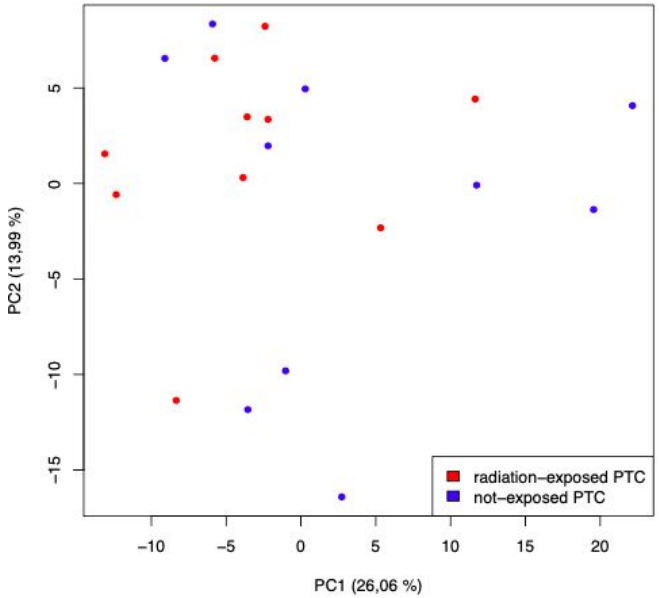
PCA and k-means clustering with 30 chips instead of 14 to get a better idea of the gene expression pattern in the three groups

PC1 vs. PC2 for radiation-exposed PTC, not-exposed PTC and healthy thyroid tissue (Handkiewicz-Junak et all., 2016)

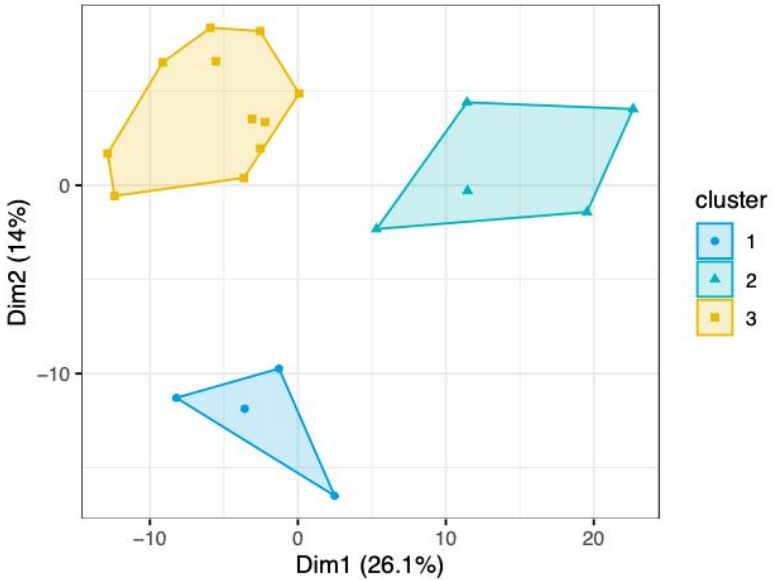


Is there really no difference between the sick samples?

PC1 vs. PC2 for radiation-exposed PTC and not-exposed PTC (Handkiewicz-Junak et all., 2016)



PC1 vs. PC2 for radiation-exposed PTC and not-exposed PTC colored according to k-means clustering (Handkiewicz-Junak et all., 2016)



The regression model

goal: create a regression model to predict the subtypes

columnnames: actual subtypes the chips belong to

rownames: subtypes the chips were assigned to by the regression model

training.model	subtype 1	subtype 2	subtype 3
subtype 1	3	0	0
subtype 2	0	4	0
subtype 3	0	0	7

