Analysis of Cellular Proteomic Processes

Keerat Gill, Michelle Wang, Kefira Wang, Ali Hafeez

December 8, 2022

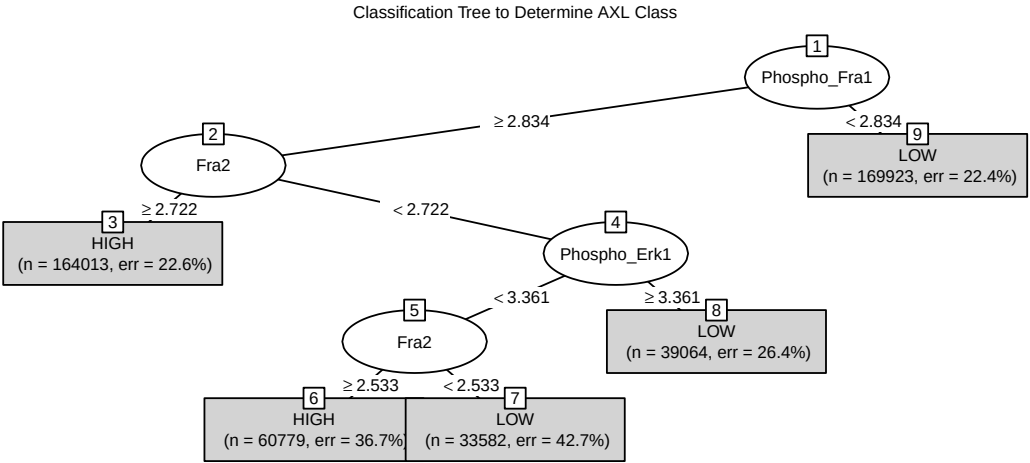
- The data is based on Flow Cytometry for single-cell analysis and Mass Spectrometry for measurement of cellular proteomic processes.
- Measures 22 AP-1 transcription factors and 4 phenotype proteins.
  What is *"good"* cellular homeostasis? **And can bad proteins become good?**

Our main research questions are:
1. Can we predict cellular phenotypical outcomes from transcription factors?
2. Do the phenotype levels change over time in different drug conditions?
3. What is the relationship between different proteins at a fixed time?

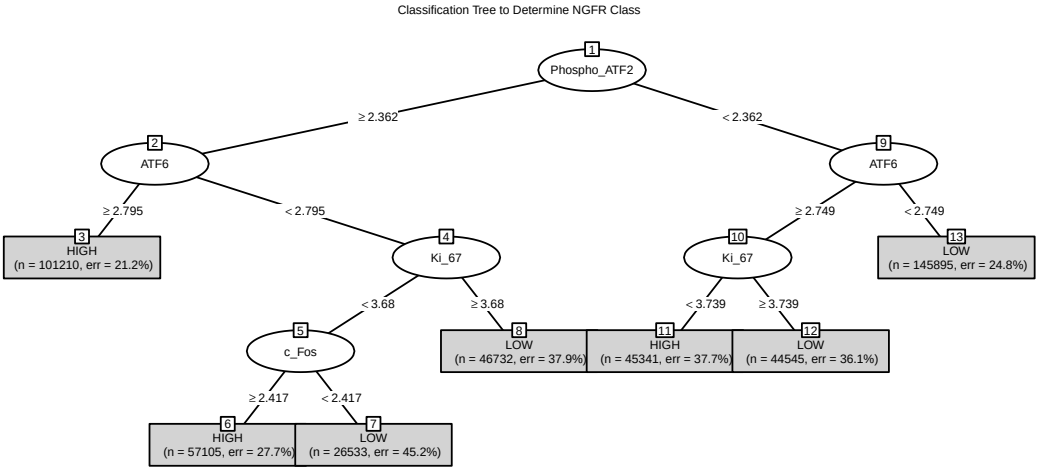
Statistical Methods

- We first omitted all NA values to make sure our data was complete.
- We designated each phenotype a HIGH or LOW classification depending on if the recorded value was higher than the mean.
- Using the HIGH/LOW designation, we designated each observation to be melanocytic or undifferentiated.

