

Using mass spectrometry to
discover new proteins

"Nothing in Biology Makes Sense Except in
the Light of Evolution"
- Theodosius Dobzhansky

John Anders

University Leipzig

Leipzig - February, 2019

- My phd-project
 - SIHUMI
 - The ms/ms-thingy
 - The present
 - The new and shiny future

- My phd-project
- SIHUMI
- The ms/ms-thingy
- The present
- The new and shiny future

- My phd-project
- SIHUMI
- The ms/ms-thingy
- The present
- The new and shiny future

- My phd-project
- SIHUMI
- The ms/ms-thingy
- The present
- The new and shiny future

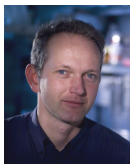
- My phd-project
- SIHUMI
- The ms/ms-thingy
- The present
- The new and shiny future



Small Proteins in Prokaryotes, an Unexplored World



Small Proteins in Prokaryotes, an Unexplored World



Martin von Bergen



Hannes Petruschke



Nico Jehmlich



Small Proteins in Prokaryotes, an Unexplored World



Martin von Bergen



Nico Jehmlich

Anders (Leipzig)



Hannes Petruschke

Mass Spectrometry



UNIVERSITÄT
LEIPZIG



John Anders



Peter Stadler



Sven Findeiss

Leipzig - February, 2019

3 /
16



Florian Eggenhofer

Small Proteins in Prokaryotes, an Unexplored World



Martin von Bergen



Hannes Petruschke

UNIVERSITÄT
LEIPZIG

John Anders



Peter Stadler



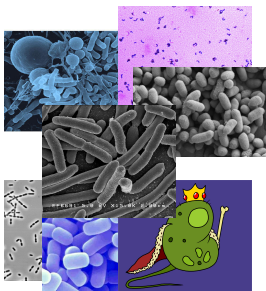
Nico Jehmlich



Sven Findeiss



simplified human intestinal microbiota
(SIHUMI)
[Becker 2001]

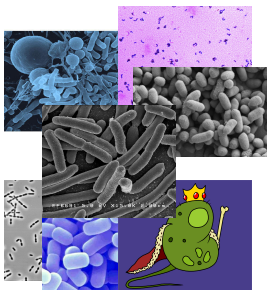


from indiamart.com



simplified human intestinal microbiota
(SIHUMI)
[Becker 2001]

Anaerostipes caccae
Bacteroides thetaiotaomicron
Bifidobacterium longum
Blautia producta
Clostridium ramosum
Escherichia coli and
Lactobacillus plantarum



from indiamart.com

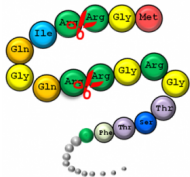


simplified human intestinal microbiota
(SIHUMI)
[Becker 2001]

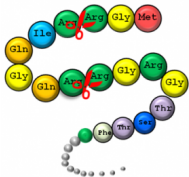
Anaerostipes caccae
Bacteroides thetaiotaomicron
Bifidobacterium longum
Blautia producta
Clostridium ramosum
Escherichia coli and
Lactobacillus plantarum