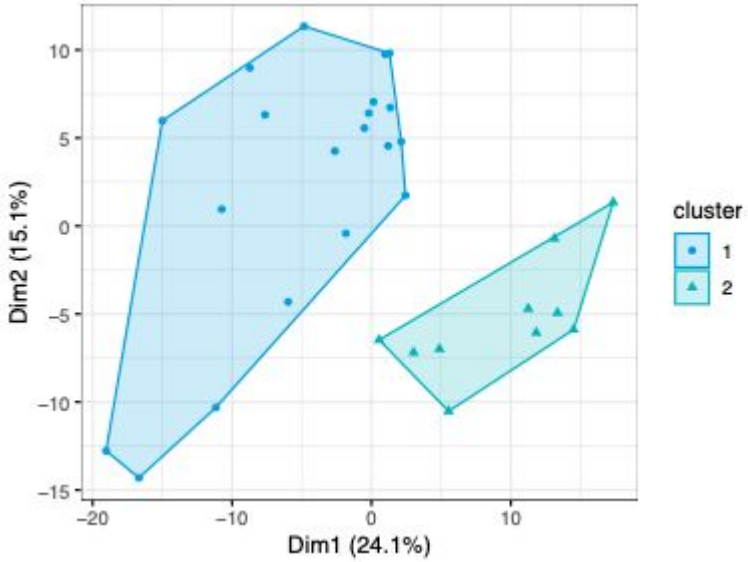
test.model	subtype 1	subtype 2	subtype 3
subtype 1	1	0	0
subtype 2	0	1	0
subtype 3	0	0	4

Results

Cancer/healthy

- 360 thyroid specific antigens
- 153 significantly different expressed genes
- no difference radiation exposed/ not exposed samples

PC1 vs. PC2 for radiation-exposed PTC, not-exposed PTC and healthy thyroid tissue (Handkiewicz-Junak et all., 2016)