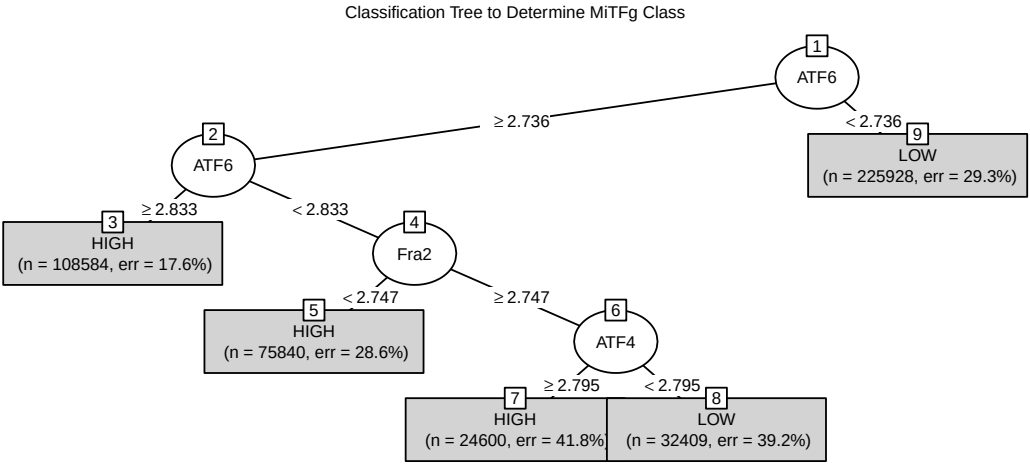
AXL Classification Tree



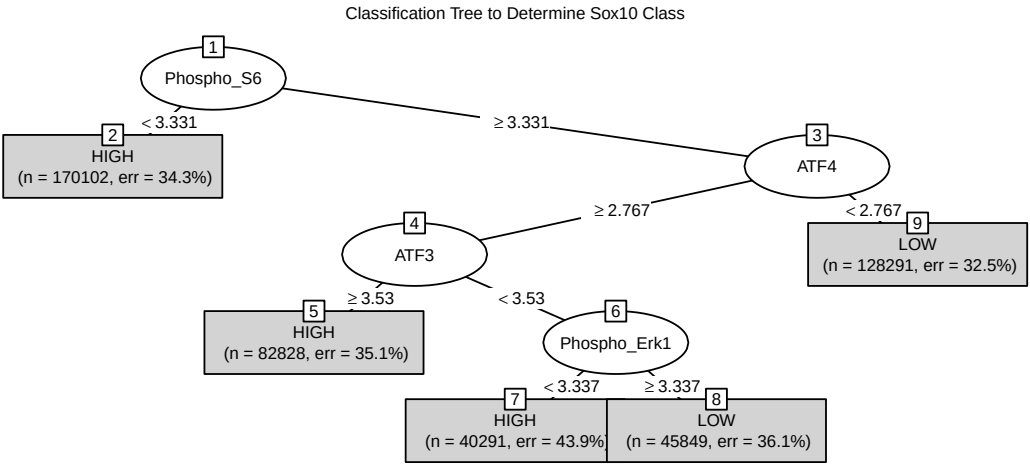
NGFR Classification Tree



MiTFg Classification Tree



Sox10 Classification Tree



Findings from Classification Trees - Cancerous Cells (Melanocytic)

AXL: LOW
- Phospho_Fra1 < 2.834
- Phospho_Fra1 >= 2.834, Fra2 < 2.722 AND Phospho_Erk1 >= 3.361
- Phospho_Fra1 >= 2.834, Fra2 < 2.722, Phospho_Erk1 < 3.361 AND Fra2 < 2.533