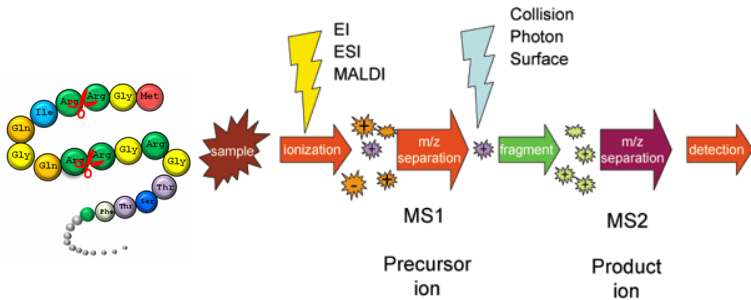




Protein

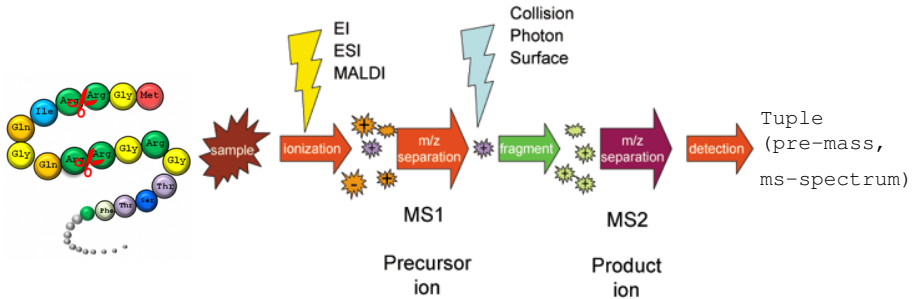


Protein



Protein

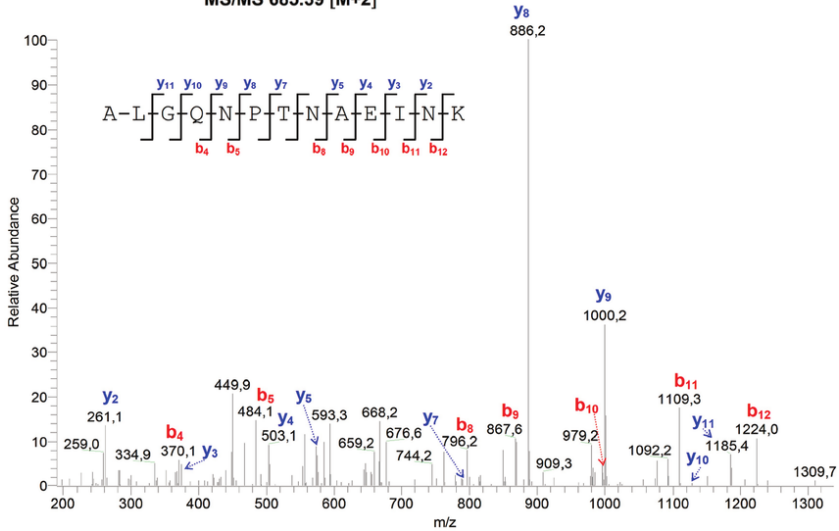
Peptide



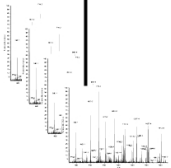
Protein

Peptide

MS/MS 685.39 [M+2]²⁺



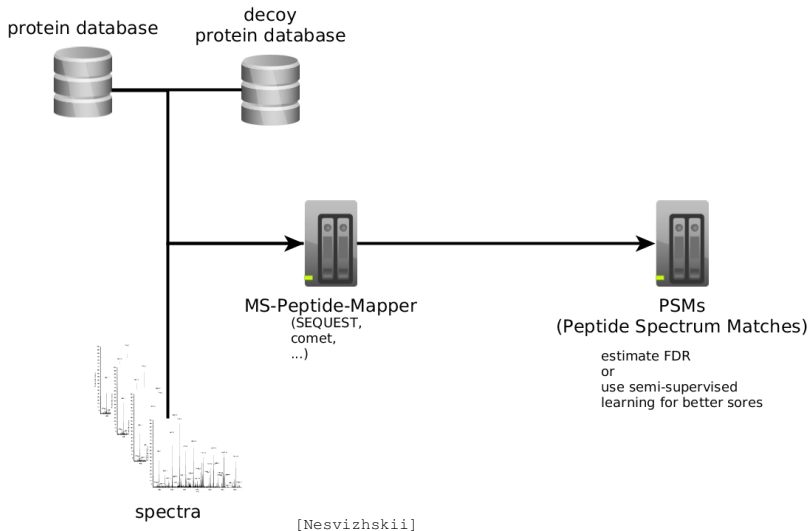
protein database



MS-Peptide-Mapper
(SEQUEST,
comet,
...)



PSMs
(Peptide Spectrum Matches)

