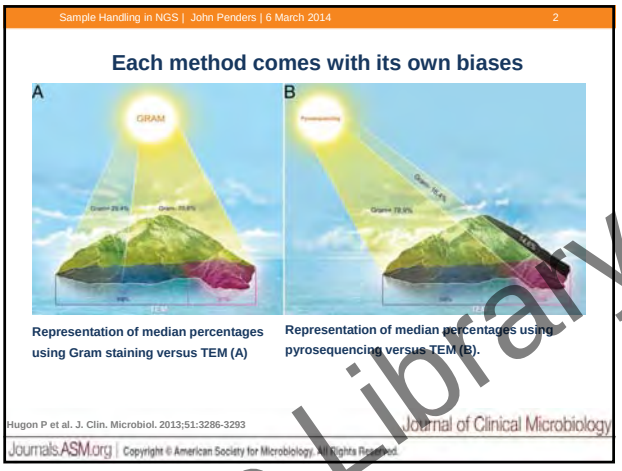
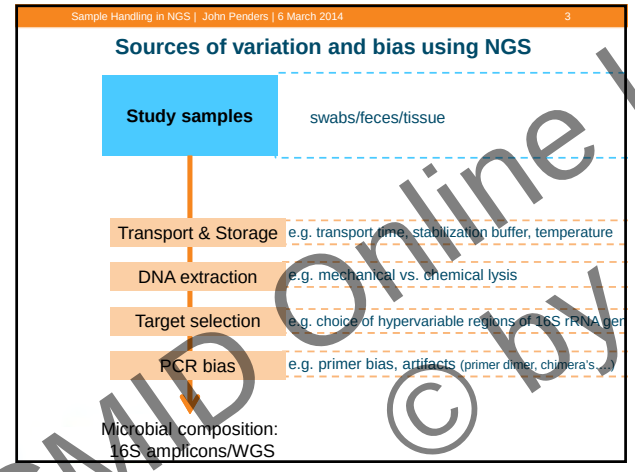
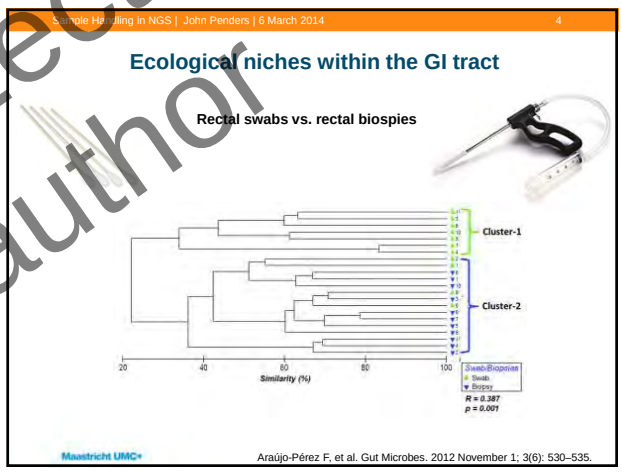
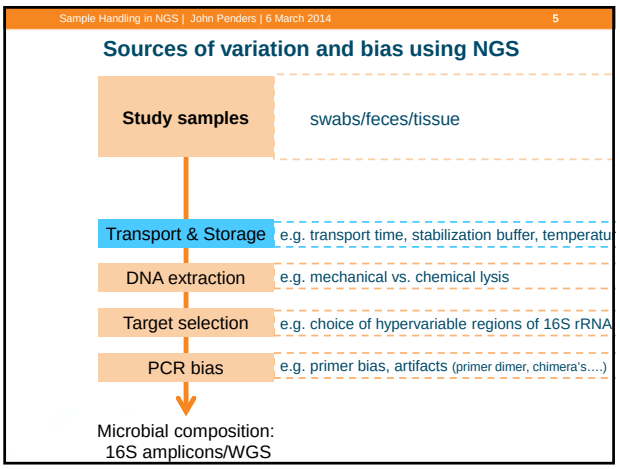
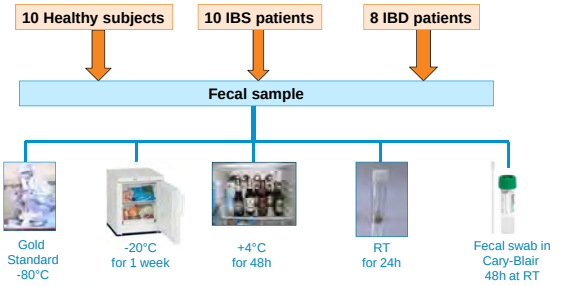








