

Longitudinal Analysis of Child Microbiomes and Brain Development

Kevin Bonham, PhD

Wellesley College

2021-12-17

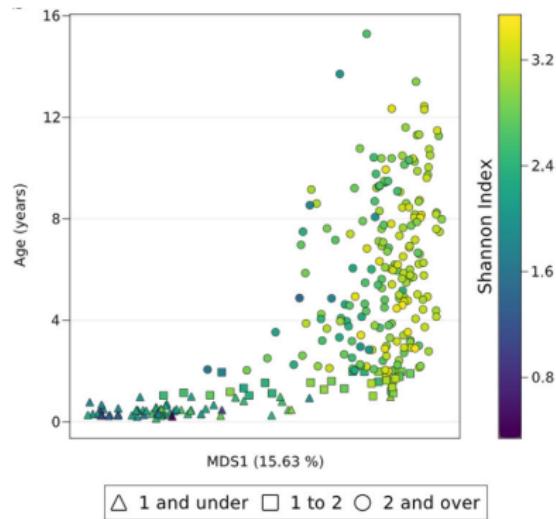
1 Background

2 The RESONANCE cohort of child brain development

3 The GP tool - GaPLAC

The microbiome develops rapidly in the first year of life

- No (or little) prenatal microbiome
- Perinatal microbiome strongly influenced by delivery, antibiotics, diet (breast milk vs formula)
- Gut microbiome is seeded by delivery (birth canal), skin of caregivers, and milk source
- Slowly increasing diversity, conversion to “adult-like” microbiome after transition to solid food

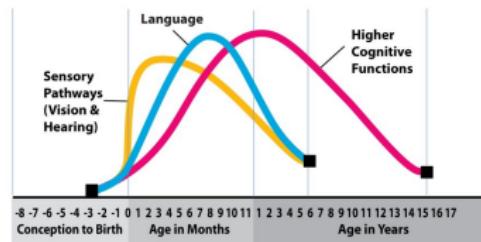


The brain develops rapidly in the first years of life

- By age 5, brain has reached 85% of adult size
- Gross patterns of axonal connections are established, and near-adult levels of myelination also achieved by 5 years of age
- Development driven in part by environmental exposures (care-giver attention, stress, diet, etc)

Human Brain Development

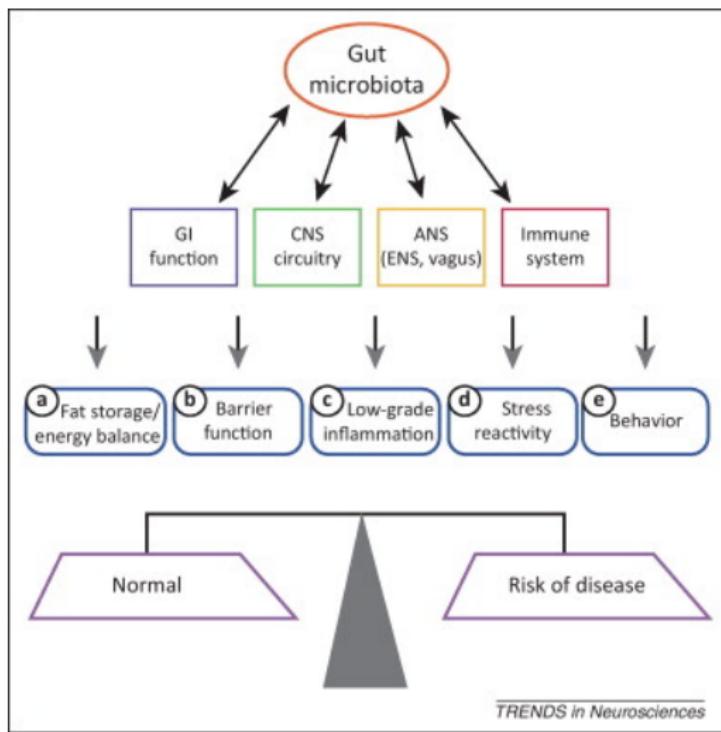
Synapse formation is dependent on early experiences



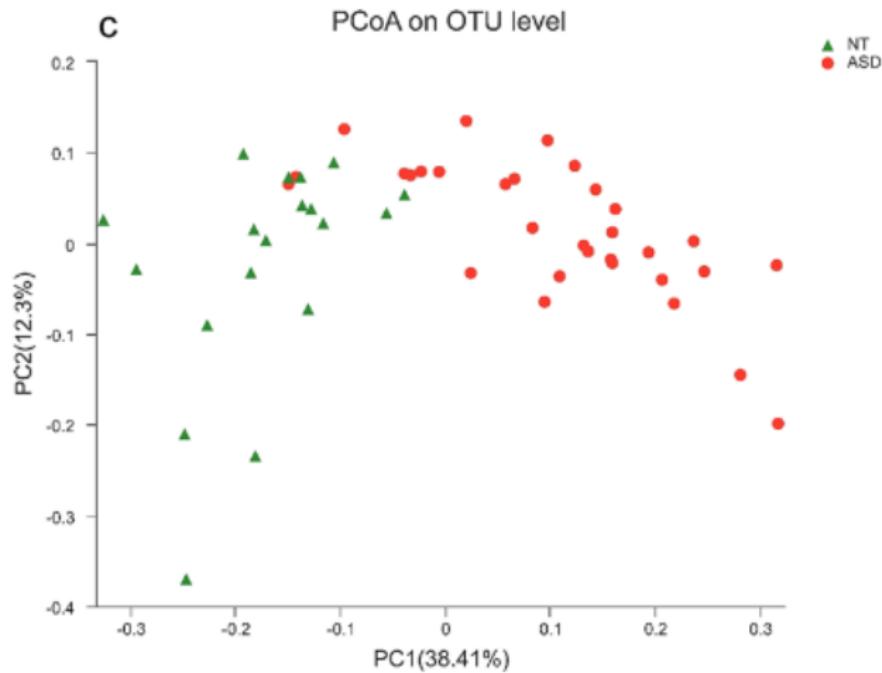
Synapse formation begins declining before Age 3

leelanauearlychildhood.org/brain-development

The gut-brain-microbiome axis: bi-directional interactions between microbiome and nervous system

