

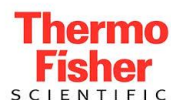
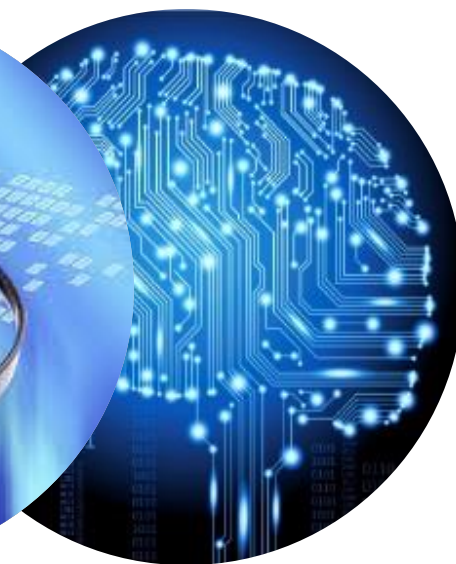
Safer Food with BIG data

The Bigger, the Better?



Alex Bossers (WBVR) & Martijn Bouwknecht (Vion)

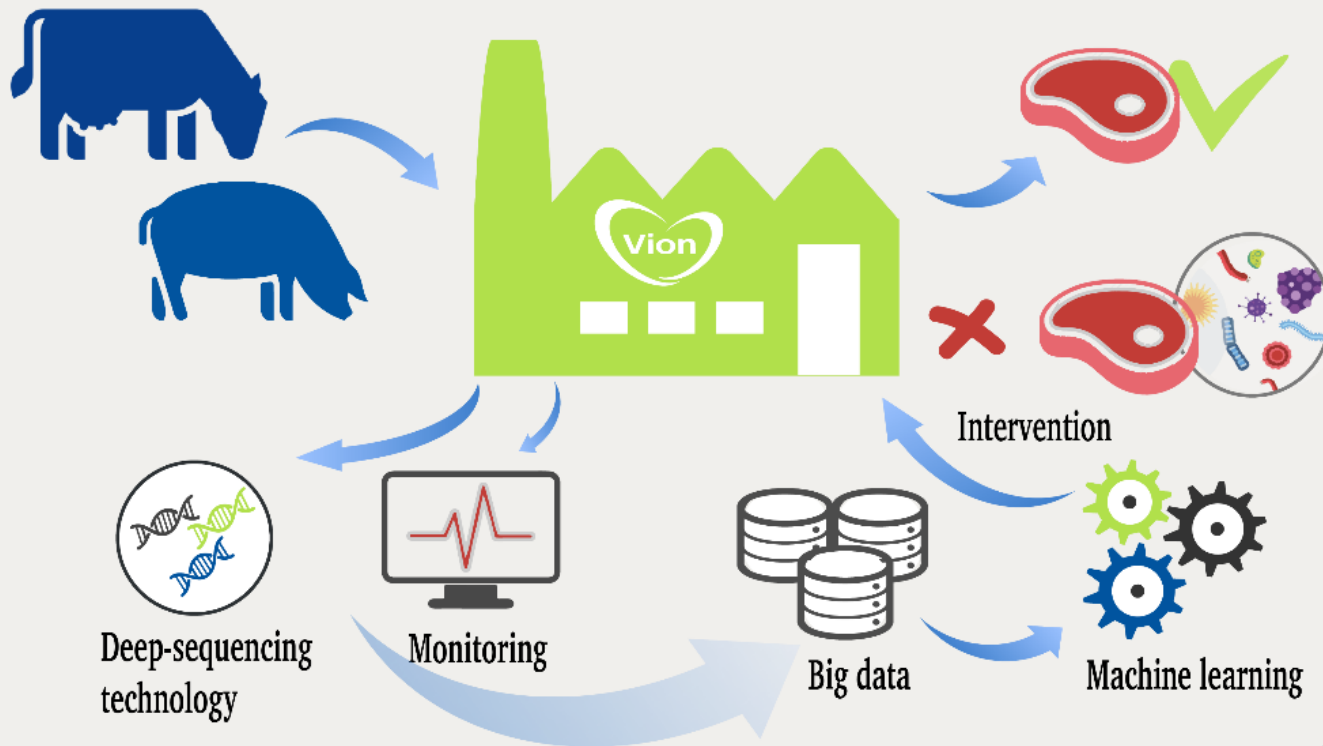
1H4F Najaarscongres
8 November 2018



Summary

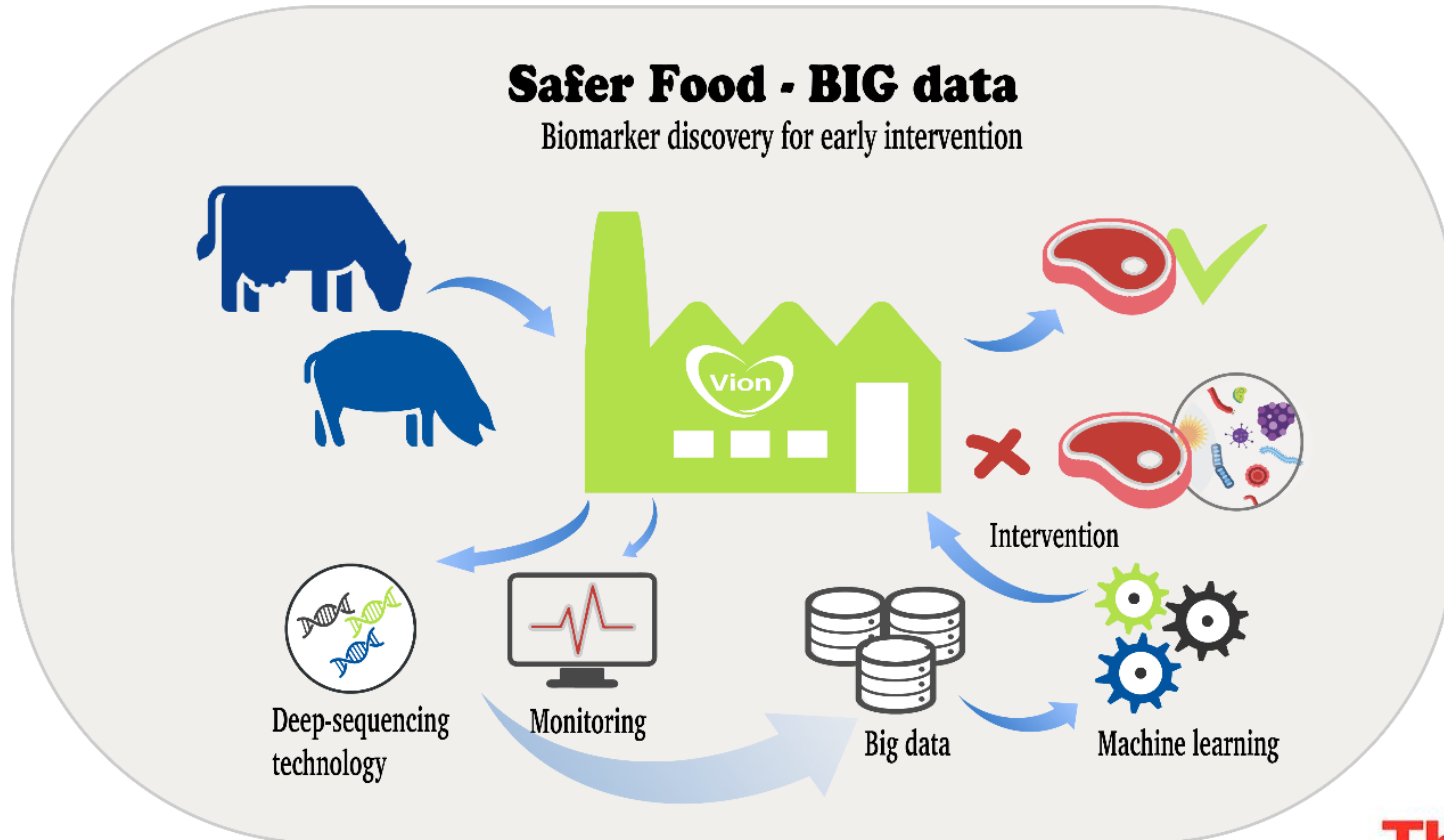
Safer Food - BIG data

Biomarker discovery for early intervention



- Evaluation of methodologies for sample processing
- Deep sequencing and biomarker discovery
- Pathogen clonality analysis by WGS
- WBVR link WU Food Microbiology laboratory

- Data storage and integration
- WATSON: Machine learning - Identify factors leading to emergence of food-borne pathogens