Study design



Quality parameters

- High total number of phylogenetic units detected
- High species richness and diversity estimates *(alpha diversity)*

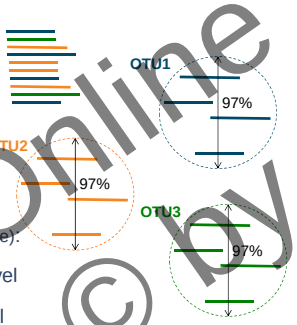
  Detection (high relative proportion) of extremely oxygen sensitive species
- Low intra-sample *(low beta-diversity)* variation as compared to -80 aliquot

Operational Taxonomic Unit (OTU) approach

- **Multiple sequence alignment**
- **Distance/similarity matrix**
- **Clustering into OTU's**

(de novo or against reference database):

- 97% similarity ~ species level
- 95% similarity ~ genus level

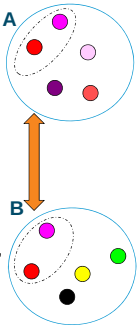


How do we describe and compare diversity?

- **α Diversity**:
  - "How many taxa/OTUs are in a sample?"

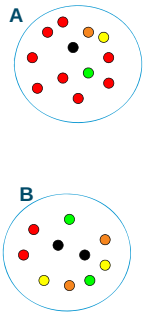
    (e.g. 5 different colors/taxa in sample A and 5 in sample B)
  - Are samples kept at RT less diverse than samples immediately stored at -80°C?
- **β Diversity:**
  - "How many taxa/OTUs are shared between samples?"

    (e.g. 2 shared colors/taxa between samples A and B)
  - Does the microbiota differ with different disease states/sampling method?

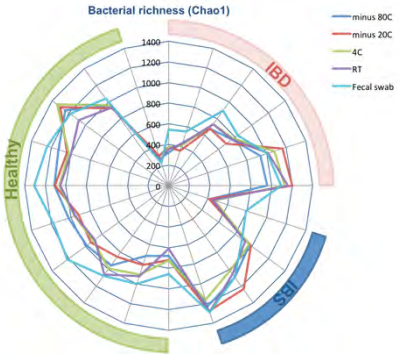


Quantitative versus Qualitative measures

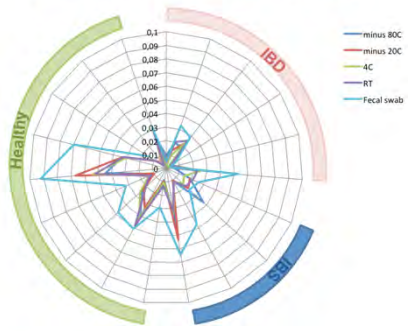
- **Qualitative alpha- & beta-diversity indices consider presence absence only**
- **Quantitative alpha- & beta-diversity indices also considers relative abundance**



Effects of sampling & storage on microbial composition



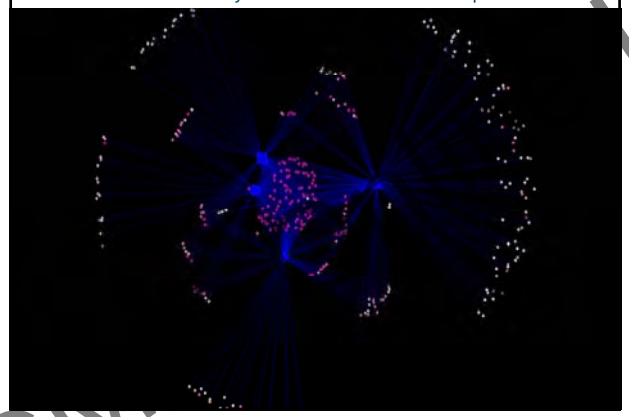
Relative abundance of ruminococcus species



Hierarchical clustering shows clustering accoring to subjects and not to storage method

Clustering according to weighted Unifrac distances by means of UPGMA method

Network-analysis: shared OTUs within sample



Sources of variation and bias using NGS

Study samples	swabs/feces/tissue
Transport & Storage	e.g. transport time, stabilization buffer, temperatu
DNA extraction	e.g. mechanical vs. chemical lysis
Target selection	e.g. choice of hypervariable regions of 16S rRNA
PCR bias	e.g. primer bias, artifacts (primer dimer, chimera's….)

