

DMBT1 Longitudinal Differential Abundance - Two-Way Results

Simon Fontaine

18/04/2023

Contents

Analysis description	1
Goal	1
Methodology	1
Genotype and histopathology	2
Hyperplasia vs CIS/SCC	2
Hyperplasia/CIS vs SCC	2
Hyperplasia vs CIS vs SCC	3
Summary	3
Genotype and sex	3
Notes	3
Representative OTUs	3

Analysis description

Goal

Investigate the temporal relationship between

- **genotype** (KO vs WT),
- **histopathology** (hyperplasia, CIS and SCC) and
- **sex** (M vs F)

in terms of predictive power of the **microbial composition** (clr-transformed OTUs or aggregated at higher taxonomic levels).

Methodology

Since we are limited by the sample size and wish to preserve some power, we limit our analysis to **second-order interactions**. This also simplifies interpretation.

For histopathology, we have three levels: we consider three different encodings:

- hyperplasia vs CIS/SCC: $\beta_{hp}(t) + \beta_{cis/scc}(t)\mathbb{I}[cis/scc]$
- hyperplasia/CIS vs SCC: $\beta_{hp/cis}(t) + \beta_{scc}(t)\mathbb{I}[scc]$
- hyperplasia vs CIS vs SCC, with a monotone contrast encoding: $\beta_{hp}(t) + \beta_{cis/scc}(t)\mathbb{I}[cis/scc] + \beta_{scc}(t)\mathbb{I}[scc]$, so the effect for subject with CIS is $\beta_{hp}(t) + \beta_{cis/scc}(t)$ and the effect for subjects with SCC is $\beta_{hp}(t) + \beta_{cis/scc}(t) + \beta_{scc}(t)$

For the other two groupings, we have only two levels; the encoding is as follows:

- Sex: $M = 0$, $F = 1$
- Genotype: $WT = 0$, $KO = 1$

For microbial compositions, I used the CLR-transformed counts (after zero-replacement). I applied our methodology to

- Selected OTUs (11)
- Genuses previously identified (10)

Genotype and histopathology

Hyperplasia vs CIS/SCC

The model is given by

$$y(t) \sim \beta_{hp}(t) + \beta_{cis/scc}(t)\mathbb{I}[cis/scc] + \beta_{ko}(t)\mathbb{I}[ko] + \beta_{cis/scc,ko}\mathbb{I}[cis/scc, ko]$$

In particular, the fitted mean function per groups are:

- HP, WT: $\beta_{hp}(t)$
- CIS/SCC, WT: $\beta_{hp}(t) + \beta_{cis/scc}(t)$
- HP, KO: $\beta_{hp}(t) + \beta_{ko}(t)$
- CIS/SCC, KO: $\beta_{hp}(t) + \beta_{cis/scc}(t) + \beta_{ko}(t) + \beta_{cis/scc,ko}$

Similarly,

- finding $\beta_{cis/scc,ko} \equiv 0$ implies there are no interactions; then,
- finding $\beta_{ko}(t) \equiv 0$ implies there are no differences between genotype;
- finding $\beta_{cis/scc}(t) \equiv 0$ implies there are no differences between diagnosis.

Hence, we are interested in finding if and when any of the three curves are non-zero. We are not interested in checking whether the intercept $\beta_{hp}(t) \equiv 0$.

Summary of results:

- *Otu0007* show an increase $WT \rightarrow KO$ at week 22
- Genus *Bradyrhizobium* show an increase $WT \rightarrow KO$ at weeks 16 & 22
- Genus *Caulobacter* show an increase $WT \rightarrow KO$ at week 16
- No interaction detected

Hyperplasia/CIS vs SCC

The model is given by

$$y(t) \sim \beta_{hp/cis}(t) + \beta_{scc}(t)\mathbb{I}[scc] + \beta_{ko}(t)\mathbb{I}[ko] + \beta_{scc,ko}\mathbb{I}[scc, ko]$$

Summary of results:

- *Otu0007* show an increase $WT \rightarrow KO$ at week 16
- *Otu0013* show an increase $WT \rightarrow KO$ at week 22
- Genus *Rothia* show an increase $WT \rightarrow KO$ at week 8 and a decrease at week 22
- Genus *Cupriavidus* show an increase $WT \rightarrow KO$ at week 16
- Genus *Bradyrhizobium* show an increase $WT \rightarrow KO$ at weeks 12, 16 & 22
- Genus *Eisenbergiella* show a decrease $WT \rightarrow KO$ at week 12
- Genus *Caulobacter* show a decrease $WT \rightarrow KO$ at weeks 0 & 4 and an increase $HP/CIS \rightarrow SCC$ at week 0

Hyperplasia vs CIS vs SCC

The model is given by

$$y(t) \sim \beta_{hp}(t) + \beta_{cis/scc}(t)\mathbb{I}[cis/scc] + \beta_{scc}(t)\mathbb{I}[scc] + \beta_{ko}(t)\mathbb{I}[ko] + \beta_{cis/scc,ko}\mathbb{I}[cis/scc, ko] + \beta_{scc,ko}\mathbb{I}[scc, ko]$$

Summary of results:

- Genus *Bradyrhizobium* show an increase $WT \rightarrow KO$ at week 16
- Genus *Atopostipes* show a decrease $CIS \rightarrow SCC$ at week 4

Summary

- It appears the best split to find differences is Hyperplasia/CIS vs SCC
- We mostly find differences between genotypes, and a few between diagnosis
- We find no interactions

Genotype and sex

The model is given by

$$y(t) \sim \beta_M(t) + \beta_F(t)\mathbb{I}[F] + \beta_{ko}(t)\mathbb{I}[ko] + \beta_{F,ko}\mathbb{I}[F, ko]$$

Summary of results:

- *Otu0006* show a decrease $WT \rightarrow KO$ at week 22
- *Otu0007* show a complex interaction throughout
- *Otu0013* show an interaction
- Genus *Rothia* show a decrease at weeks 16 & 22
- Genus *Cupriavidus* show a complex interaction throughout (very similar to *Otu0007*, which is of that Genus; notice the change in intercept)
- Genus *Bradyrhizobium* show an interaction
- Genus *Atopostipes* show an interaction
- Genus *Caulobacter* show an interaction
- Generally no $F \rightarrow M$ main effects, but some interactions

Notes

Representative OTUs

- Genus *Atipostipes*: **Otu0052**
- Genus *Bradyrhizobium*: **Otu0018**
- Genus *Caulobacter*: **Otu0016**
- Genus *Cupriavidus*: **Otu0007**
- Genus *Eisenbergiella*: **Otu0212**
- Genus *Rothia*: **Otu0251, Otu0351**