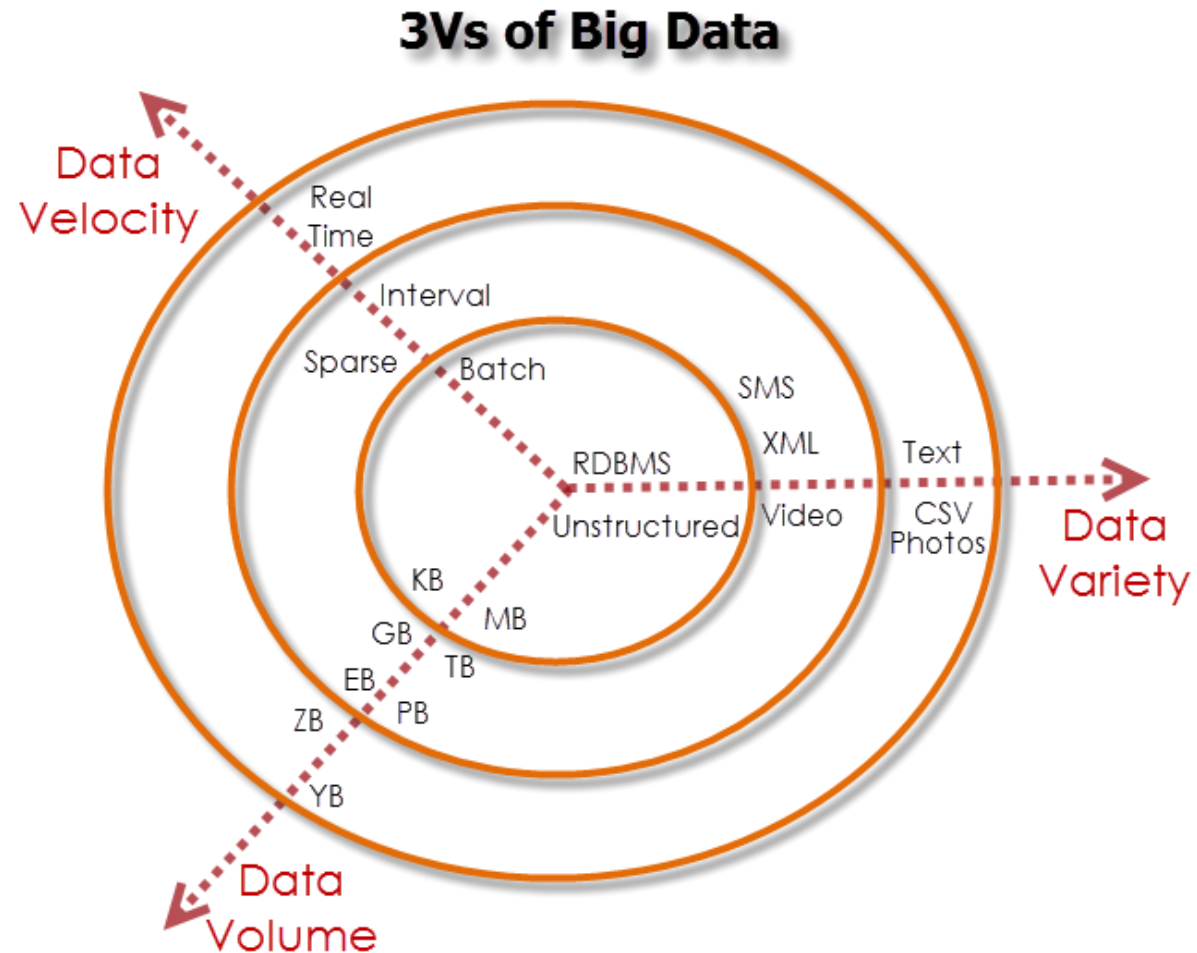
- Delivery of samples,
- meta-data on environmental, physical, biological and chemical variables.
- Intervention strategies
- ...

- Knowledge and tools for sample preservation, processing and preparation of RNA/DNA
- Biomarker screening / detection kits, sample prep kits.

BIG data

■ 3Vs

- Velocity
- Volume
- Variety



BIG data

■ 3Vs

- Velocity
- Volume
- Variety

■ Types

- Sensor data (temperature, humidity, belt speed)
- Continuous generic microbial monitoring
- Specific data types like microbiomes, imaging, ...



PPS – Safer Food with BIG data



- Develop blue print how to use BIG data approaches to enable early detection of future potential product contamination
 - Blue print can be re-used for other agents / species / products / slaughterhouses

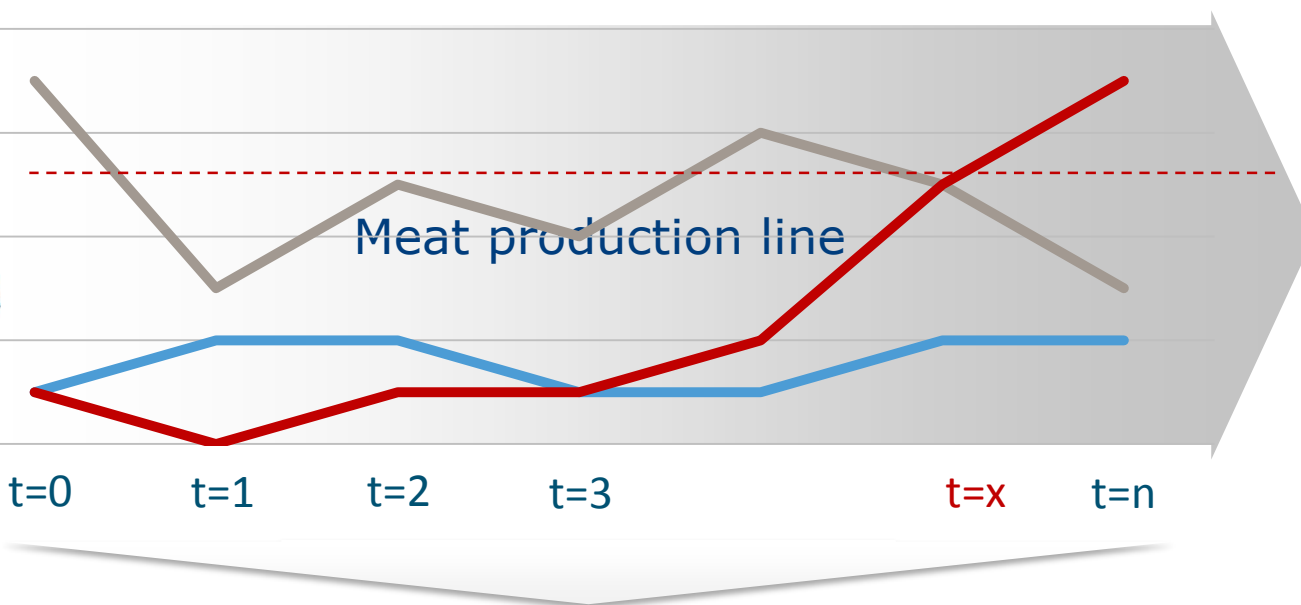
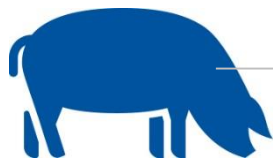
- Demonstration use case:
 - Prevention of Salmonella contamination pig meat
 - Develop methods to determine early (bio)markers
 - Using BIG data methodologies
 - Model intervention strategies

“DATA”

Environmental factors (humidity/temperature)