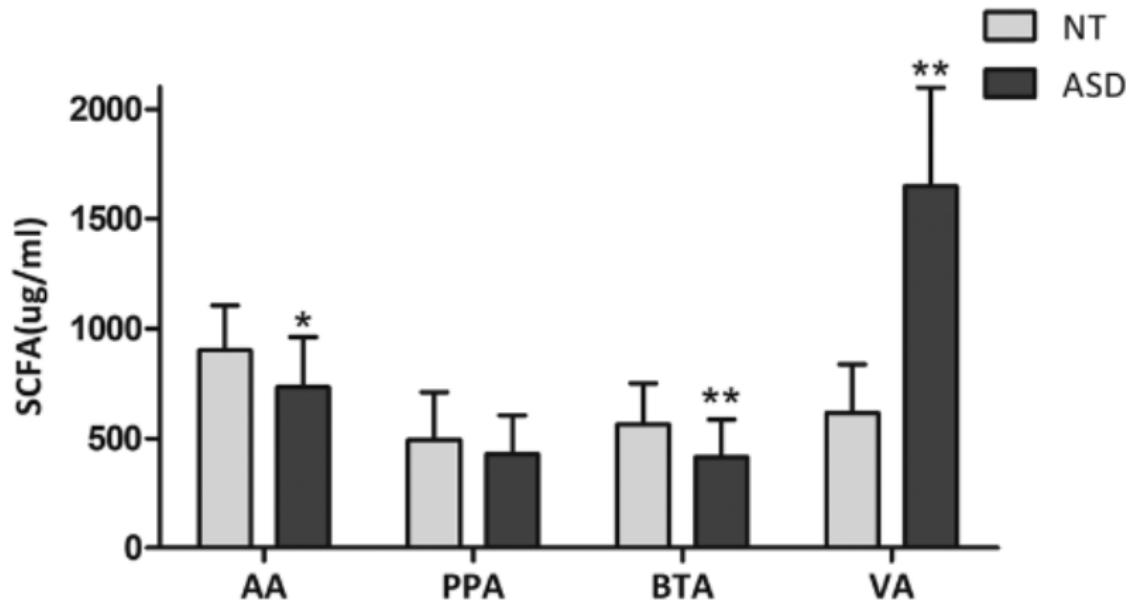


Evidence of microbiome-brain interactions from autism in humans



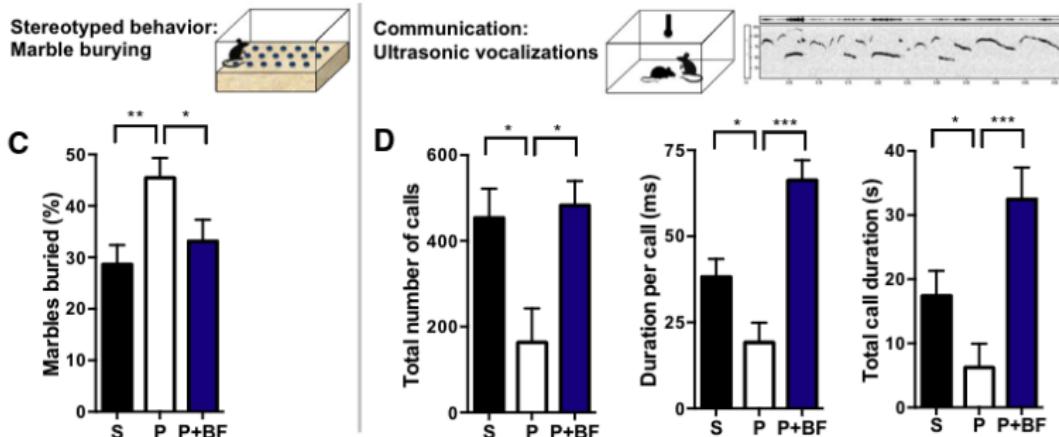
<https://doi.org/10.1038/s41598-018-36430-z>

Evidence of microbiome-brain interactions from autism in humans



<https://doi.org/10.1038/s41598-018-36430-z>

Evidence of microbiome-brain interactions from “autism” in mice



<http://dx.doi.org/10.1016/j.cell.2013.11.024>

Open questions

What effects does the microbiome have on normal brain development?

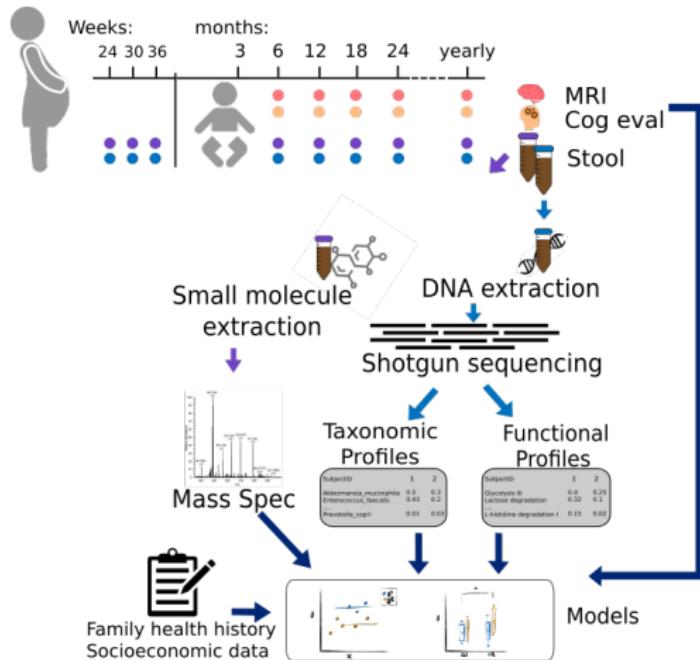
Open questions

What effects does the microbiome have on normal brain development?

What are the links between microbial metabolism, gut metabolites, and neurocognitive function?

Cohort design

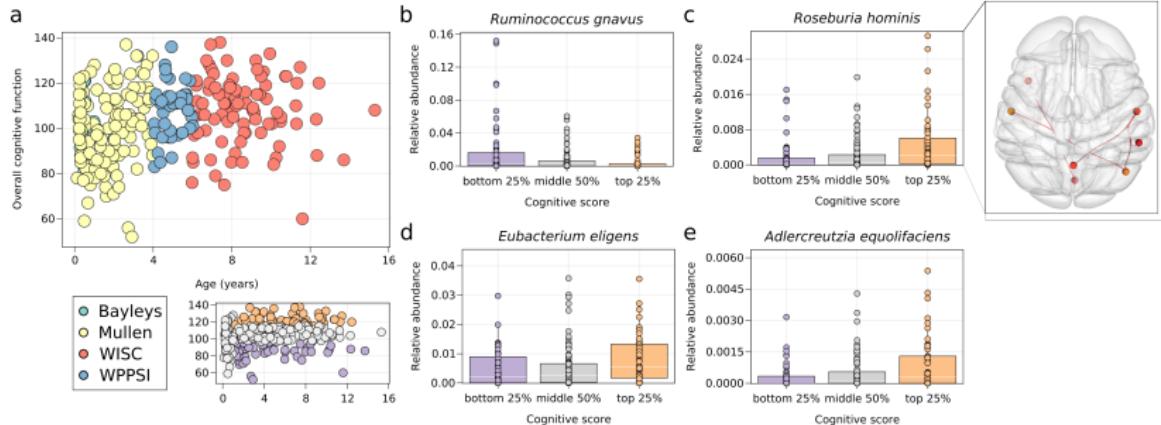
- Shotgun metagenomic sequencing
- LCMS Metabolomics
- Neuroimaging (MRI)
- Cognitive assessments (IQ-like)
- Sleep, genetics, nutrition
- Other clinical covariates



Starting with cross-sectional data:

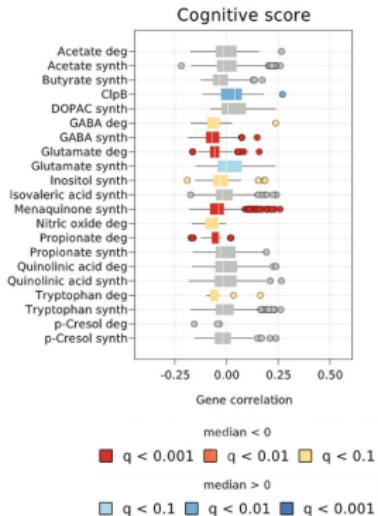
Are specific taxa, gene functions, or metabolites associated with neurocognition?