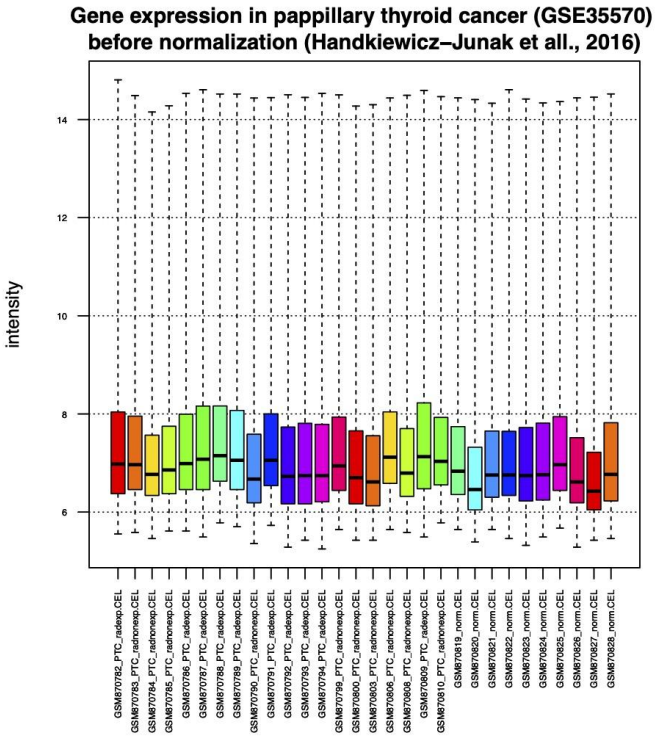
Cancer subtypes

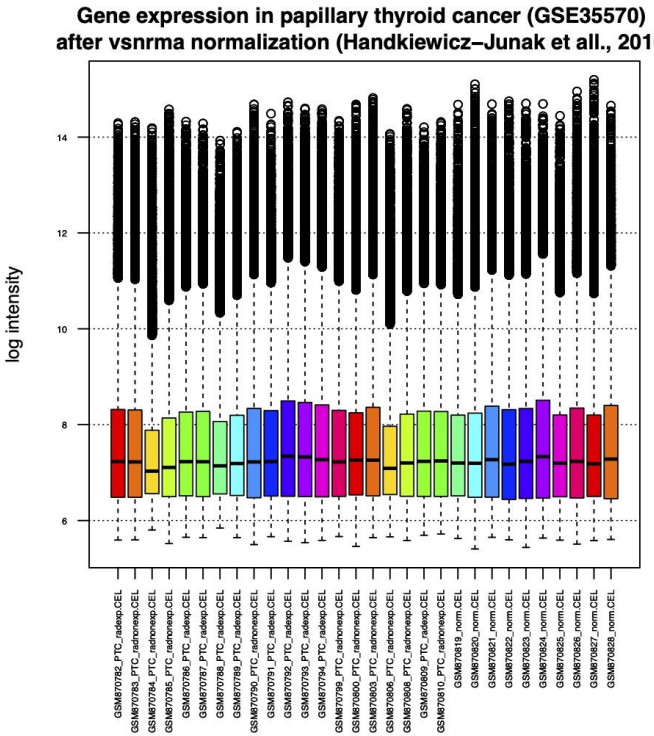
- 3 PTC subtypes:
  - conventional variant
  - follicular variant
  - tall cell variant
- 134 significantly different expressed genes
- 11 genes different in each subtype

Main 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, 2007.

Dinkelacker, 2019, PhD thesis, "Chromosomal clustering of tissue-restricted antigens", University of Heidelberg.


Gene expression in pappillary thyroid cancer (GSE35570)
before normalization (Handkiewicz–Junak et all., 2016)
intensity
14
12
10
8
6
GSM870782_PTC_radexp.CEL
GSM870783_PTC_radnonexp.CEL
GSM870784_PTC_radnonexp.CEL
GSM870785_PTC_radnonexp.CEL
GSM870786_PTC_radexp.CEL
GSM870787_PTC_radexp.CEL
GSM870788_PTC_radexp.CEL
GSM870789_PTC_radexp.CEL
GSM870790_PTC_radnonexp.CEL
GSM870791_PTC_radexp.CEL
GSM870792_PTC_radexp.CEL
GSM870793_PTC_radexp.CEL
GSM870794_PTC_radexp.CEL
GSM870799_PTC_radnonexp.CEL
GSM870800_PTC_radnonexp.CEL
GSM870803_PTC_radnonexp.CEL
GSM870806_PTC_radnonexp.CEL
GSM870808_PTC_radnonexp.CEL
GSM870809_PTC_radexp.CEL
GSM870810_PTC_radnonexp.CEL
GSM870819_norm.CEL
GSM870820_norm.CEL
GSM870821_norm.CEL
GSM870822_norm.CEL
GSM870823_norm.CEL
GSM870824_norm.CEL
GSM870825_norm.CEL
GSM870826_norm.CEL
GSM870827_norm.CEL
GSM870828_norm.CEL


Gene expression in papillary thyroid cancer (GSE35570)
after vsnrma normalization (Handkiewicz–Junak et all., 2016)
log intensity
14
12
10
8
6
GSM870782_PTC_radexp.CEL
GSM870783_PTC_radnonexp.CEL
GSM870784_PTC_radnonexp.CEL
GSM870785_PTC_radnonexp.CEL
GSM870786_PTC_radexp.CEL
GSM870787_PTC_radexp.CEL
GSM870788_PTC_radexp.CEL
GSM870789_PTC_radexp.CEL
GSM870790_PTC_radnonexp.CEL
GSM870791_PTC_radexp.CEL
GSM870792_PTC_radexp.CEL
GSM870793_PTC_radexp.CEL
GSM870794_PTC_radexp.CEL
GSM870799_PTC_radnonexp.CEL
GSM870800_PTC_radnonexp.CEL
GSM870803_PTC_radnonexp.CEL
GSM870806_PTC_radnonexp.CEL
GSM870808_PTC_radnonexp.CEL
GSM870809_PTC_radexp.CEL
GSM870810_PTC_radnonexp.CEL
GSM870819_norm.CEL
GSM870820_norm.CEL
GSM870821_norm.CEL
GSM870822_norm.CEL
GSM870823_norm.CEL
GSM870824_norm.CEL
GSM870825_norm.CEL
GSM870826_norm.CEL
GSM870827_norm.CEL
GSM870828_norm.CEL