A Novel Microarray Design Strategy for Studying the Gingival Flora of Noma Patients

Antoine Huyghe[1], Yvan Charbonnier[1], Patrice Francois[1], Manuela Bento[1], Ignacio Bolivar[2], Denise Baratti-Mayer[3], Jacques Schrenzel[1]

1 Genomic Research Laboratory - University Hospital of Geneva, Switzerland.
2 Department of Animal Biology - University of Geneva, Switzerland.
3 Geneva Study Group on Noma (GESNOMA), Faculty of Medicine, University of Geneva, Geneva, Switzerland.

Mail: Antoine.Huyghe@genomic.ch
Web: www.genomic.ch

Overview

Overview **Noma is a gangrenous disease of unknown etiology** leading to the destruction of hard and soft tissues of the face. This disorder mainly affects young children of poor or less developed countries weakened by malnutrition and poor hygiene. Although the exact prevalence is unknown, recent studies have shown that the disease has a global yearly incidence of 140'000 cases and a mortality rate of approximately 90%.

Objective The exact etiology of the noma is not well established but the disease has been associated with the presence of *Fusobacterium necrophorum* and *Prevotella intermedia* in the oral flora. In order to further study the microbial community of noma samples we developed an **original microarray based on a phylogenic approach.** Microarray design was performed by selecting 16S rDNA probes specific to nodes matching each level of the phylogenetic tree (domain, phylum, class, order, family, genus, specie). This approach is based on sequence conservation. While providing information on the taxonomic composition of microbial communities, this approach should prove useful for the **detection of uncharacterized species**.

Results Preliminary results show that this **hierarchical approach is able to reveal minor changes in flora composition between two samples**. Large series of microarray experiments are currently performed to monitor gingival contents of bacteria in noma patients compared to matched healthy controls from the same geographical origin.

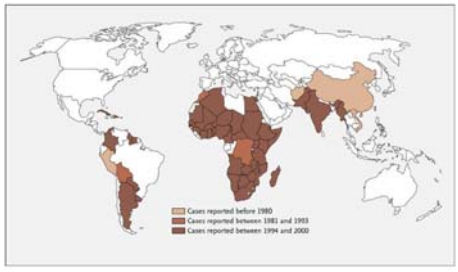


Figure 1. Noma cases reported worldwide before 1980 and until 2000 according to WHO.

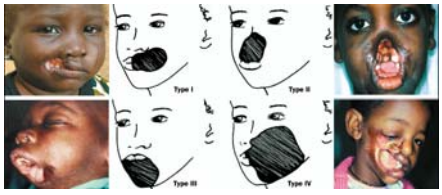


Figure 2. Classification of noma sequelae with corresponding clinical cases (From Baratti-Mayer D et al, *Lancet Infect Dis. 2003 Jul;3(7):419-31*).

Methods

Microarray Design

Design is based on 16S rDNA sequences deposited on the Ribosomal Database Project (RDP; release 9.34: ~194'000 sequences) which provides aligned sequences classified according to Bergey's taxonomic classification.

Step 1: Identification of all possible 25-nt long oligonucleotide probe sequences fitting within a range of homogeneous Tm.

Step 2: Each candidate probe was assigned to a hierarchical node (phylogenic level) representing all target sequences, as depicted below.

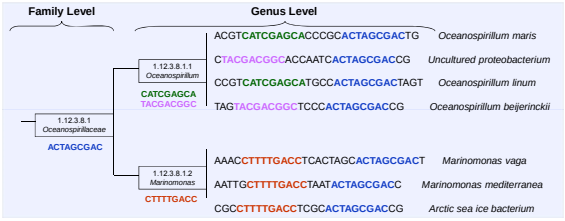


Figure 3. Simplified representation of Node assignment process.

Step 3: Candidate probes maximizing coverage were selected first, while probes providing a limited gain in coverage were discarded.

These 3 steps yielded to a set of about 9'900 oligonucleotide probes with specificities ranging from the species to the domain level. Furthermore, probe set is specific to 152'000 sequences (78%) deposited in the RDP.

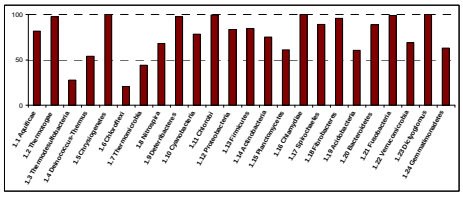


Figure 4. % coverage obtained by our strategy for each Phylum.

In vitro transcription and labelling

RNA samples were poly(A) tailed, converted to cDNA, amplified and labelled using in vitro transcription; then hybridized on the arrays.

2 samples were hybridized in duplicate:

Sample A
Sample of Healthy human oral flora.