Some taxa are associated with cognition



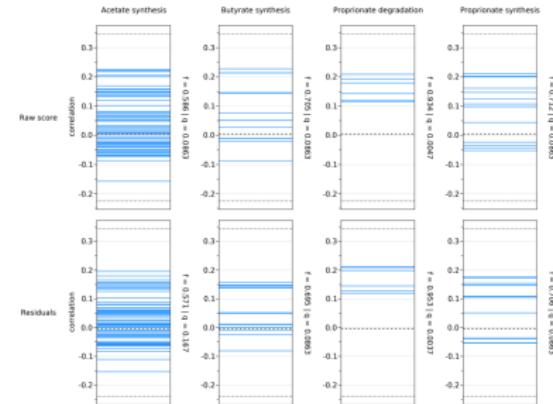
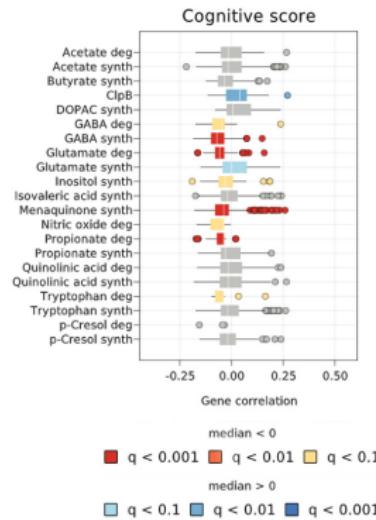
Upper and lower quartiles of cognitive function

Potentially neuroactive genes are associated with cognition and social responsiveness



FSEA of neuroactive genes for cognitive function

Potentially neuroactive genes are associated with cognition and social responsiveness

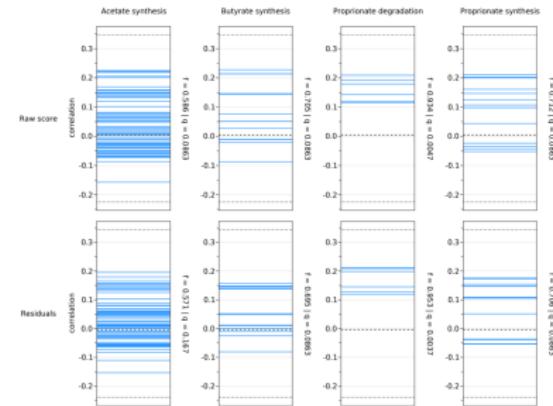
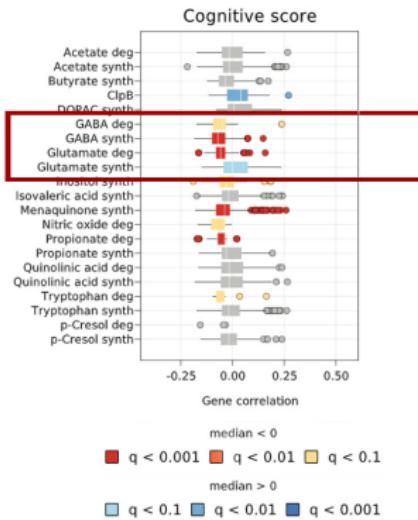


FSEA of neuroactive genes with SRS2 scores
in collaboration with Hannah Laue

FSEA of neuroactive genes for cognitive function

Valles-Colomer *et. al.* (2019)

Potentially neuroactive genes are associated with cognition and social responsiveness

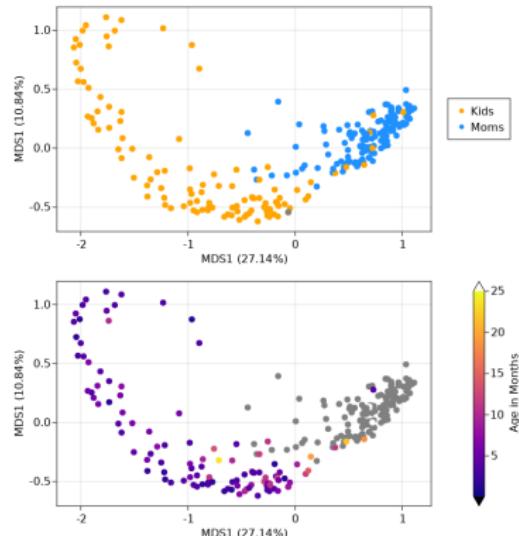


FSEA of neuroactive genes with SRS2 scores
in collaboration with Hannah Laue

FSEA of neuroactive genes for cognitive function

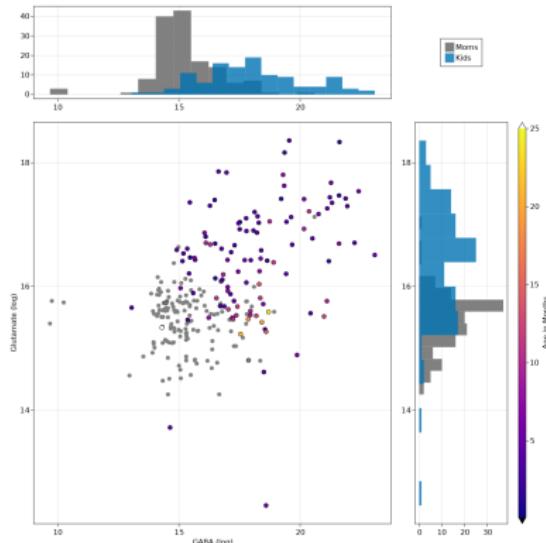
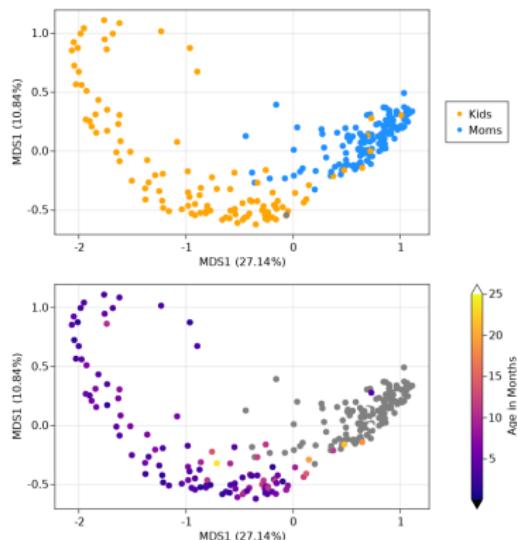
Valles-Colomer *et. al.* (2019)

Infants have very different metabolomes than their mothers



PCoA of metabolites measured
with 4 methods

Infants have very different metabolomes than their mothers



PCoA of metabolites measured with 4 methods

Question:

How are microbial metabolic potential and gut metabolome linked, esp for neuroactive genes?

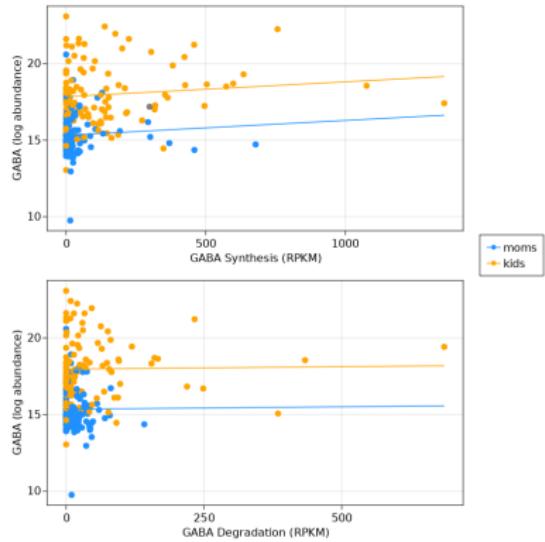
Question:

How are microbial metabolic potential and gut metabolome linked, esp for neuroactive genes?