NGFR: HIGH
- Phospho_ATF2 < 2.362 AND ATF6 >= 2.749 AND Ki_67 < 3.739
- Phospho_ATF2 >= 2.362 AND ATF6 >= 2.795
- Phospho_ATF2 >= 2.362 AND ATF6 < 2.795 AND Ki_67 < 3.68 AND c_Fos >= 2.417

MiTFg: HIGH
- ATF6 >= 2.736 AND ATF6 >= 2.833
- ATF6 >= 2.736 AND ATF6 < 2.833 AND Fra2 < 2.747
- ATF6 >= 2.736 AND ATF6 < 2.833 AND Fra2 >= 2.747 AND ATF4 >= 2.795

SOX10: LOW
- Phospho_S6 >= 3.331 AND ATF4 < 2.767
- Phospho_S6 >= 3.331, ATF4 >= 2.767, ATF3 < 3.53 AND Phospho_Erk1 >= 3.337

Findings from Classification Trees - Non-Cancerous Cells (Undifferentiated)

AXL: HIGH
- Phospho_Fra1 >= 2.834 AND Fra2 >= 2.722
- Phospho_Fra1 >= 2.834, Fra2 < 2.722, Phospho_Erk1 < 3.361 AND Fra2 >= 2.533

NGFR: LOW
- Phospho_ATF2 < 2.362 AND ATF6 < 2.749
- Phospho_ATF2 < 2.362 AND ATF6 >= 2.749 AND Ki_67 >= 3.739
- Phospho_ATF2 >= 2.362 AND ATF6 < 2.795 AND Ki_67 >= 3.68
- Phospho_ATF2 >= 2.362 AND ATF6 < 2.795 AND Ki_67 < 3.68 AND c_Fos < 2.417

MiTFg: LOW
- ATF6 < 2.736
- ATF6 >= 2.736 AND ATF6 < 2.833 AND Fra2 >= 2.747 AND ATF4 < 2.795

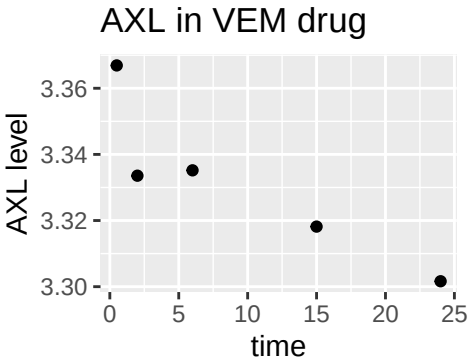
SOX10: LOW
- Phospho_S6 >= 3.331 AND ATF4 < 2.767
- Phospho_S6 >= 3.331, ATF4 >= 2.767, ATF3 < 3.53, Phospho_Erk1 >= 3.337

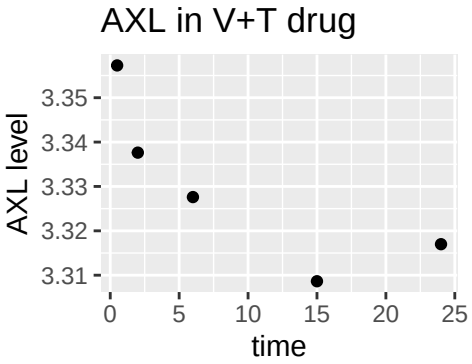
Do the phenotype levels change over time in different drug conditions?

- We first plotted the average level of phenotypes in melanocytic cells over time, grouped by drug type, to see if there was a noticable difference to justify a two-sample hypothesis test.
- NGFR, AXL and MiTFg had noticable linear trends in their means overtime.
- Sox10 had an abnormal jump at the 2h timepoint for both drugs.
- This was taken into consideration when conducting the two-sample test.
- The following are the average means and p-values for the change in phenotype level between the 0.5h and 24h timepoint.
- Alpha-significance level is 0.05, as per scientific convention.