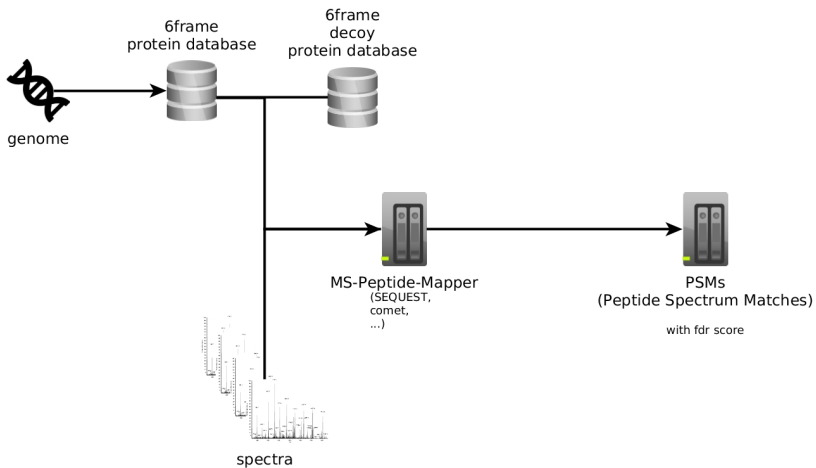


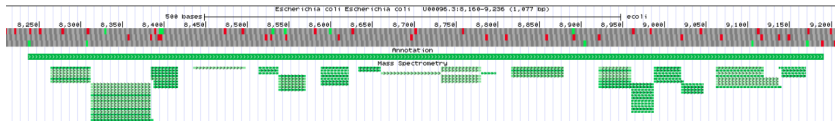
"Nothing in ~~Biology~~ Mass Spectrometry Makes
Sense Except in the Light of ~~Evolution~~ History"

```

1 > sixFrame
2     S F S F * L Q R A I C L C V D * K K S V   F1
3     A F H S D C N G Q Y V S V W I K K R V S   F2
4     L F I L T A T G N M S L C G L K K E C L   F3
5     1 AGCTTTTCATTCTGACTGCAACGGGCAATATGTCTCTGTGGATTAAGAGTGTC 60
6     ----:----|----:----|----:----|----:----|----:----|----:----|
7     1 TCGAAAAGTAAGACTGACGTTGCCGTTATACAGAGACACACCTAATTTTTTCTCACAG 60
8     L K E N Q S C R A I H R Q T S * F F L T   F6
9     X S K M R V A V P L I D R H P N F F S H   F5
10    A K * E S Q L P C Y T E T H I L F L T D   F4
11

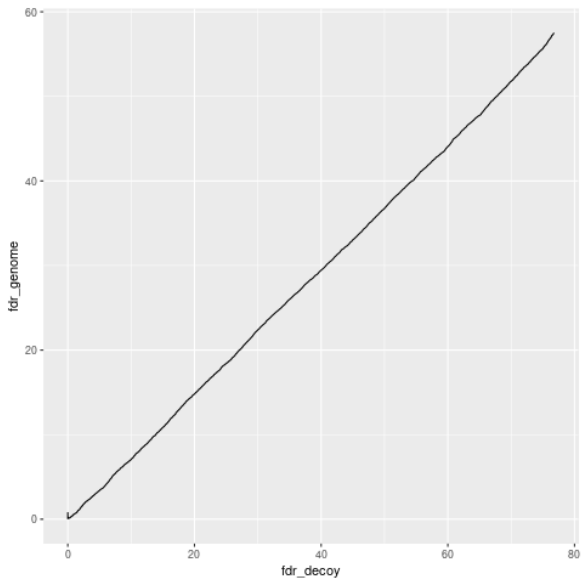
```





- `fdr_decoy` = % of false PSMs predict by decoy-DB
- `fdr_genome` = % of PSMs on not annotated frames

- `fdr_decoy` = % of false PSMs predict by decoy-DB
- `fdr_genome` = % of PSMs on not annotated frames



- True de-novo prediction is not feasible
- Current method compare against a protein database
- I want to use direct mapping on to the genome
- More error but also more context. How to use it?
- Convince biological community
- Compare prediction of genome mapping against extended protein database

- True de-novo prediction is not feasible
- Current method compare against a protein database
- I want to use direct mapping on to the genome
- More error but also more context. How to use it?
- Convince biological community
- Compare prediction of genome mapping against extended protein database

- True de-novo prediction is not feasible
- Current method compare against a protein database
- I want to use direct mapping on to the genome
- More error but also more context. How to use it?
- Convince biological community
- Compare prediction of genome mapping against extended protein database

- True de-novo prediction is not feasible
- Current method compare against a protein database
- I want to use direct mapping on to the genome
- More error but also more context. How to use it?
- Convince biological community
- Compare prediction of genome mapping against extended protein database

- True de-novo prediction is not feasible
- Current method compare against a protein database
- I want to use direct mapping on to the genome
- More error but also more context. How to use it?
- Convince biological community
- Compare prediction of genome mapping against extended protein database

- True de-novo prediction is not feasible
- Current method compare against a protein database
- I want to use direct mapping on to the genome
- More error but also more context. How to use it?
- Convince biological community
- Compare prediction of genome mapping against extended protein database



Thank you
and I am happy for some input.