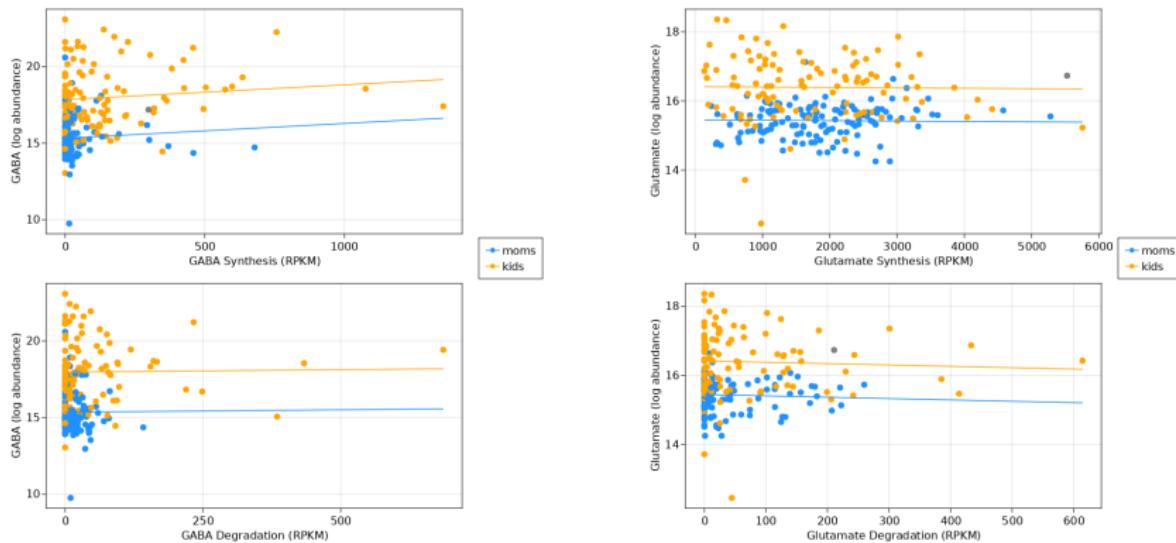
GABA and Glutamate are particularly important in neuronal development in infants

And there are known links between these molecules in the gut and brain function

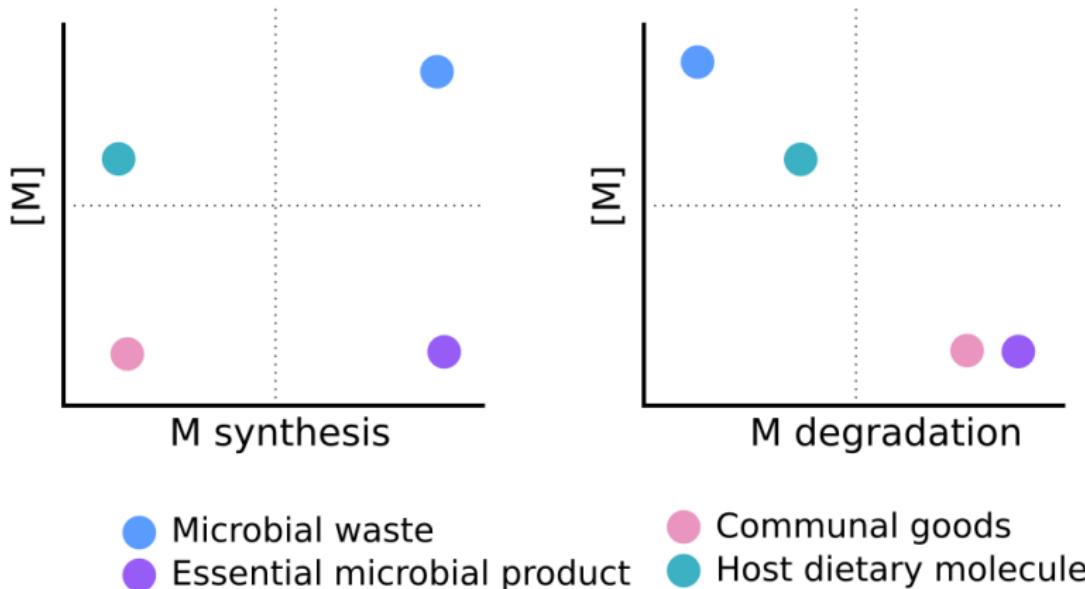
Microbial metabolism of GABA and Glutamate metabolism is not significantly associated with gut concentrations



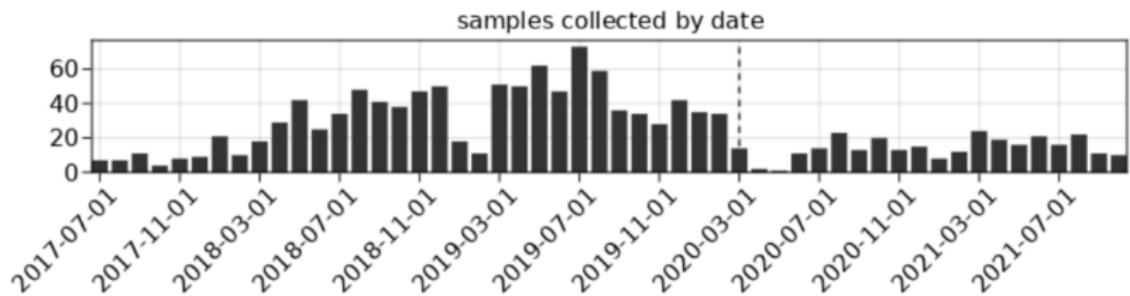
Microbial metabolism of GABA and Glutamate metabolism is not significantly associated with gut concentrations



Relationships between microbial metabolism and molecules



Limitations due to pandemic



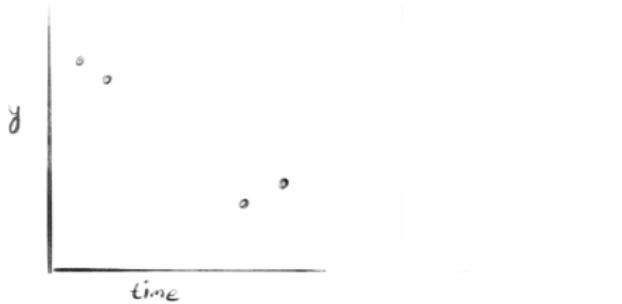
Conclusions

- The early childhood microbiome is associated with brain development
- Possible effects through SCFAs and metabolism of other neuroactive molecules
- Don't try to do longitudinal human cohort stuff during a global pandemic

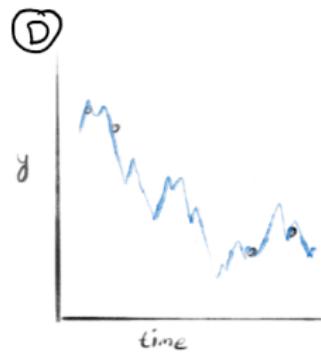
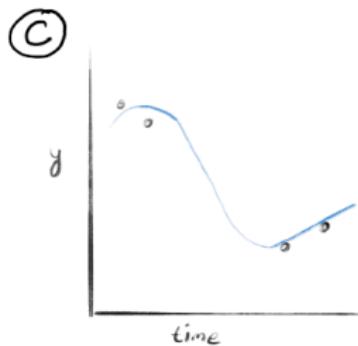
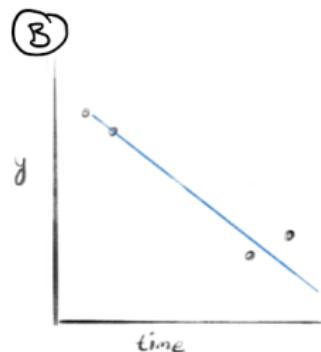
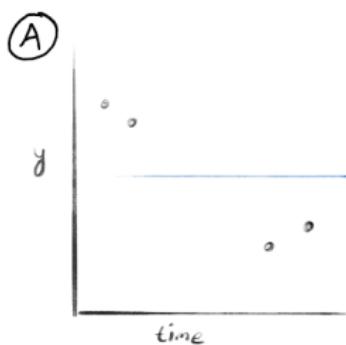
“When self and other are the same, mind and dharmas microbes are one.”