Microbial composition:
16S amplicons/WGS

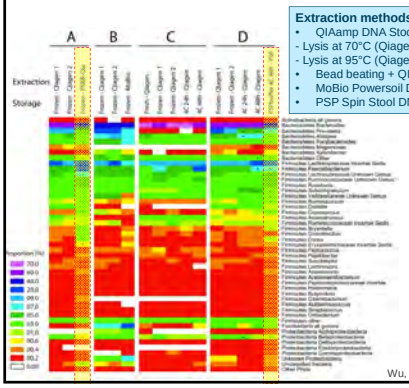
DNA extraction

- Efficient disruption of bacterial cell walls (lysis) is crucial
- Cell lysis can be achieved by:
  - mechanical disruption (e.g. bead beating, sonication)
  - enzymatic (lysozyme, lysostaphin) or chemical disintegration of cell walls
- Lysis of some bacteria is more difficult than others
  - o Microbes most resistant to lysis (e.g. gram-positive cocci and methanogenic archaea) require harsh physical and chemical treatments
  - o Microbes which are easy to disrupt (e.g. gram-negative Bacteroidetes) may tend to yield denatured DNA under drastic treatment conditions
- Efficiency of lysis protocol differs considerably according to type of sample (e.g. removal of inhibitors)

Ideally, lysis of G+ and G- microbes with equal efficiency, but with limited shearing of bacterial DNA

DNA extraction

Impact of DNA extraction methods on recovery of bacterial taxa



Extraction methods
- QIAamp DNA Stool Minikit:
  - Lysis at 70°C (Qiagen 1)
  - Lysis at 95°C (Qiagen 2)
- Bead beating + QIAamp DNA Stool Minikit (PhBB-Qia)
- MoBio Powersoil DNA Isolation Kit (MoBio)
- PSP Spin Stool DNA Plus Kit (PSPbuffer)

Wu, G.D. et al. BMC Microbiol. 2010; 10: 206.

DNA extraction

- DNA isolation method has strong effect on microbial composition
- Bead beating combined with chemical lysis led to recovery of a greater proportion of *Firmicutes*
- In cases where it is important to know the exact proportions of *Firmicutes*, it may be best to use the bead beating combined with chemical lysis
- Other studies also suggest using a protocol combining bead beating followed by enzymatic/chemical lysis of bacterial cells

International Human Microbiome Standards

The overall concept of IHMS is to promote development and implementation of standard procedures and protocols in three separate but related fields:

Collecting and processing of human samples

- Review of current protocols of >20 groups involved in human intestinal metagenomics, across 4 continents (Asia, Australia, Europe, North America)
- Experimentally test these protocols. A single stool sample was aliquoted and aliquots were sent to the participating laboratories for DNA extraction. In parallel, a mock bacterial community was established and aliquots sent for the same purpose.

http://www.microbiome-standards.org

Sources of variation and bias using NGS

Study samples	swabs/feces/tissue
Transport & Storage	e.g. transport time, stabilization buffer, temperatu
DNA extraction	e.g. mechanical vs. chemical lysis
Target selection	e.g. choice of hypervariable regions of 16S rRNA
PCR bias	e.g. primer bias, artifacts (primer dimer, chimera's…)

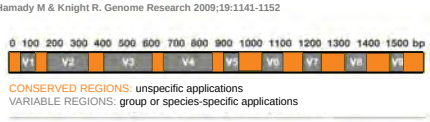
Microbial composition: 16S amplicons/WGS

Target selection

Impact of 16S rRNA gene target region on taxonomic classification

Hamady M & Knight R. Genome Research 2009;19:1141-1152

- Combination of variable and moderately conserved regions optimal for performing analyses at different phylogenetic depths
- Amplicons containing V2 or V4 give lowest error rates when assigning taxonomy (Liu Z, et al. Nucleic Acid Res 2007;35:e120)