Other “routine” lab-measurements

Microbiomes



Contaminated products



Internal recall

External recall

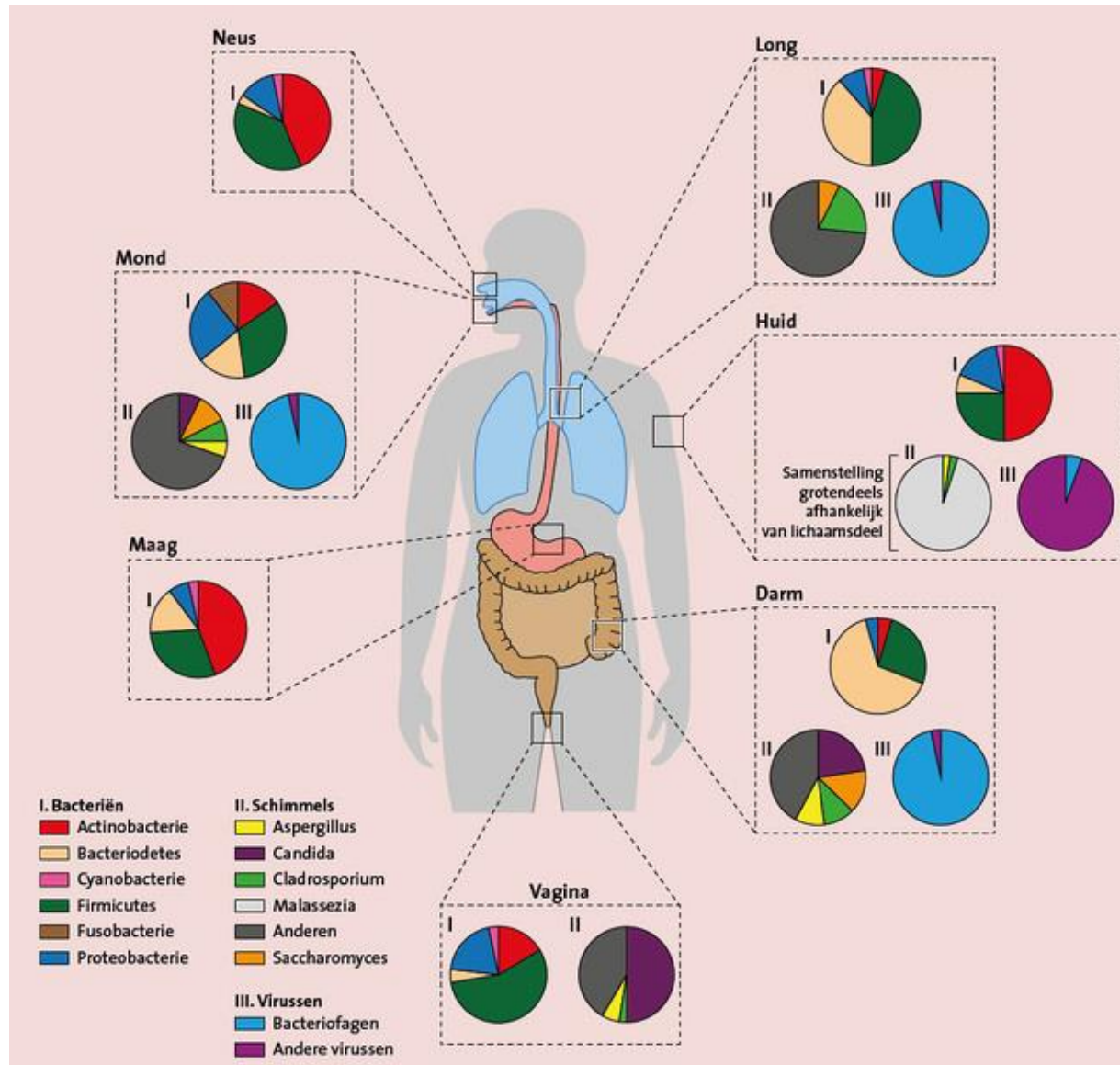


Microbiom



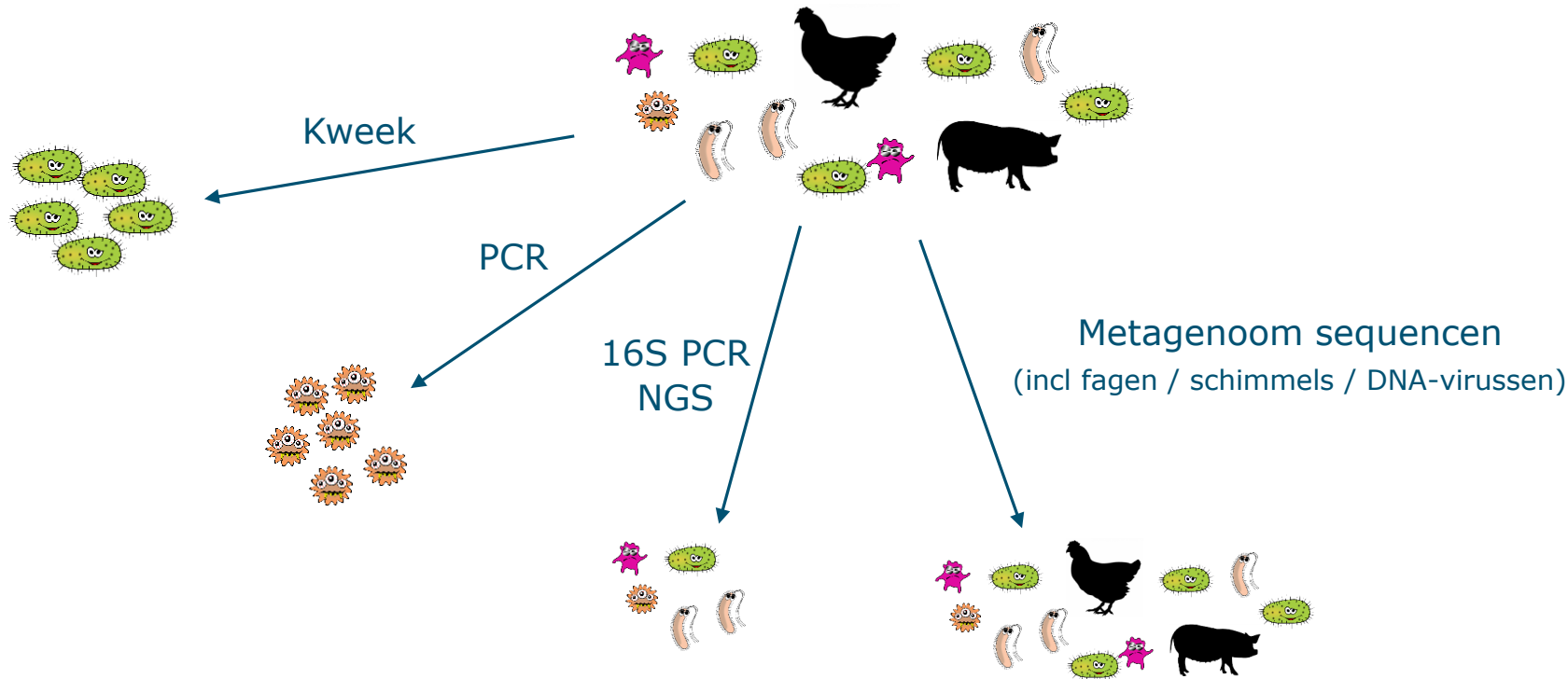
Microbioom samenstelling - functie

Spatio- (en temporale) variabiliteit



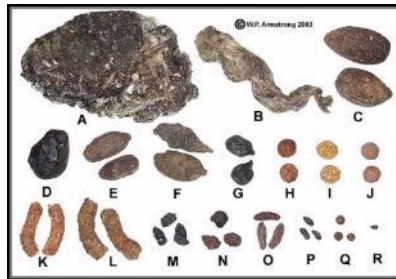
Microbioom screenen

- Metagenoom = alle aanwezige DNA (ook van evt gastheer/omgeving)



Metagenomic profiling

Deep-sequencing



DNA extraction



Total DNA
(chromosomal + plasmid)

Fragmentation
Library generation

16S-amplicons



Deep-sequencing

Environmental sample

Typical metagenome analysis

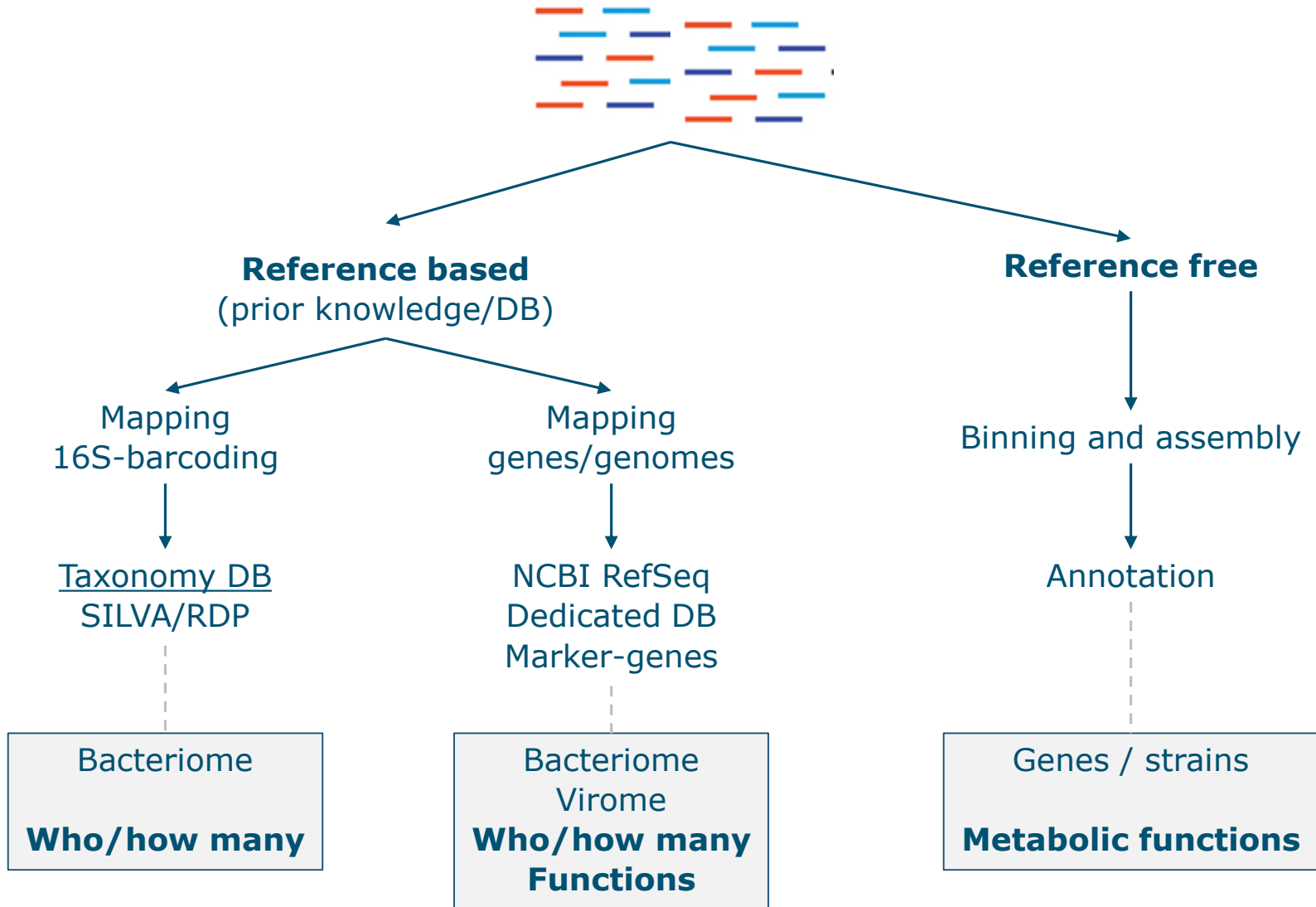
(40M DNA clusters PE150)
(12 billion bases per sample)