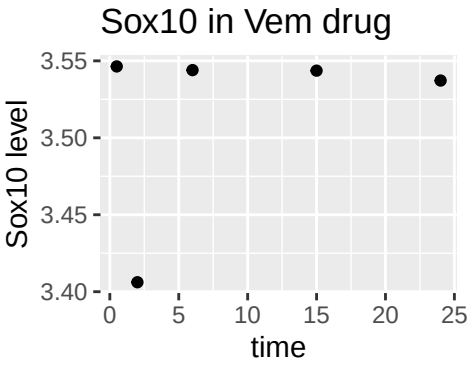
NULL HYPOTHESIS: There is no change in protein level over time.

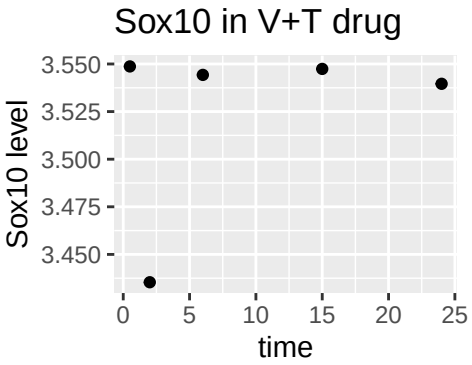
AXL Levels Change Over Time in VEM and VEM+TRAM Drug



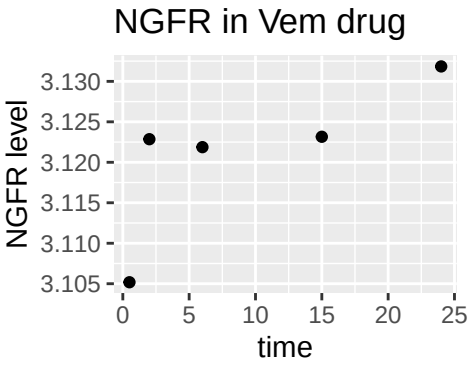


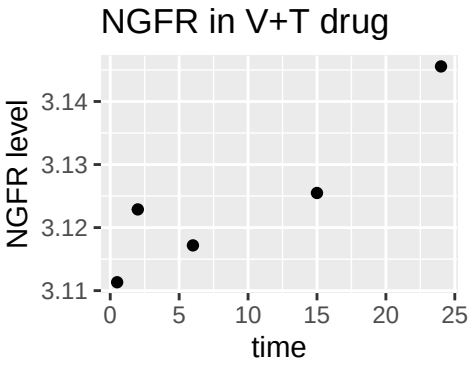
Sox10 Level Change Over Time in VEM and VEM+TRAM Drug



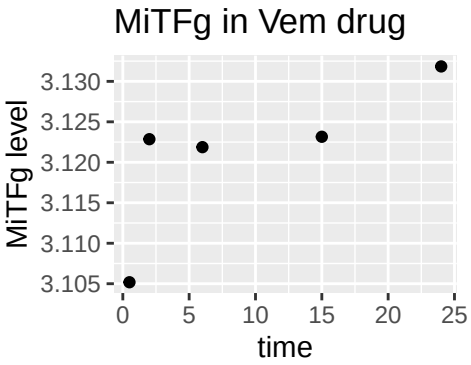


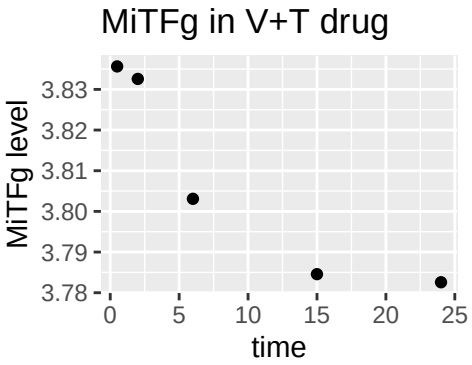
NGFR Level Change Over Time in VEM and VEM+TRAM Drug