~ Hongzhi Zhengjue *Cultivating the Empty Field*

Which would you choose?



Which would you choose?



Gaussian Process models (GPs)

- Collection of arbitrarily many random variables linked by a covariance function (kernel).

Gaussian Process models (GPs)

- Collection of arbitrarily many random variables linked by a covariance function (kernel).
- Entirely described by this kernel, which dictates how “nearby” points relate to each other.

Gaussian Process models (GPs)

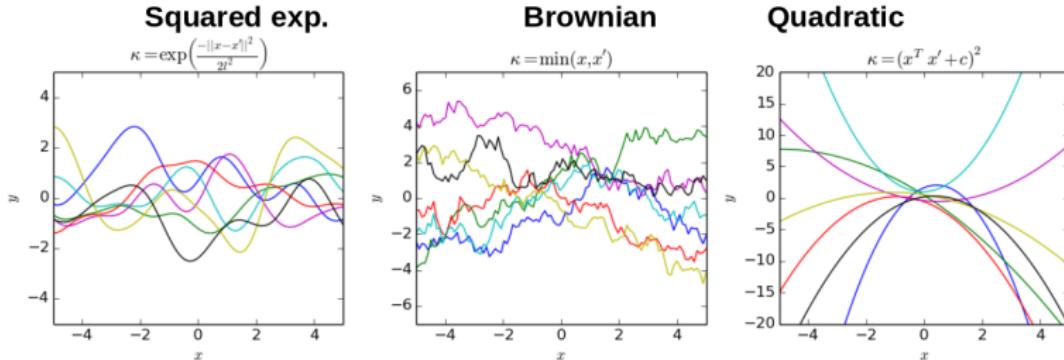
- Collection of arbitrarily many random variables linked by a covariance function (kernel).
- Entirely described by this kernel, which dictates how “nearby” points relate to each other.
- Typically over time, but can be any other continuous variable).

Gaussian Process models (GPs)

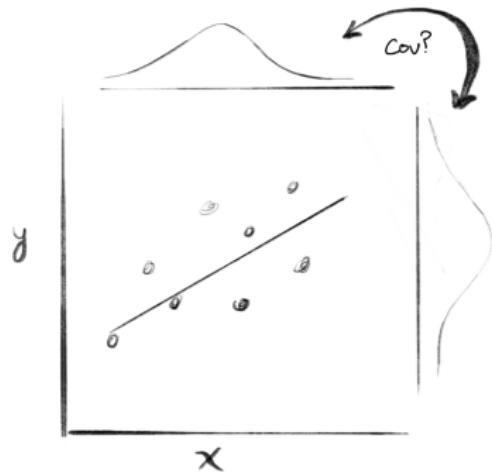
- Collection of arbitrarily many random variables linked by a covariance function (kernel).
- Entirely described by this kernel, which dictates how “nearby” points relate to each other.
- Typically over time, but can be any other continuous variable).
- More constrained than freeform nonparametric Bayes, more flexible than regression.

Gaussian Process models (GPs)

- Collection of arbitrarily many random variables linked by a covariance function (kernel).
- Entirely described by this kernel, which dictates how “nearby” points relate to each other.
- Typically over time, but can be any other continuous variable).
- More constrained than freeform nonparametric Bayes, more flexible than regression.