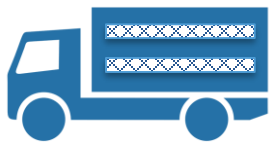
~24 billion data points per sample

Millions of DNA reads
~ 150-300 bp

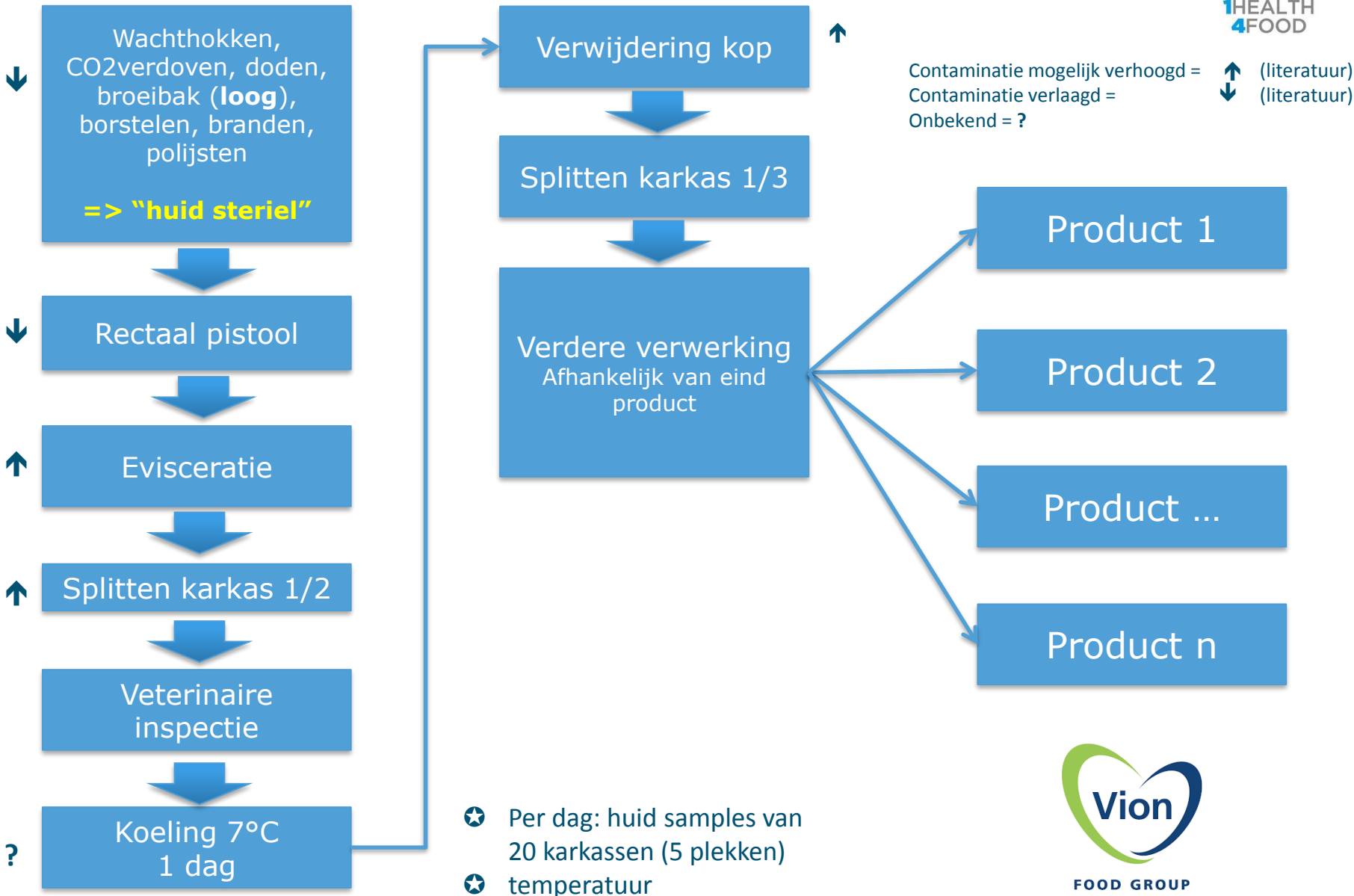


Sequences to interpretation



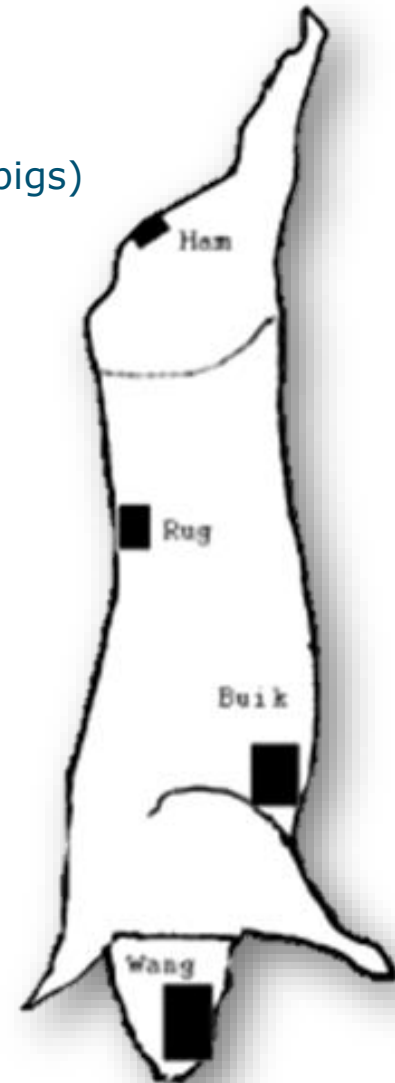


Overview meat production line Vion Boxtel



Sampling procedure for carcasses

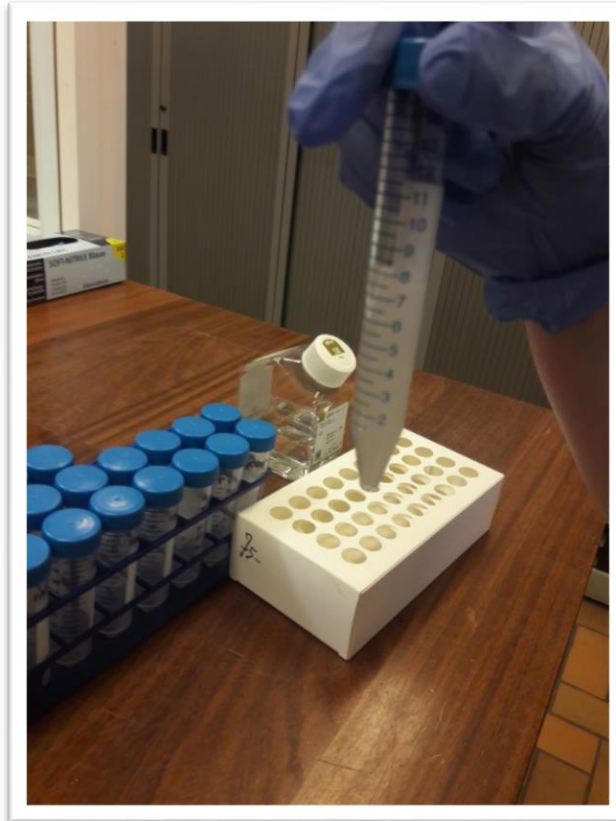
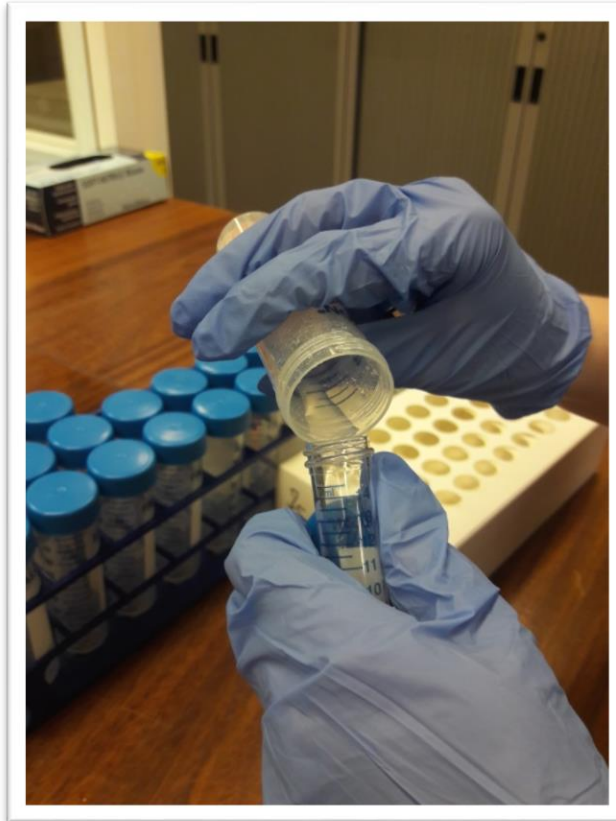
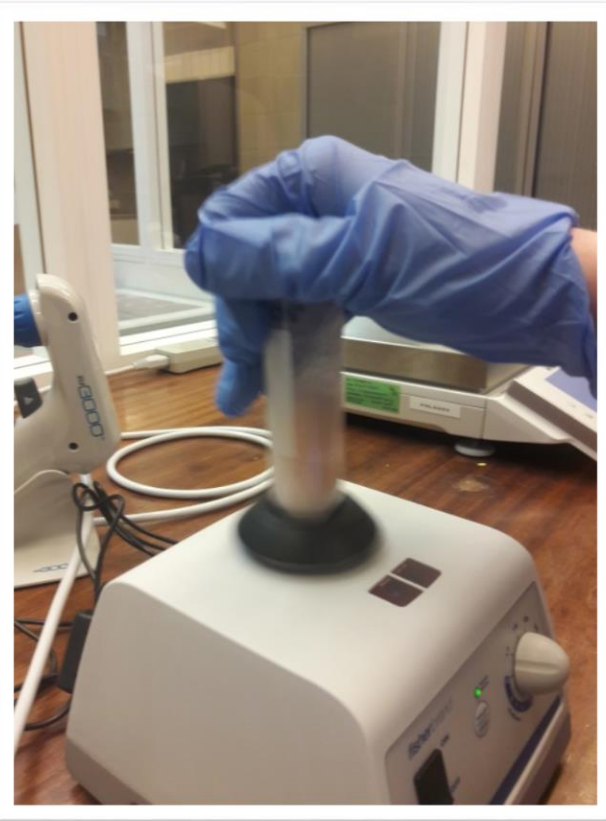
- Uniform procedure on all sites
- Trained staff
- Daily *at random* sampling ~20 carcasses in Boxtel (1 / 1,000 slaughtered pigs)
Oct 2017 – Oct 2018: ~4,100 samples collected (sampling 4 dy/week)
- Routine excision sampling:
4× 5 cm² per carcass, for TVC and enterobacteriaceae
- Five carcasses selected for Salmonella pool: additional excision samples