Sample B
Sample A + *Fusobacterium necrogenes (4% total quantity)* .

Analysis

- Use of ANOVA (Analysis Of Variance) statistical test to discriminate bacteria between duplicated samples.
- Use of the "Volcano Plot" representation to interpret the ANOVA (Scatter plot of the relative expression values against the *p* value for each probe).

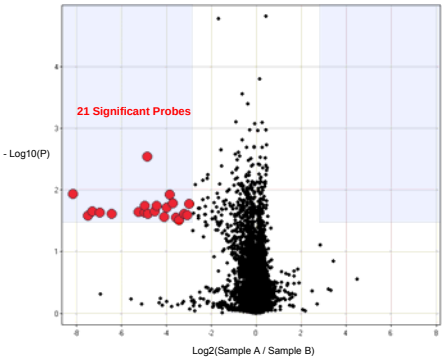
Results

- Fluorescence of probes mapping the *Fusobacteriales* Order to the *Fusobacterium* Genus are statistically more predominant in the spiked sample.

Probe Name	Node	Description	Fold Change	P Value
N02784	1.21.1.1	Fusobacteriales(order)	283	0.012
N04174	1.21.1.1.1	Fusobacterium(genus)	177	0.026
N02783	1.21.1.1	Fusobacteriales(order)	157	0.022
N04172	1.21.1.1.1.1	Fusobacterium(genus)	124	0.023
N04173	1.21.1.1.1.1	Fusobacterium(genus)	86	0.025
N04176	1.21.1.1.1.1	Fusobacterium(genus)	38	0.023
N05977	1.21.1.1.1	Fusobacteriaceae(family)	31	0.024
N02313	1.13.2.3.1.1	Acholeplasma(genus)	31	0.018
N02785	1.21.1.1	Fusobacteriales(order)	28	0.003
N00401	1.12.2.8.1.6	Dechloromonas(genus)	28	0.024
N05979	1.21.1.1.1	Fusobacteriaceae(family)	23	0.022
N06825	1.13.1.1.2.4	Anaerostipes(genus)	21	0.018
N05988	1	Bacteria(domain)	17	0.028
N04985	1.13.1.1.1.11	Faecalibacterium(genus)	16	0.019
N04180	1.23.1.1.2.1	Victivallis(genus)	14	0.012
N02723	1.12.4.6.1.4	Desulfovirga(genus)	13	0.017
N02036	1.21.1.1.1.2	Ilyobacter(genus)	12	0.028
N05989	1	Bacteria(domain)	11	0.031
N01684	1.12.4.2.3.3	Desulfonatronovibrio(genus)	9	0.025
N05978	1.21.1.1.1	Fusobacteriaceae(family)	8	0.026
N05983	1.21.1.1.1	Fusobacteriaceae(family)	8	0.017

Table 1. 21statistically significant probes arising from ANOVA analysis of microarray data.

Figure 5. Volcano Plot.
Statistical significance by ANOVA [-Log(P)] vs ratio [mean Sample A / mean Sample B]
P ≤ 0.03 ; Fold Change ≥ 8



- Our approach appears useful to characterize species present in the healthy buccal flora and supports published data.

Probe #	Description	Level
1	Abiotrophia	genus
2	Streptococcus	genus
3	Rubrobacter	genus
4	Streptococcaceae	family
5	Firmicutes	phylum
6	Bacteria	domain
7	unclassified	-
8	unclassified	-
9	Cellulosimicrobium	genus
10	Paenibacillus	genus
11	Peptococcaceae	family
12	unclassified	-
13	Lactobacillales	order
14	unclassified	-
15	Lactobacillus	genus
16	unclassified	-
17	Desemzia	genus
18	unclassified	-
19	Desulfobacula	genus
20	Streptococcus	genus

Table 2. Top 20 fluorescent probes, representing the most predominant families found in a healthy human gingival sample.

Conclusions

- Our oligoarray is composed of ~9'900 probes matching 78% of the total sequences (~194'000) deposited in the RDP.
- The main advantage of our strategy is the gain of time compared to the 'gold standard' 16S rDNA cloning and sequencing.
- Results show that our approach is able to detect significant differences between spiked samples.
- Large series of microarray experiments are currently performed to monitor gingival contents of bacteria in noma patients compared to matched healthy controls from the region of Zinder, Niger.
- This design should prove useful for characterizing species present in complex microbial mixtures, such as body fluids and environmental samples, i.e. applicable beyond this research project.
- The limitation of our approach is that it cannot measure unknown species sharing poor relatedness with known sequences.

Acknowledgements

This project is supported by the Geneva Study Group on Noma (GESNOMA).

www.gesnoma.org