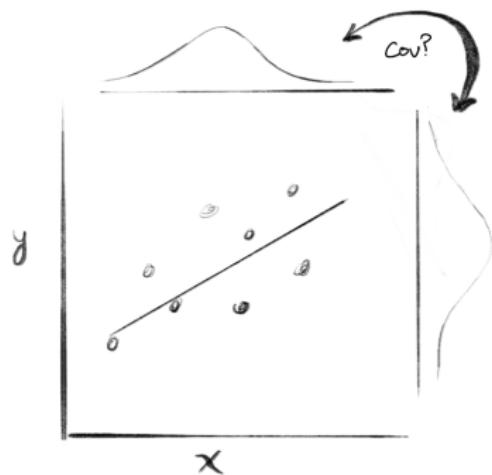


Linear models vs GPs

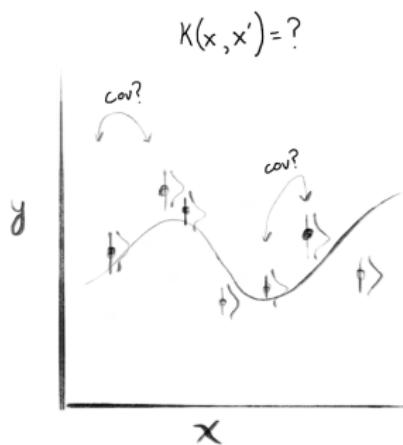


Fitting a linear model

Linear models vs GPs

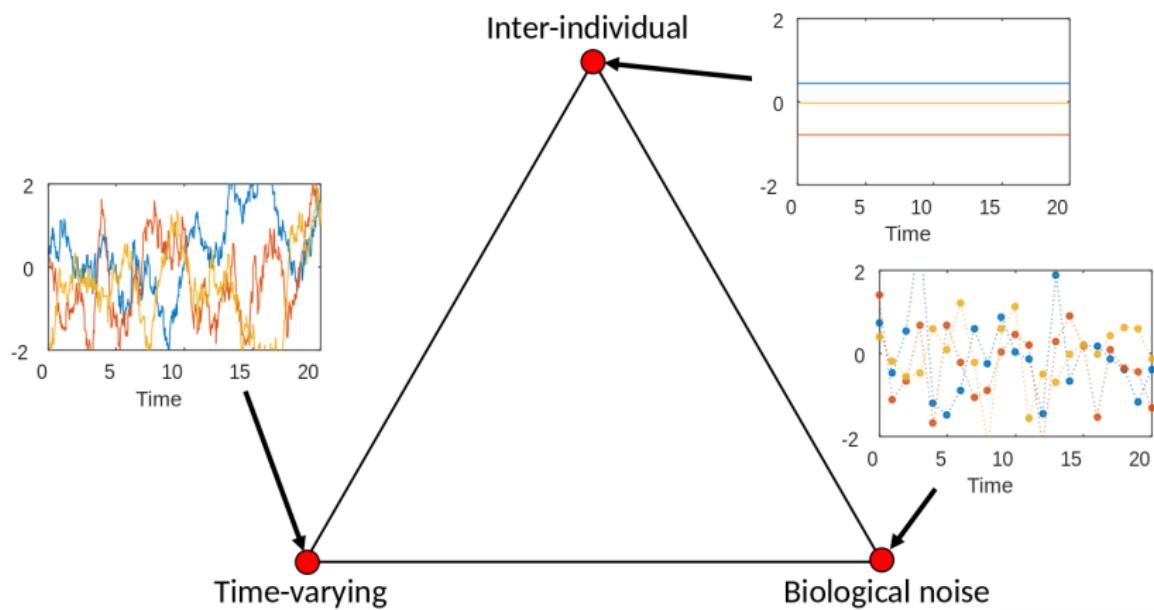


Fitting a linear model



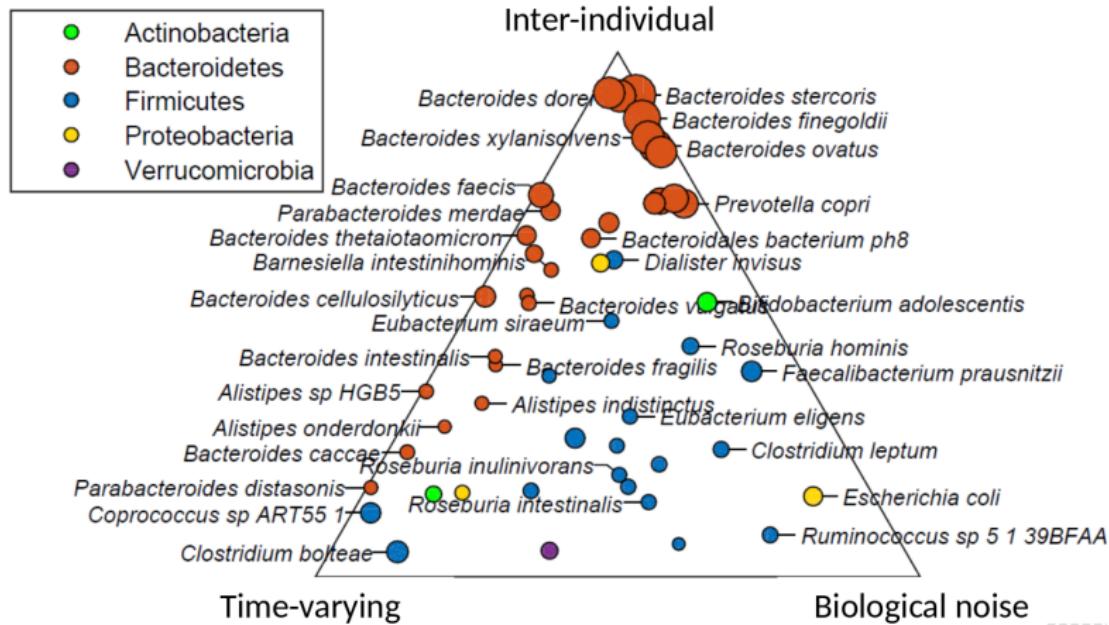
Fitting a GP

GPs in HMP1-II to model different sources of variation



Credit: Jason Lloyd-Price <https://doi.org/10.1038/nature24485>

GPs in HMP1-II to model different sources of variation



Credit: Jason Lloyd-Price <https://doi.org/10.1038/nature24485>

GPs in the literature to model other dynamic processes

Method

Detecting differential growth of microbial populations with Gaussian process regression

Peter D. Tonner,^{1,2} Cynthia L. Darnell,² Barbara E. Engelhardt,³ and Amy K. Schmid^{1,2}

¹Program in Computational Biology and Bioinformatics, Duke University, Durham, North Carolina 27708, USA; ²Biology Department, Duke University, Durham, North Carolina 27708, USA; ³Computer Science Department, Center for Statistics and Machine Learning, Princeton University, Princeton, New Jersey 08540, USA

Bioinformatics, 34(3), 2018, 372–380

doi: 10.1093/bioinformatics/btx549

Advance Access Publication Date: 13 September 2017

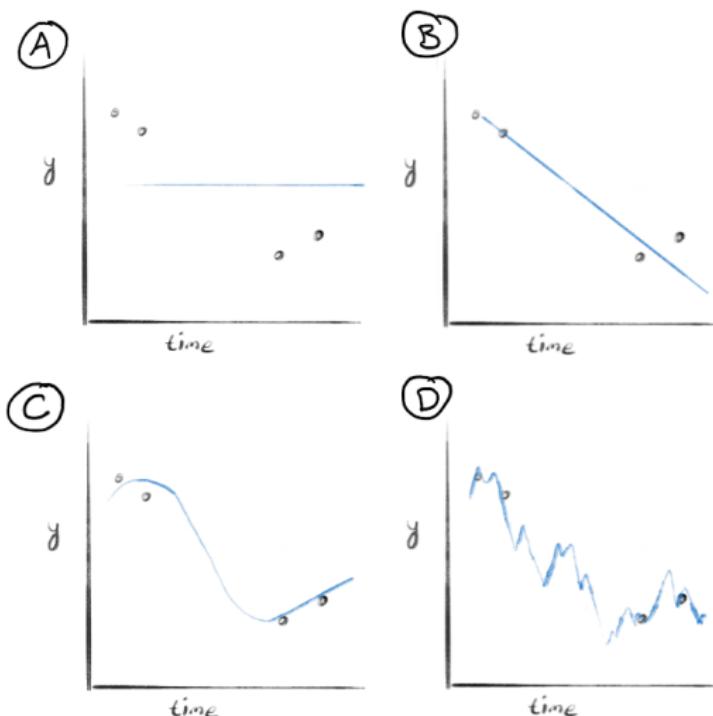
Original Paper

Genome analysis

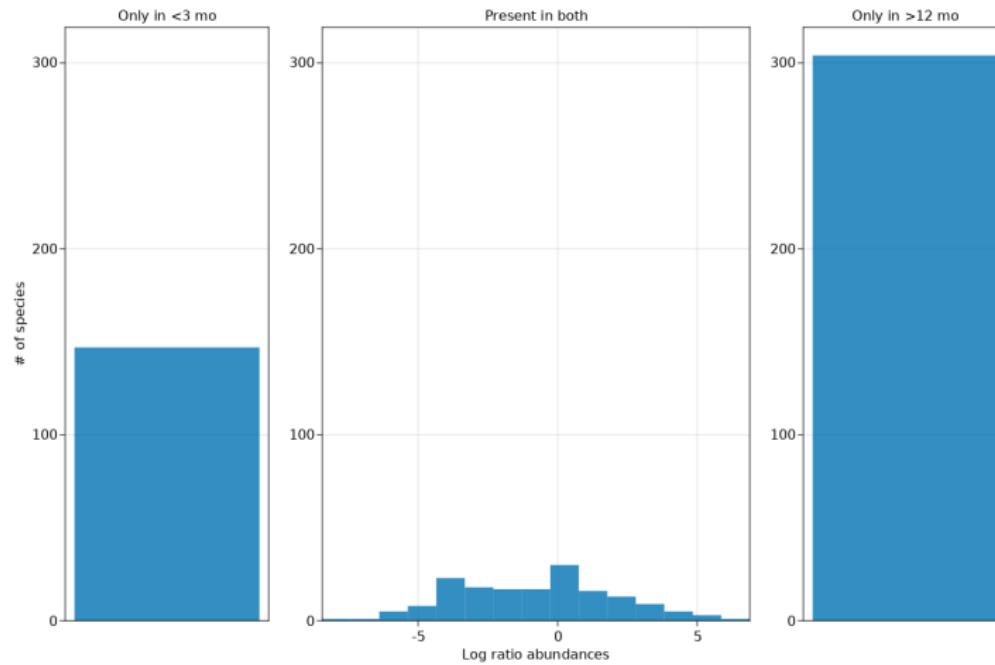
Temporal probabilistic modeling of bacterial compositions derived from 16S rRNA sequencing