Skin patch sampling



Skin sample prep

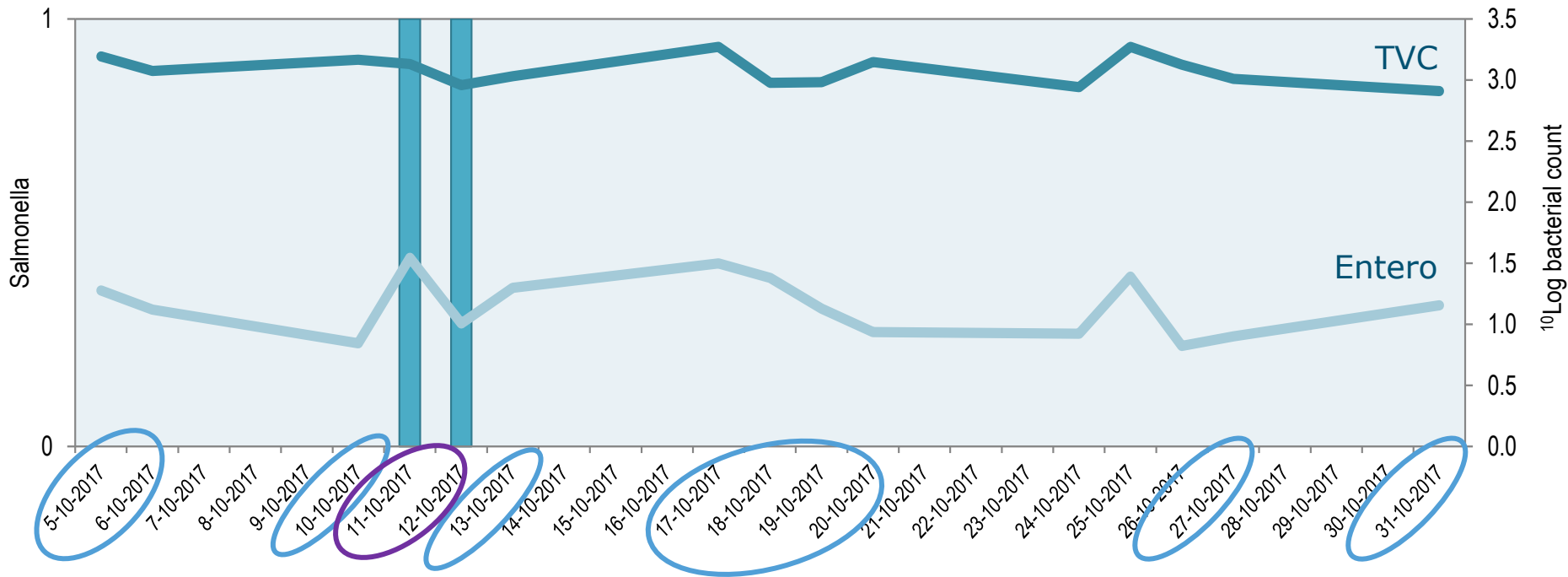


Use case Salmonella

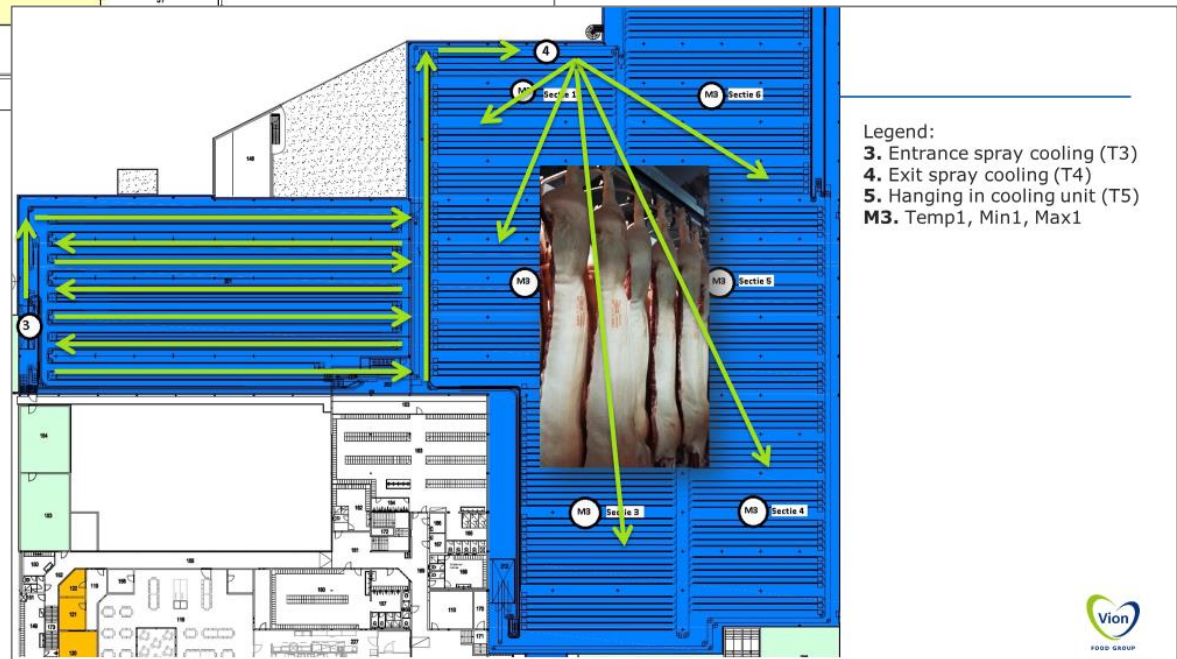
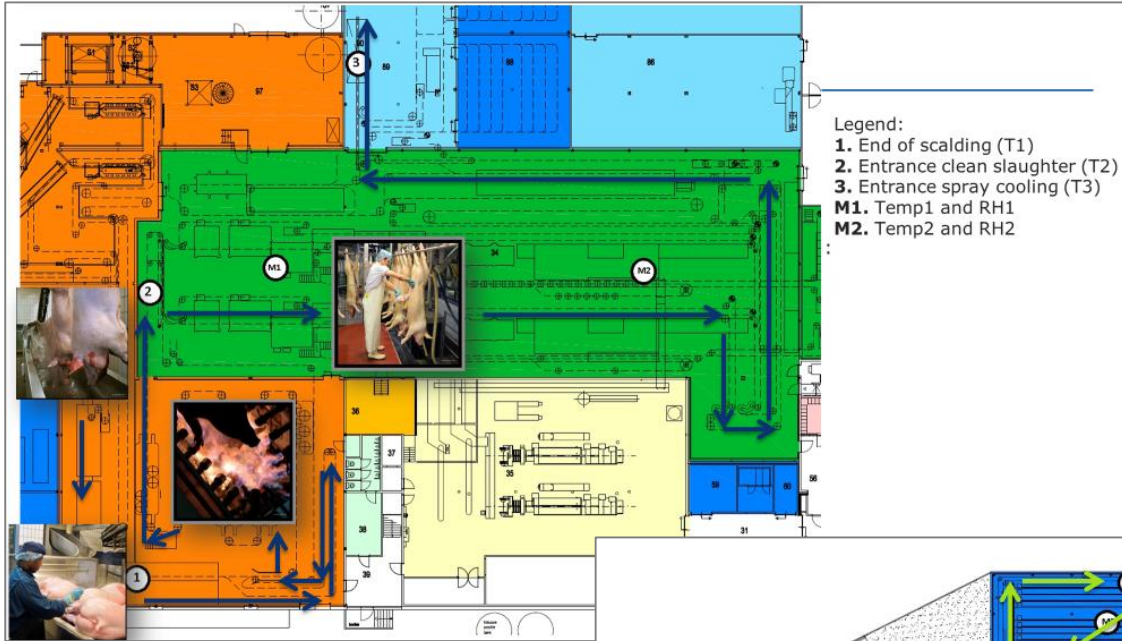
Microbiological results



- Results obtained from 15 salmonella pools (1 month)
- 2 Salmonella positive pools detected



Temperature & Relative humidity



Microbiome analysis



■ Conditions for sample storage/prep

- Sample suitable?
- DNA extraction
- Bacteriome vs full microbiome
 - 16S-barcoding vs shotgun seq (full metagenome)

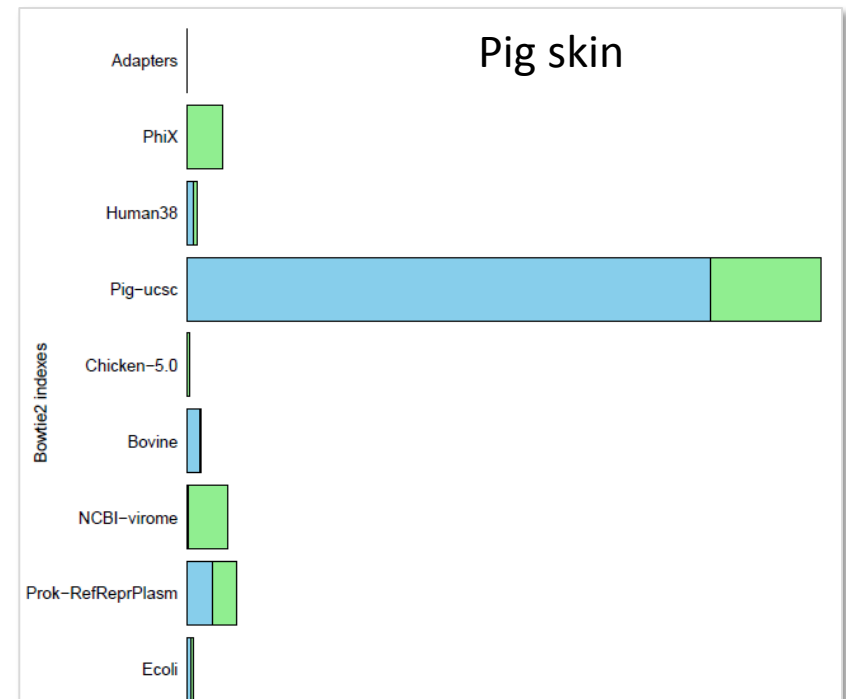
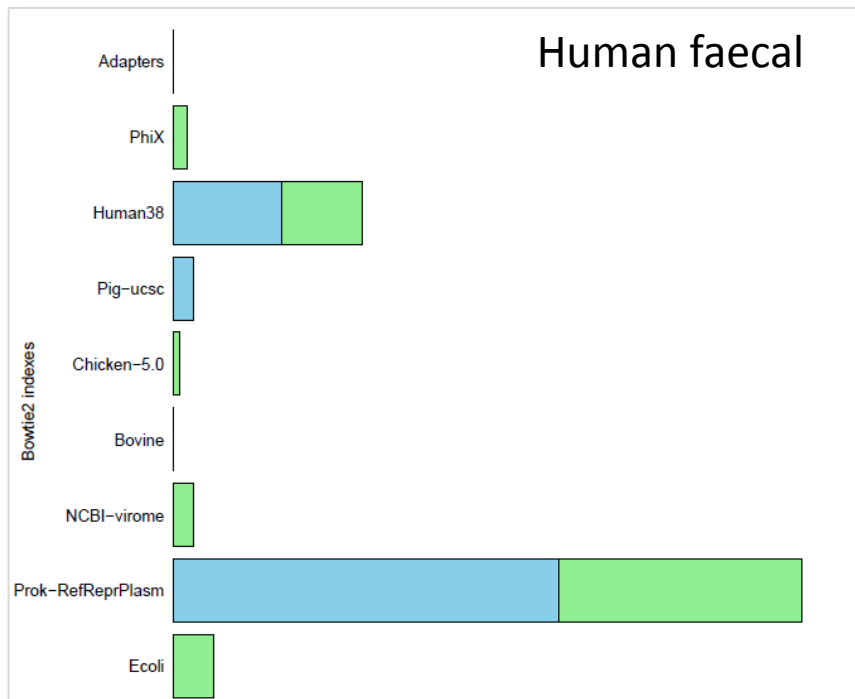
■ Future

- Dead/alive microbiome
 - 16S-barcoding sample pre-treatment
 - Meta-transcriptomics

Bigger/more = better?