MiTFg Level Change Over Time in VEM and VEM+TRAM Drug





Resulting P-Values for Cancerous Cells

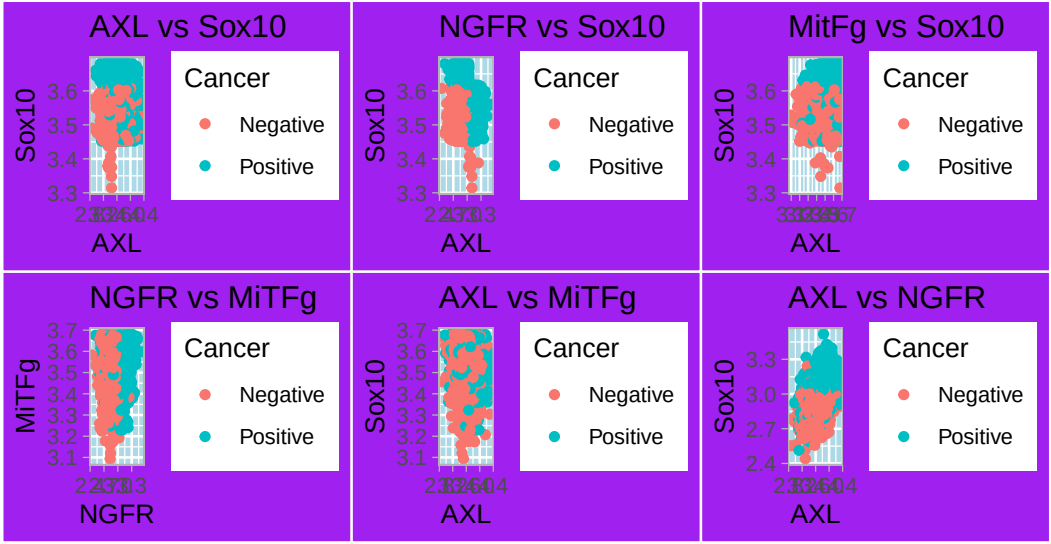
From the two-sample hypothesis tests, our main findigns were;

- AXL_VEM drug → P-value was so small, it showed up as 0 on the code.
- AXL_VEM+TRAM drug → P-value was so small, it showed up as 0 on the code.
- NGFR_VEM drug → P-value was so small, it showed up as 0 on the code.
- NGFR_VEM+TRAM drug → P-value was so small, it showed up as 0 on the code.
- MiTFg_VEM drug → P-value was so small, it showed up as 0 on the code.
- MiTFg_VEM+TRAM drug → P-value was so small, it showed up as 0 on the code.
- Sox10 VEM drug → P-value was 0.74. Fail to reject the null hypothesis.
- Sox10 VEM+TRAM drug → P-value was 0.36. Fail to reject the null hypothesis.

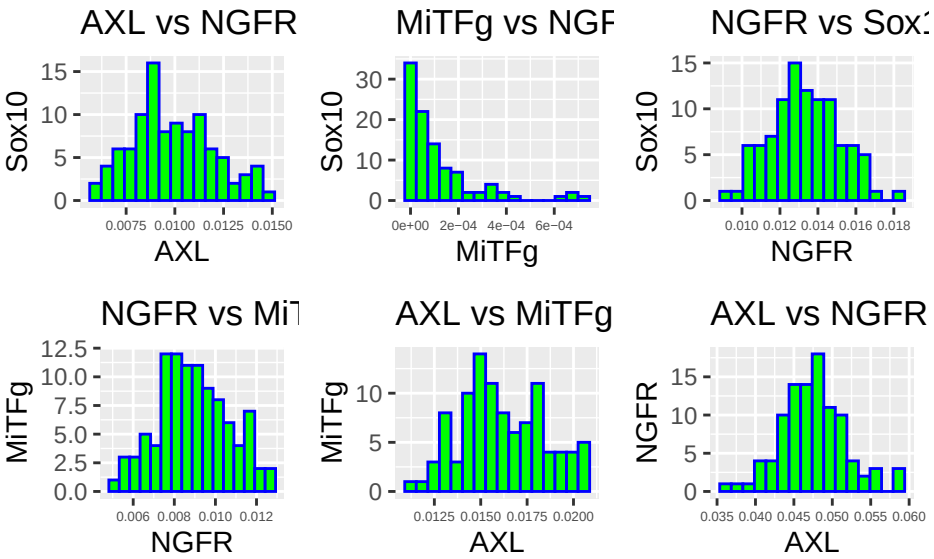
What is the relationship between different proteins at a fixed timepoint and environmental condition?