CONSERVED REGIONS: unspecific applications
VARIABLE REGIONS: group or species-specific applications

Sources of variation and bias using NGS

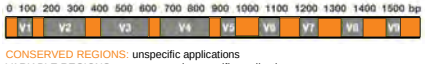
Study samples	swabs/feces/tissue
Transport & Storage	e.g. transport time, stabilization buffer, temperatu
DNA extraction	e.g. mechanical vs. chemical lysis
Target selection	e.g. choice of hypervariable regions of 16S rRNA
PCR bias	e.g. primer bias, artifacts (primer dimer, chimera's…)

Microbial composition: 16S amplicons/WGS

PCR bias (1): primer design

Poor choice of primers can lead to radically different biological conclusions

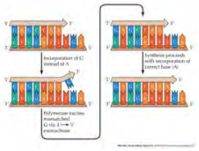
Primer	% match of sequences in RDP from all major gut bacterial phyla
8F	>95%
337F	>95%
338R	>95%
515F	>95%
784F	Biased against Verrucomicrobia
915F	>95%
930F	>95%
967F	<5% of Bacteroidetes phylum
1046R	>95%
1061R	>95%
1492R	61% of Actinobacteria, 54% Proteobacteria, <50% of other phyla



CONSERVED REGIONS: unspecific applications
VARIABLE REGIONS: group or species-specific applications

PCR bias (2): artifacts

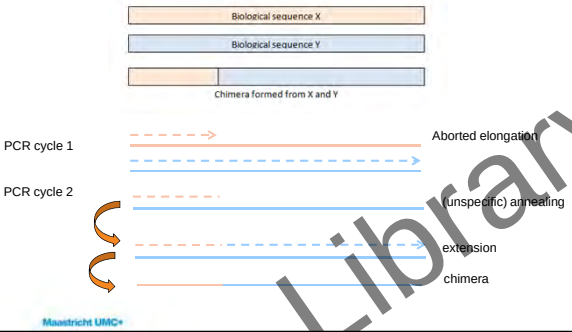
- Point mutations introduced during PCR
  - More errors in longer fragments
  - More errors with increasing number of cycles
  - DNA polymerases with 3' – 5' proofreading function have very low error rates



- Chimeric PCR products
  - Results from incomplete synthesis of rRNA gene fragments

PCR bias (2): artifacts

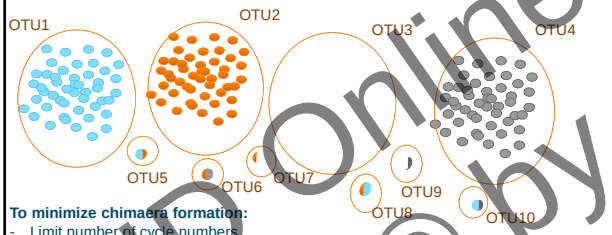
Chimaeras in PCR amplification



PCR bias (2): artifacts

Few chimaeras may greatly affect data

Only ~ 1% of amplicons is estimated to be chimeric,
however impact on microbial diversity (number of OTU's) is much larger



To minimize chimaera formation:
- Limit number of cycle numbers
- Allow for sufficient elongation time during each cycle

Remaining chimeric sequences can be identified and removed with specific software (e.g. ChimeraSlayer)

Sample Handling in NGS

- Carefully design PCR protocol for creating amplicon libraries
  - Make use of high fidelity polymerase
  - Minimize number of PCR cycles, maximize elongation time
  - Choose primers and 16S hypervariable regions providing optimal resolution
- Be aware of variations introduced by choice of sample collection & handling protocols
- Never ever let your samples thaw prior to DNA isolation!
- Optimally, pre-test different protocols before starting a study
- Be cautious when comparing data from studies using different protocols
  - especially when looking at relative abundance of different bacterial taxa

And when your sequencing data are available:
quality filtering, quality filtering, quality filtering……..

Thank you!

Chronic GI Diseases Research Group
Dept. Medical Microbiology, MUMC+

Prof. Paul Savelkoul
Danyta Tedjo
Mayk Lucchesi
Christel Driessen

Div. Gastroenterology & Hepatology
Internal Medicine, MUMC+
Prof. Ad Masclee
Daisy Jonkers
Marieke Pierik