Full metagenome microbiome analysis method?

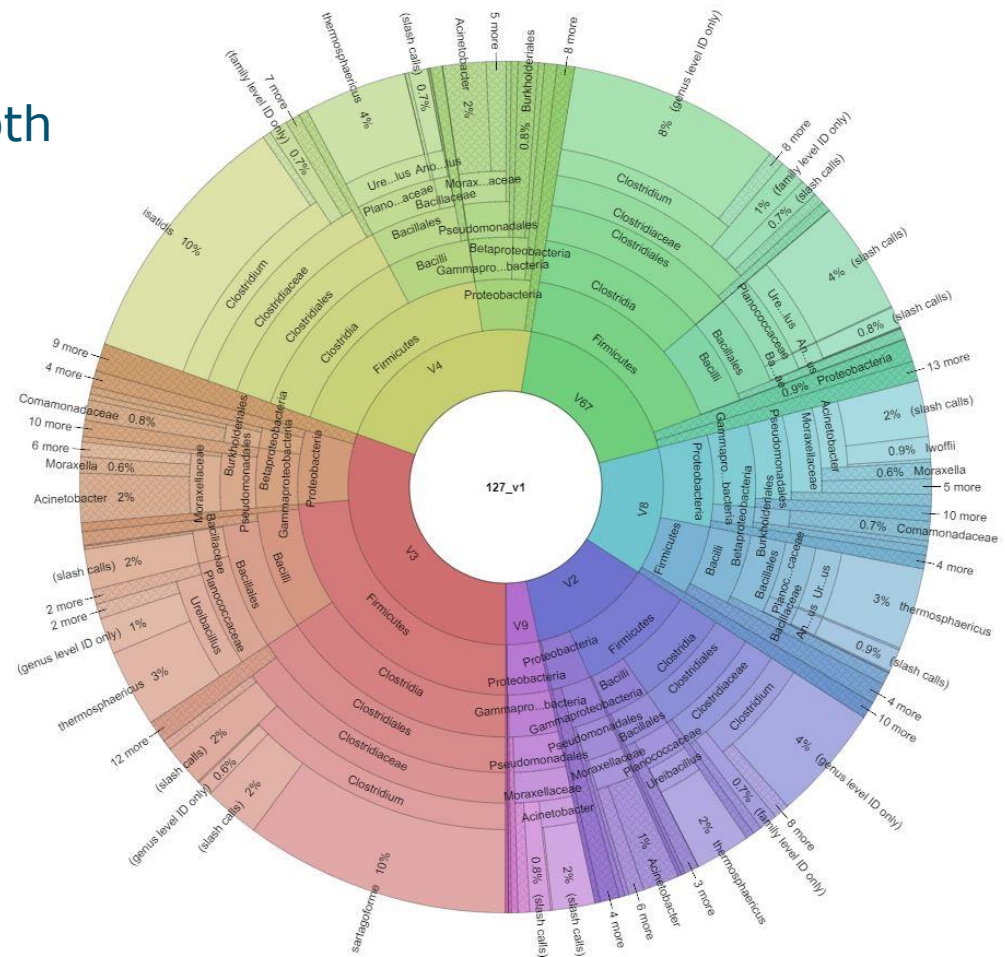
- Mapping 80M sequences against several species databases --> counts



16S-barcoding

- Ion-torrent system (~240-400 bases/read)
- Multiple 16S rRNA gene variable regions (v2, v3, v4, v6-7, v8, v9)
- Bacteriome ~ 1-2M seq depth

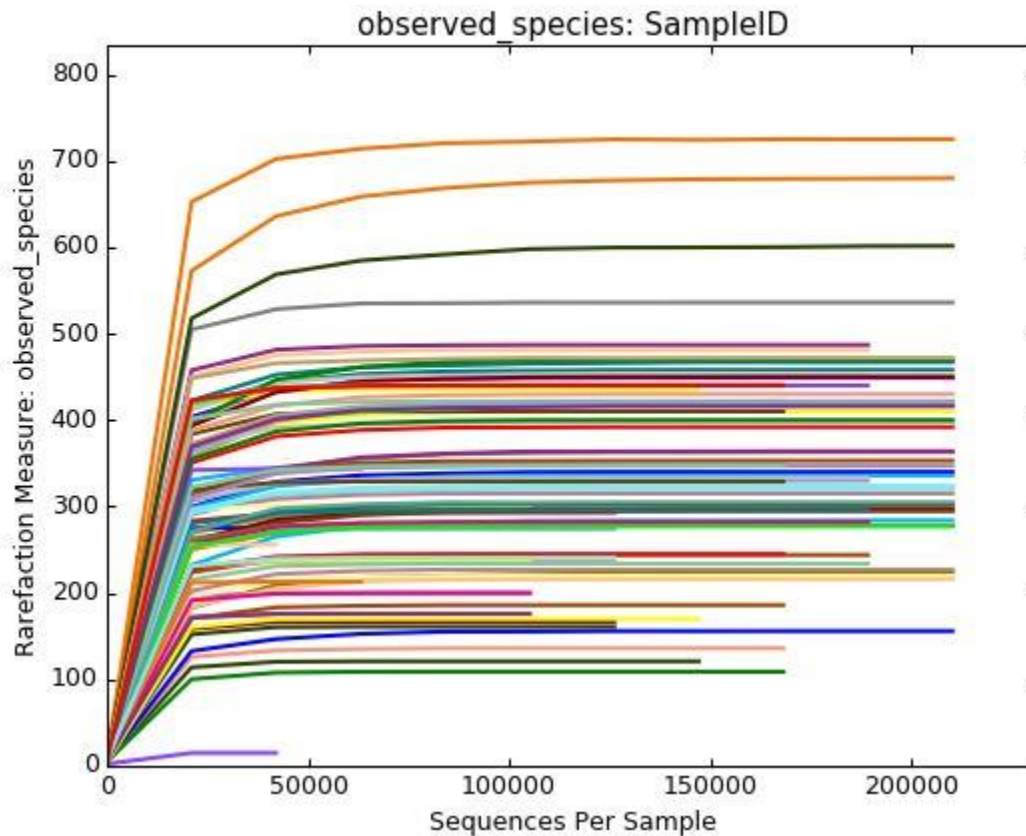
**Thermo
Fisher**
SCIENTIFIC



Detected taxonomic units

Select a Metric: Select a Category:

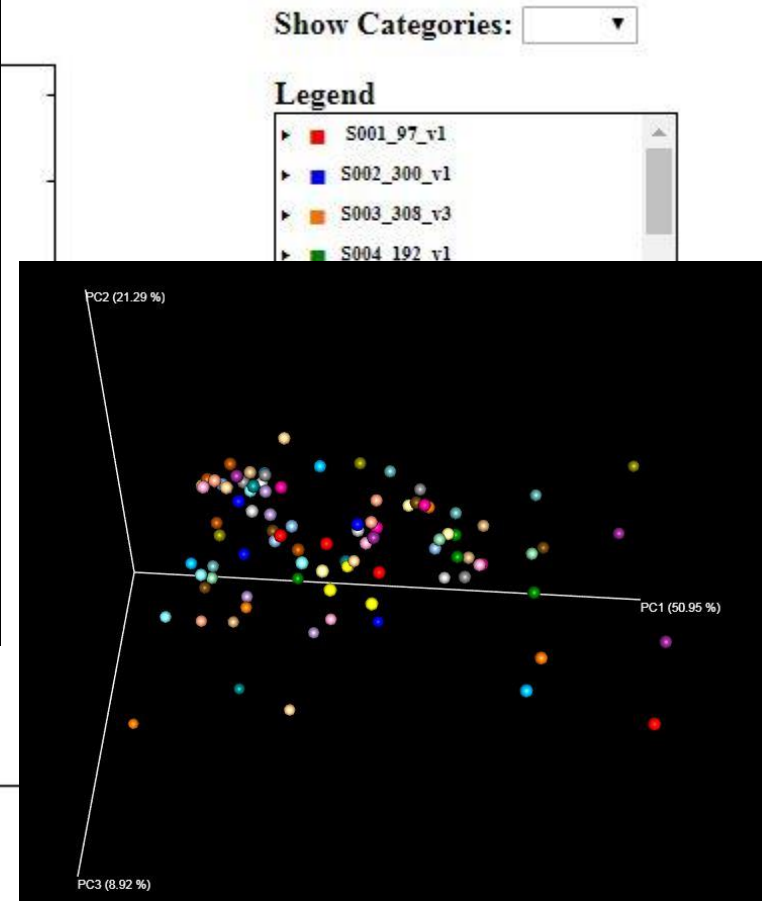
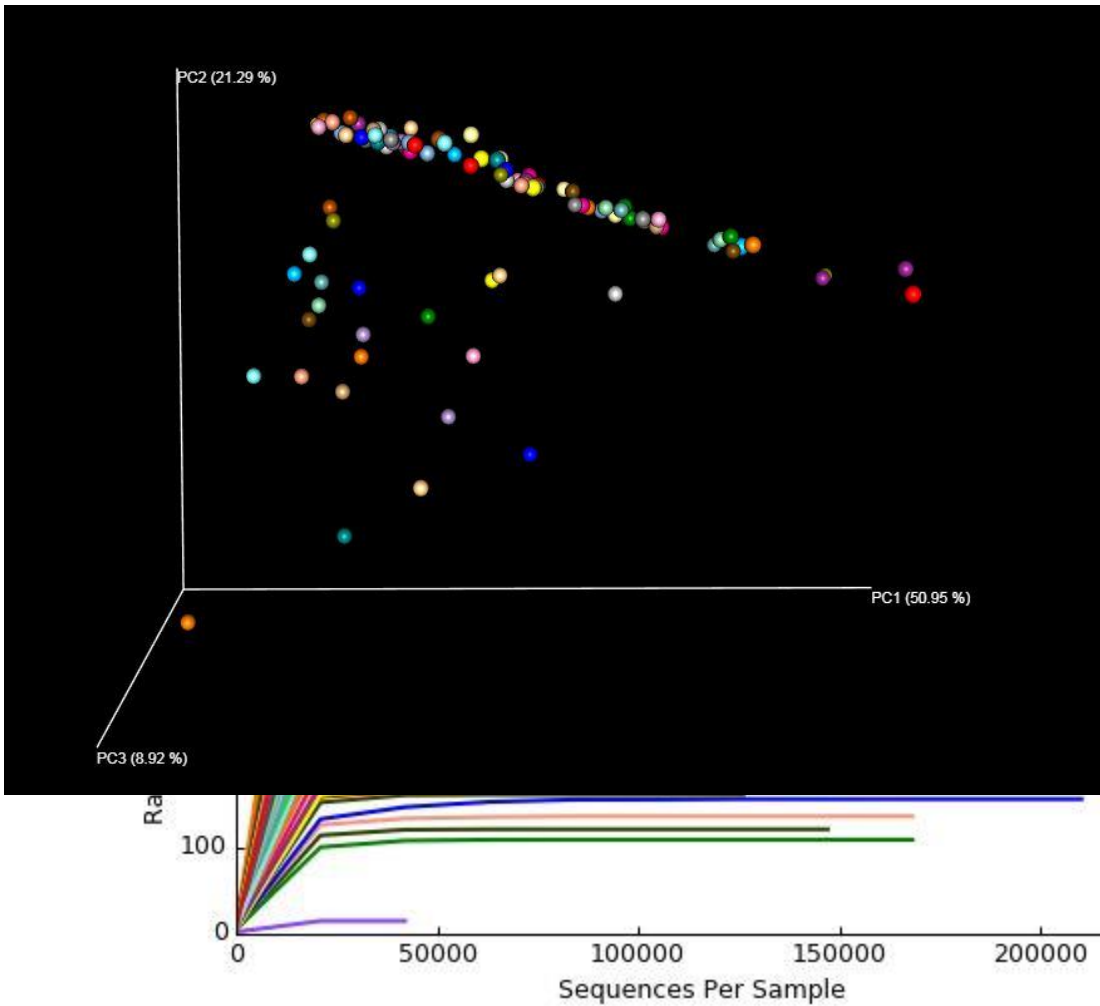
Show Categories:



Legend

- ▶ S001_97_v1
- ▶ S002_300_v1
- ▶ S003_308_v3
- ▶ S004_192_v1
- ▶ S005_111_v3
- ▶ S006_49_v1
- ▶ S007_80_v1
- ▶ S009_130_v2
- ▶ S010_173_v1
- ▶ S011_35_v1
- ▶ S012_148_v1
- ▶ S013_58_v1
- ▶ S014_196_v1
- ▶ S015_177_v1
- ▶ S016_310_v1

Detected taxonomic units



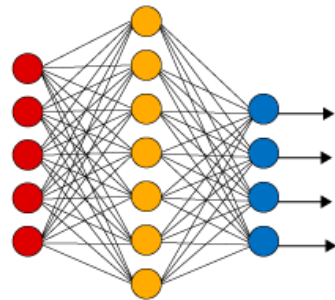
Probably deep-learning needed

IBM – Watson :: exploratory studies

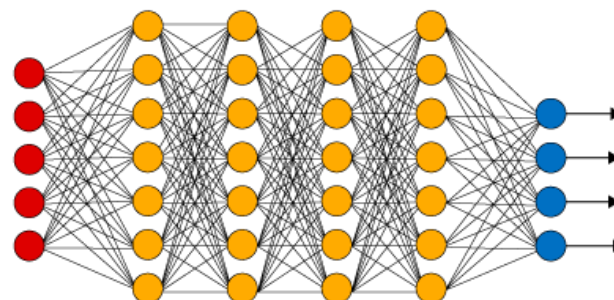


- Find non-obvious correlations --> early signatures
- Integrate other (meta) data
 - TVC, enterobacteriaceae, Salmonella count, ...
 - 16S data
 - Sensor data (temperature/humidity)

Simple Neural Network



Deep Learning Neural Network



● Input Layer

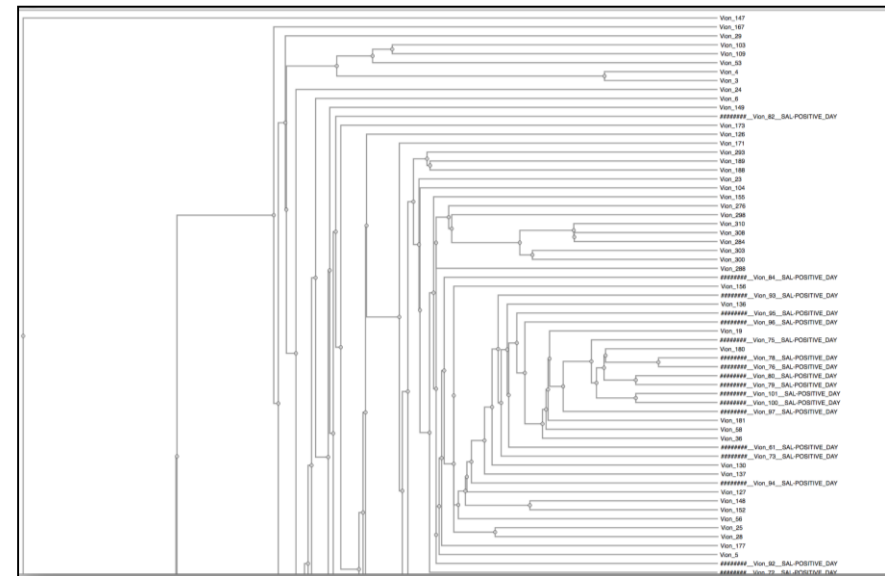
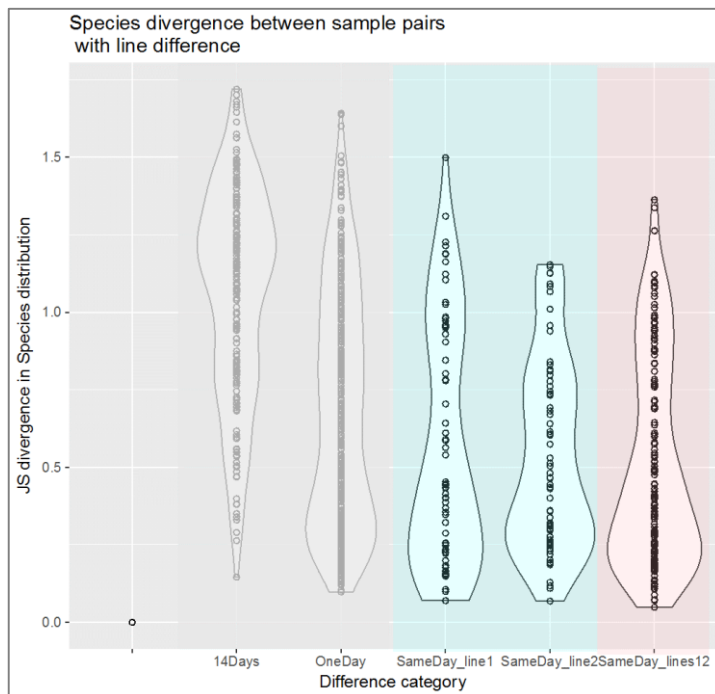
● Hidden Layer

● Output Layer

Preliminary results

On a subset of analysed samples (first month)

- In rank-based correlations a few taxa informative (@genus)
- Microbiomes of two processing lines not different
- Reference-free clustering:
No correlations yet



Future

Just started...



- Pooling strategies to increase data coverage
- More data --> complete year
 - Integrate sensor data
 - Deep analysis for early signatures
- If early signatures exist --> biomarkers
 - > fast and cheap but still sensitive
- Intervention modelling when signatures are discovered