- We wanted to investigate this as it is an essential component to understanding how phenotypes behave and what environmental conditions, such as other phenotypes, can affect their own phenotype level.
- If we find trends, we can use this to justify further scientific exploration into how phenotypes themselves can be used to reduce melanocytic cells.

```
## [1] 0.0002527696
## [1] 0.01324962
## [1] 0.1483614
## [1] 0.003353127
## [1] 0.001668293
## [1] 0.03019172
```


AXL vs Sox10
Sox10
AXL
Cancer
Negative
Positive
NGFR vs Sox10
Sox10
AXL
Cancer
Negative
Positive
MitFg vs Sox10
Sox10
AXL
Cancer
Negative
Positive
NGFR vs MiTFg
MiTFg
NGFR
Cancer
Negative
Positive
AXL vs MiTFg
Sox10
AXL
Cancer
Negative
Positive
AXL vs NGFR
Sox10
AXL
Cancer
Negative
Positive

BOOTSTRAPPING CORRELATION ESTIMATES FOR CONFIDENCE INTERVALS



Our main results were...

From the classification trees, it can be concluded that;

- If we know the level of the TFs Phospho_Fra1, Fra2, Phospho_Erk1, Phospho_ATF2, ATF6, Ki_67, c_Fos, ATF4, Phospho_S6, and ATF3, then we can tell if each protein level is HIGH or LOW. From that, we combine our results and figure out if the cell is cancerous or not.

From the two-sample hypothesis tests, it can be concluded that:

- There is a significant change overtime for the AXL, NGFR and MiTFg proteins in the VEM and VEM+TRAM drug.
- In the VEM drug, AXL decreases overtime, NGFR increases overtime and MiTFg increases overtime.
- In the VEM+TRAM drug, AXL decreases overtime, NGFR increases overtime and MiTFg decreases overtime.
- We fail to reject the null hypothesis for Sox10.

From the correlation estimation and bootstrapping, it can be concluded that;
- There exists a weak correlation between different proteins which may force us to reassess the importance of the relative quantities of proteins.

Conclusion

We have concluded that...

- In the presence of the VEM+TRAM drug, MiTFg levels reduce. This is a significant finding since we know cells are cancerous when they have high MiTFg, high NGFR, low Sox10 and low AXL.
- Although Sox10 is relatively unchanged overtime, AXL further decreases and NGFR further increases, the significant change of one protein away from the regular cancerous protein configuration is promising and provides a scientific basis to further research how proteins respond to the Vem and Tram drug.
- Furthermore, given the weak correlation between key phenotypes, the findings from this investigation provide a basis for further scientific investigation into how phenotypes and various drugs may reduce the development of melanocytic cells.

Some considerations include;

- Each experiment was made under the condition that everything else remained constant except the variables under consideration.
- Sox10 has a noticable change in protein levels at the 2h timepoint which, due to the limitations of linear regression, could not be aptly included in this investigation.

Acknowledgements

All code and functions used during this investigation come from the following R libraries; tidyverse, rpart, partykit and cowplot.