Tarmo Äijö^{1,*}, Christian L. Müller¹ and Richard Bonneau^{1,2,3,*}

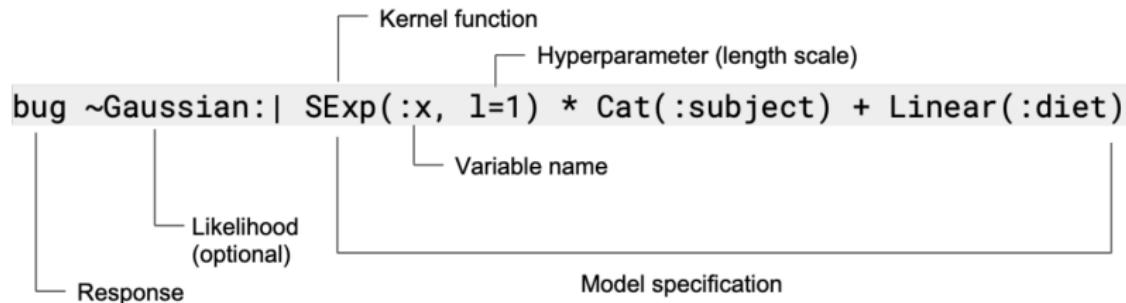
GPs could provide more realistic model for longitudinally sampled microbiomes



Could GPs better handle transitions between infant and adolescent communities?

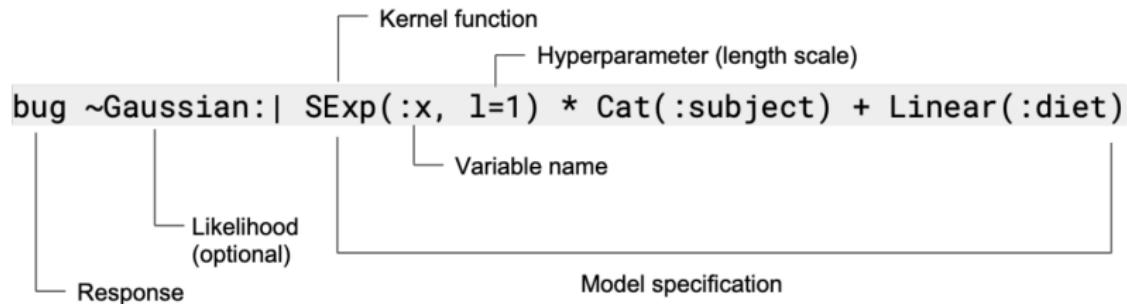


(Ga)ussian (P)rocess models for (L)ongitudinal (A)nalysis of (C)ommunities



<https://en.wikipedia.org/wiki/Geplak>

(Ga)ussian (P)rocess models for (L)ongitudinal (A)nalysis of (C)ommunities



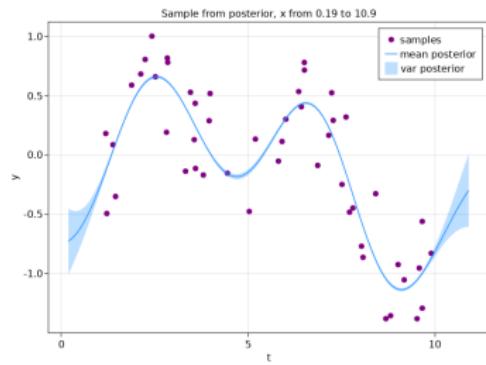
<https://en.wikipedia.org/wiki/Geplak>

Goal

Make fitting GPs as easy as fitting LMs.

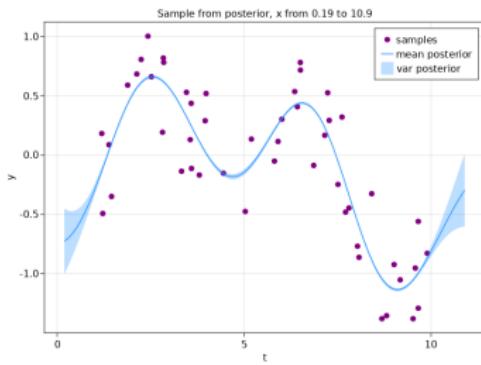
Current Functionality - Sample (simulate)

```
> gaplac -v sample "y :~| SqExp(:t', l=1.5)" \  
--at "t = rand(Uniform(1,10), 50)" \  
--output data_sqexp.tsv \  
--plot notes/assets/sqexpplot.png
```

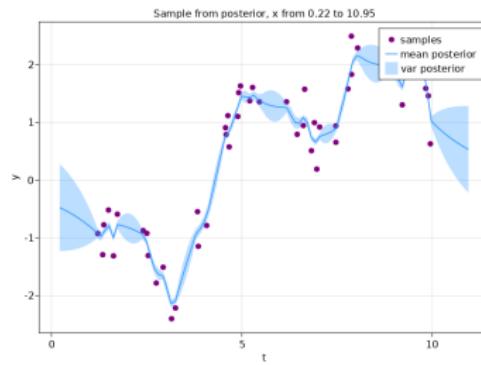


Current Functionality - Sample (simulate)

```
> gaplac -v sample "y :~| SqExp(:t; l=1.5)" \  
--at "t = rand(Uniform(1,10), 50)" \  
--output data_sqexp.tsv \  
--plot notes/assets/sqexpplot.png
```

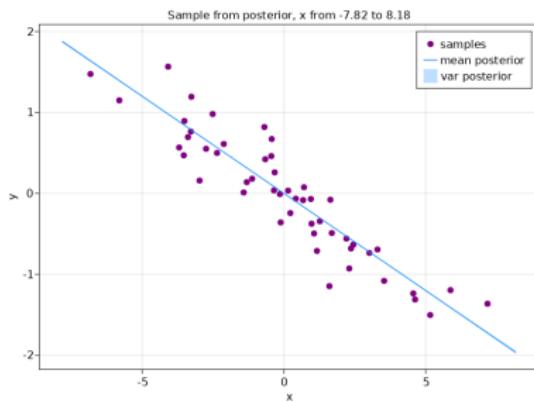


```
> gaplac -v sample "y :~| OU(:t; l=1.5)" \  
--at "t = rand(Uniform(1,10), 50)" \  
--output data_ou.tsv \  
--plot notes/assets/ouplot.png
```



Current Functionality - Sample (simulate)

```
> ./gaplac -v sample "y :~| Linear(:x)" \  
--at "x = rand(Normal(0,3), 50)" \  
--output data_linear.tsv \  
--plot notes/assets/linearplot.png
```



Current Functionality - MCMC (fit / infer parameters)

```
❯ ./gaplac -v mcmc "y ~| OU(t)" \
    --data data_ou_l3.tsv \
    --output mcmc_ou.tsv \
    --samples 500 --infer 1
[ Info: Getting started!
[ Info: running 'mcmc'
[ Info: Found initial step size
[   ε = 0.4
Sampling 19% |███████████| ETA: N/A
| ETA: 0:01:13
```