Thank you for your attention



WBVR
Alex Bossers
Cynthia Ho
Albert de Boer

WU-FM
Marcel Zwietering
Tjakko Abee



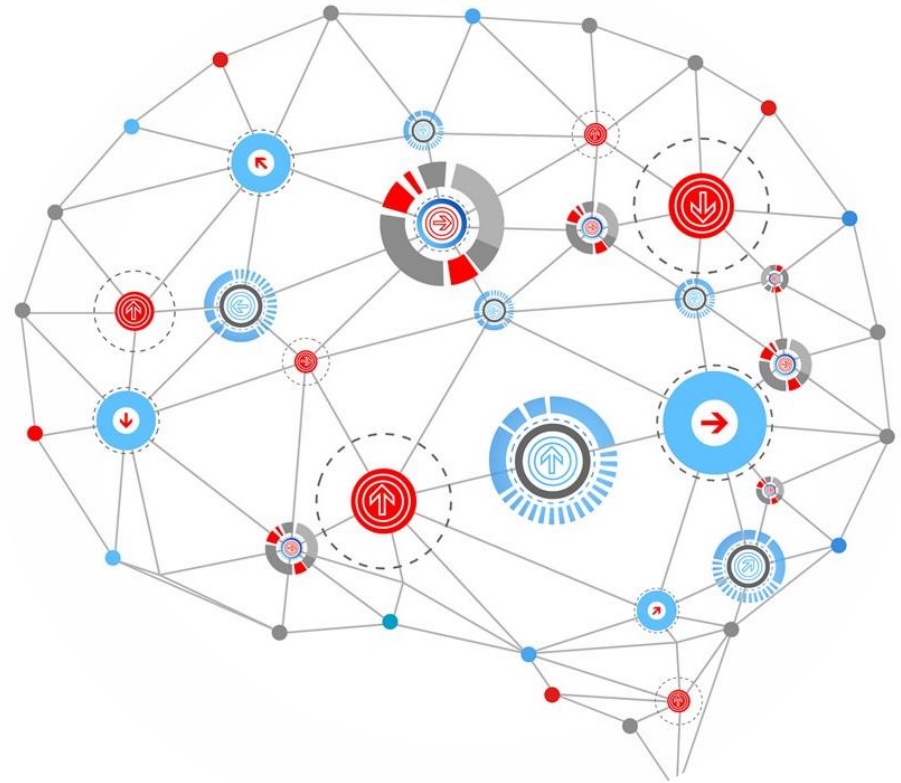
Bert Urlings
Martijn Bouwknecht



Patrik Buholzer
Livio Baselgia

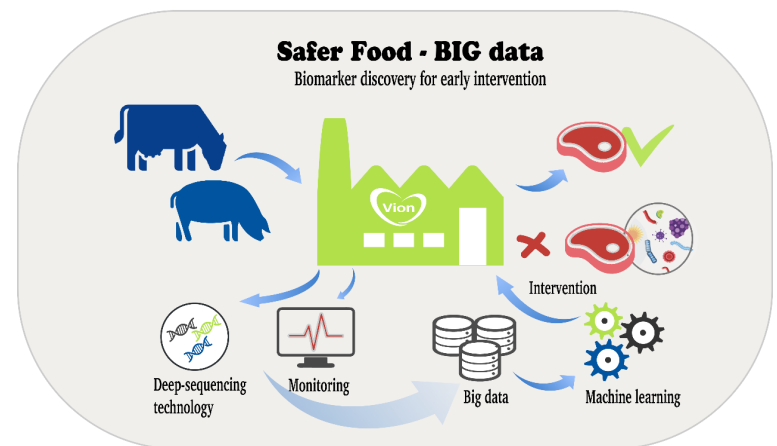


David Chambliss
Nimrod Megiddo
Ban Kawas



Stelling 1

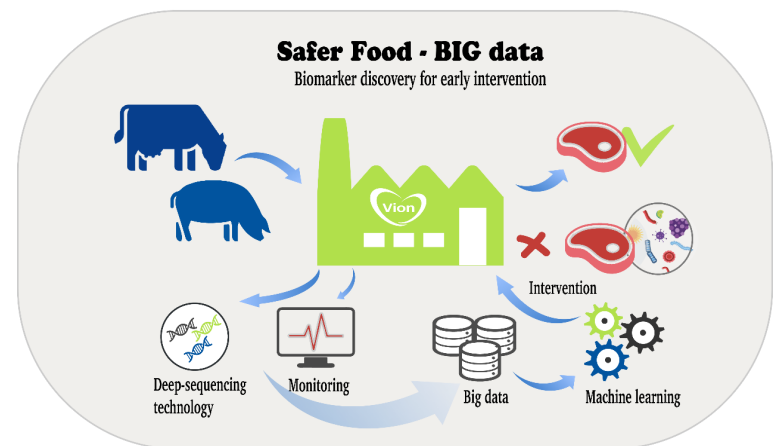
- BIG-data toepassingen zijn operationeel in het food-safety domein binnen:
 - <5 jaar
 - 5-10 jaar
 - >10 jaar
 - Nooit



Stelling 2

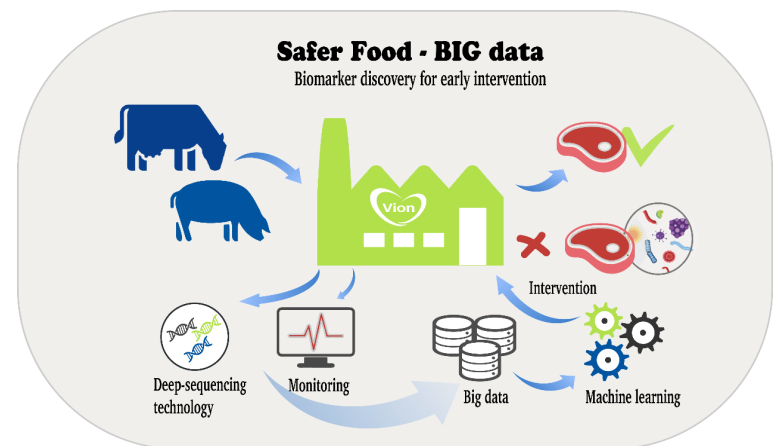
■ Klassieke microbiologie rules! *En zal altijd nodig blijven.*

- Eens
- Niet eens
- Weet niet



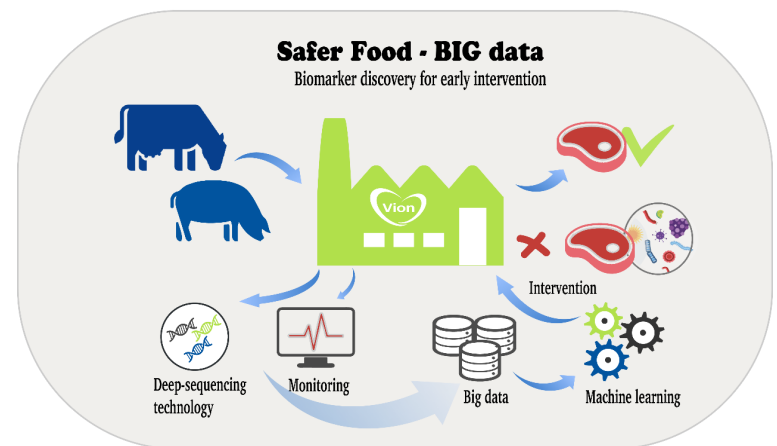
Stelling 3

- Next generation sequencing in food safety is een *pandorra's box*
 - Eens
 - Oneens
 - Geen mening



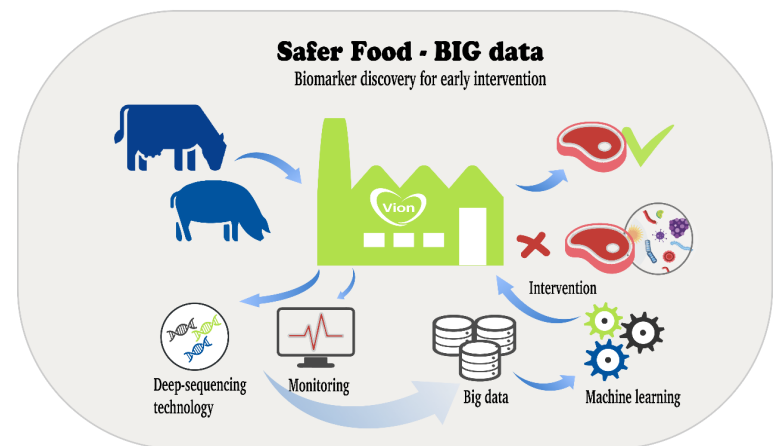
Stelling 4

- The BIGGER the data, the BETTER
 - Eens
 - Oneens
 - Geïntegreerde data is belangrijker



Stelling 5

- Correlaties via deep-learning applicaties zijn niet voldoende, causaties zijn nodig
 - Eens
 - Oneens



What else?



estNreadClade	#clade_name					
85361	k__Viruses					
112043	k__Bacteria					
658	k__Eukaryota					
85361	k__Viruses	p__Viruses_noname				
102653	k__Bacteria	p__Actinobacteria				
9839	k__Bacteria	p__Proteobacteria				
4473	k__Bacteria	p__Firmicutes				
405	k__Eukaryota	p__Apicomplexa				
9449	k__Viruses	p__Viruses_noname	c__Viruses_noname	o__Viruses_noname	f__Retroviridae	g__Gammaretrovirus
810	k__Viruses	p__Viruses_noname	c__Viruses_noname	o__Viruses_noname	f__Partitiviridae	g__Alphacryptovirus
77633	k__Bacteria	p__Actinobacteria	c__Actinobacteria	o__Actinomycetales	f__Propionibacteriaceae	g__Propionibacterium
215	k__Viruses	p__Viruses_noname	c__Viruses_noname	o__Viruses_noname	f__Potyviridae	g__Potyvirus
371	k__Viruses	p__Viruses_noname	c__Viruses_noname	o__Viruses_noname	f__Viruses_noname	g__Viruses_noname
8878	k__Bacteria	p__Proteobacteria	c__Gammaproteobacteria	o__Pseudomonadales	f__Moraxellaceae	g__Acinetobacter
5148	k__Bacteria	p__Actinobacteria	c__Actinobacteria	o__Actinomycetales	f__Propionibacteriaceae	g__Propionibacteriaceae_unclassified
3697	k__Bacteria	p__Firmicutes	c__Bacilli	o__Bacillales	f__Staphylococcaceae	g__Staphylococcus
2970	k__Bacteria	p__Proteobacteria	c__Gammaproteobacteria	o__Pseudomonadales	f__Pseudomonadaceae	g__Pseudomonas
13	k__Viruses	p__Viruses_noname	c__Viruses_noname	o__Caudovirales	f__Podoviridae	g__P22likevirus
598	k__Bacteria	p__Actinobacteria	c__Actinobacteria	o__Actinomycetales	f__Corynebacteriaceae	g__Corynebacterium
2	k__Viruses	p__Viruses_noname	c__Viruses_noname	o__Viruses_noname	f__Flaviviridae	g__Flavivirus
455	k__Bacteria	p__Firmicutes	c__Bacilli	o__Bacillales	f__Planococcaceae	g__Ureibacillus
22	k__Viruses	p__Viruses_noname	c__Viruses_noname	o__Viruses_noname	f__Baculoviridae	g__Alphabaculovirus
3	k__Viruses	p__Viruses_noname	c__Viruses_noname	o__Caudovirales	f__Siphoviridae	g__C2likevirus
723	k__Bacteria	p__Proteobacteria	c__Gammaproteobacteria	o__Enterobacteriales	f__Enterobacteriaceae	g__Salmonella
331	k__Bacteria	p__Actinobacteria	c__Actinobacteria	o__Actinomycetales	f__Micrococcaceae	g__Rothia
258	k__Bacteria	p__Firmicutes	c__Bacilli	o__Lactobacillales	f__Streptococcaceae	g__Streptococcus
293	k__Bacteria	p__Proteobacteria	c__Gammaproteobacteria	o__Pseudomonadales	f__Moraxellaceae	g__Enhydrobacter
308	k__Bacteria	p__Actinobacteria	c__Actinobacteria	o__Actinomycetales	f__Corynebacteriaceae	g__Turicella
441	k__Eukaryota	p__Apicomplexa	c__Aconoidasida	o__Haemosporida	f__Haemosporida_noname	g__Plasmodium
51	k__Bacteria	p__Proteobacteria	c__Gammaproteobacteria	o__Aeromonadales	f__Aeromonadaceae	g__Aeromonas

Biofilm: trouble ~ informer

