Laboratoire de Spectrochimie Infrarouge et Raman

# A novel statistical approach for the identification of DNA sequences characterised by single-molecule fluorescence microscopy

Raffaele Vitale, Arno Bouwens, Jochem Deen, Kris Janssen, Cyril Ruckebusch, Johan Hofkens





http://lasir.univ-lille1.fr

Chimiométrie 2018 31/01/2018

# Where everything started: the ADGut project (H2020)



These findings add Alzheimer's disease to the growing list of diseases associated with gut microbial alterations, as well as suggest that gut bacterial communities may be a target for therapeutic intervention.

Vogt, N.M. et al., Scientific Reports, 2017 (7), 13537

Determining the composition of the gut microbiome in real samples is **fundamental** to develop new diagnostic tools and maybe even come up with a way to delay the onset of this neurodegenerative disorder.



How can this be done?



## **DNA mapping by single-molecule fluorescence microscopy**



# **Species identification/matching**



Chimiométrie 2018 - 31/01/2018

### Assessing the statistical significance of the matching





### A first fast attempt: Lambda vs. T7





#### A more comprehensive test: a simulated dataset



Assignation Matrix -  $L_{eff}$  70% - Widefield -  $\alpha$  = 0.05

85.8	7.8	4.5	7.8	7.7	5.7	4.2	4.1	12.7	9.4
4.5	72	29	39.6		7.5	5.6	5.9	11	12.5
4.3	35.9	66.5	56.2	5.4	7.9		5.2	6.6	13.7
- 4	38.6		72.7	4.9	5.4	4.7	5.8	7.3	15.6
- 4	11		9.1	72.1	28.6	38.4	5.3	6.7	14.2
- 3.8	6.4	5.5		23.3	74.3	37.8	3.3		9.7
- 3.6	9.1	7.5	8.3	34.6	38.5		6.2	5.9	9.5
- 3.1	7.2	5.7	11.2					7.4	11.9
10.5	16.1	7.6	11.6	11.1	12.9		2.5	92.9	22.2
2.4	8.1	8.4	21.1	5.5	4.4	3.3	5.4	17.6	98.5
V01146	AP004402	NC_004913	AP005154	NC_019711	J02459	NC_019723	NC_019721	NC_002167	NC_002166
	Tested species								



# **Super-resolution fluorescence microscopy**



# **Reducing ambiguity by increasing resolution**



LASIR

Chimiométrie 2018 - 31/01/2018

## **Reducing ambiguity by statistical bootstrapping**



# Conclusions

- A novel permutation-based computational procedure was developed to assess the statistical reliability of the identification of DNA sequences characterised by singlemolecule fluorescence microscopy
- Such a procedure exhibited a very promising performance in a simulated and a real case-study involving viral species
- Both super-resolution and resampling techniques like bootstrapping improved the matching quality reducing the ambiguity given by multiple assignations
- The potential of the implemented methodology is currently being assessed in more complex real scenarios also involving bacterial species



Laboratoire de Spectrochimie Infrarouge et Raman

# Many thanks for your kind attention





Chimiométrie 2018 31/01/2018

http://lasir.univ-lille1.fr