Current Functionality - MCMC (fit / infer parameters)

```
› ./gaplac -v mcmc "y ~| OU(t)" \
    --data data_ou_l3.tsv \
    --output mcmc_ou.tsv \
    --samples 500 --infer 1
[ Info: Getting started!
[ Info: running 'mcmc'
[ Info: Found initial step size
[ Info: ε = 0.4
Sampling 19% |███████████| ETA: N/A
| ETA: 0:01:13
```

iteration	chain	acceptance_rate	fx[1]	fx[2]	fx[3]
1	1	0.9725847795284468	0.9226902307329367	0.876145988637838	1.17506
2	1	0.4364405143780374	0.7228242923386721	0.8541957426349509	0.64525
3	1	0.8917148506448946	1.063428648042305	1.2473254712747368	1.65505
4	1	0.9945385413302966	0.7685081280868516	1.6488844595271832	1.54130
5	1	0.6372594283619787	0.9867805712396475	1.813864011366549	2.07370
6	1	0.9634410811380992	0.4355375598469431	0.39543550617160694	0.44463
7	1	0.6236477856318113	-0.3030181226395457	0.7981933705807881	1.05997
8	1	0.9766185475287668	0.32835627064683715	1.1838954278976836	1.56570
9	1	0.7039144715653907	0.2346673985446619	1.3697179368081789	1.38056
10	1	0.7937399131149093	0.14584204022056943	0.8958713896766708	1.05198

Current Functionality - Select (compare models)

```
> ./gaplac -v select --formulae "y ~| SqExp(:t, l=3)" "y ~| OU(:t, l=3)" --data data_ou_l3.tsv
[ Info: Getting started!
[ Info: running 'select'
Info:
[ Dict{String, Any} with 4 entries:
  "plot" => nothing
  "formulae" => Any["y ~| SqExp(:t, l=3)", "y ~| OU(:t, l=3)"]
  "data" => "data_ou_l3.tsv"
  "chains" => Any[]
Info: Log2 Bayes: -3.305
  • Log(pdf) - model 1: -43.4752
  • Log(pdf) - model 2: -40.1704
Note - Positive values indicate more evidence for model 1
```

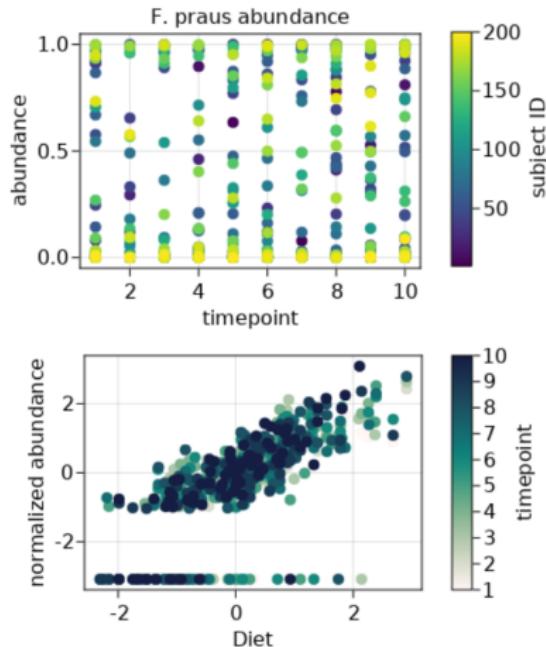
Current Functionality - Select (compare models)

```
./gplac -v select --formulae "y ~| SqExp(:t, l=3)" "y ~| OU(:t, l=3)" --data data_ou_l3.tsv
[ Info: Getting started!
[ Info: running 'select'
Info:
Dict{String, Any} with 4 entries:
  "plot" => nothing
  "formulae" => Any["y ~| SqExp(:t, l=3)", "y ~| OU(:t, l=3)"]
  "data" => "data_ou_l3.tsv"
  "chains" => Any[]
Info: Log2 Bayes: -3.305
  • Log(pdf) - model 1: -43.4752
  • Log(pdf) - model 2: -40.1704
Note - Positive values indicate more evidence for model 1
```

Performance on simulated data - it works!

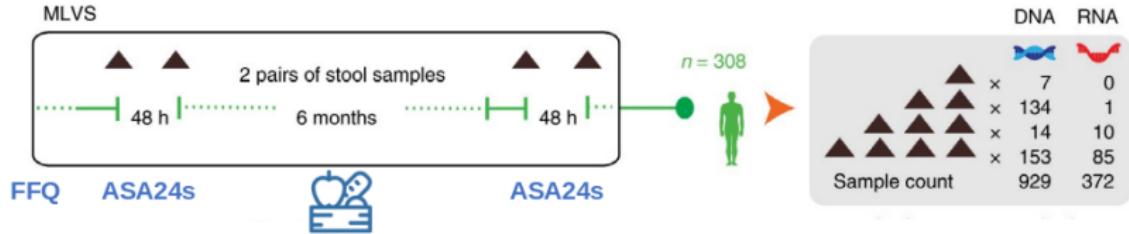
Synthetic Data:

- 200 subjects, 300 “bugs”.
- One bug “*Faecalibacterium prausnitzii*” linearly associated with continuous variable “diet.”
- Modeled from real HMP gut microbiomes using SparseDOSSA 2.
- Diet = simple continuous variable.
- Three (randomly sampled) time points / individual.
- Compared Bayes factor (essentially, fit) for model with or without diet variable



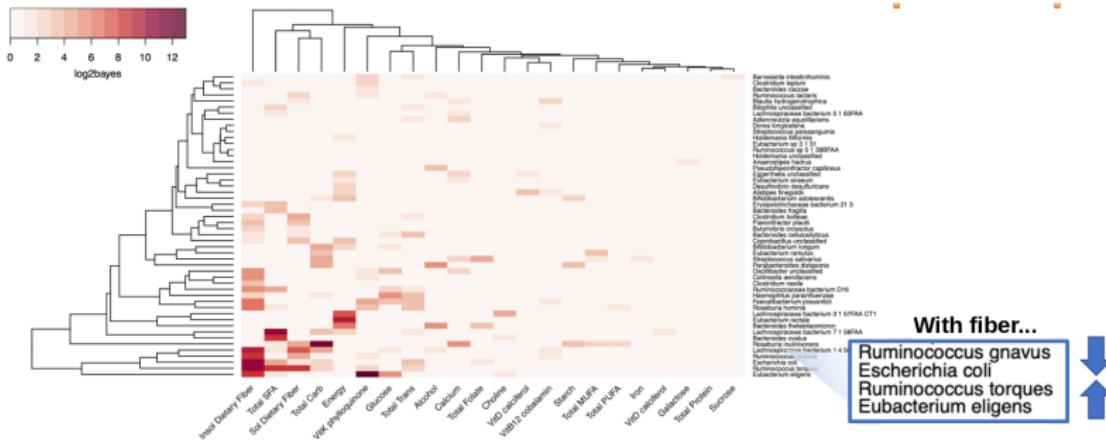
Special thanks: Meg

Conclusions from real-world data



Special thanks: Tobyn

Conclusions from real-world data



- Comparing model: $\text{Bug} \sim \text{SqExp}(\text{time}) * \text{Categorical(subject)}$
- With or without $\text{Linear}(\text{diet})$

Special thanks: Tobyn

Next steps

- Figure out what's going on with MCMC sampling
- Incorporate non-Gaussian likelihoods (similar to link functions in LMs)
- Allow different AD backends / sampling methodology (eg NUTS)
- Benchmark against MaAsLin on simulated data
- Compare against MaAsLin on real-world data

Thanks!

Vanja Klepac-Ceraj

Shelley McCann

Sophie Rowland

Danielle Peterson

Lauren Tso

Anika Luo

Annelle Abatoni

Alexa Gross

Juliette Madan

Hannah Laue

Curtis Huttenhower

Jason Lloyd-Price

Tobyn Branck

Meg Short

Andrew Ghazi

Eric Franzosa

<http://www.infinitecuriosity.org/vizgp/>