Utilisation de la métagénomique 16S pour la surveillance de l'émergence de zoonoses bactériennes dans les populations animales

Réunion Rongeur 2014 – CBGP Maxime Galan

Réunion Rongeur du CBGP, Montferrier-sur-Lez, 15-16 septembre 2014

Métagénomique 16S: Pourquoi?

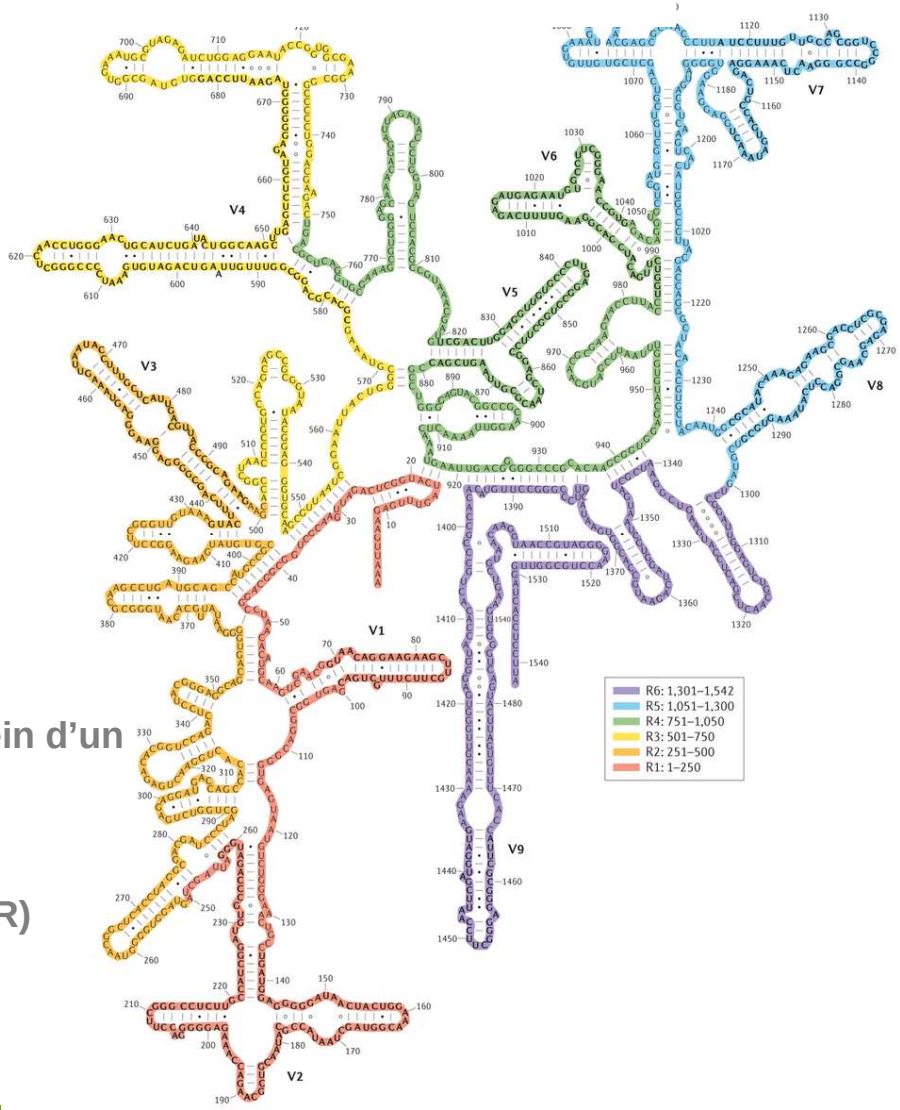
- Identification génétique de bactéries en mélange

1g de sol = 10^3 à 10^6 espèces!!

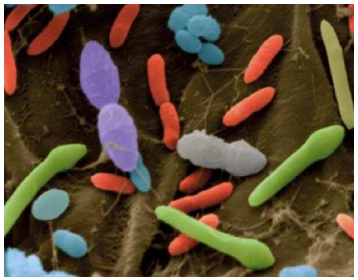
Métagénomique 16S: Principe

- Identification génétique de bactéries en mélange

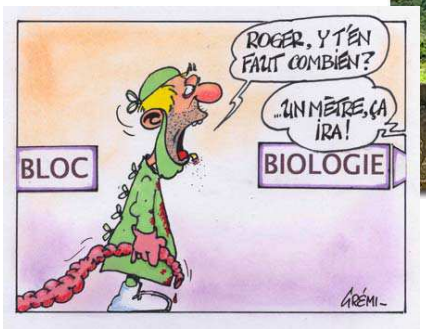
- Marqueur génétique universel: 16S rRNA
- Bases de données exhaustives et fiables
- Région variable entre taxons / conservée au sein d'un taxon
- Régions flanquantes conservées (amorces PCR)
- Longueur « séquençable » (régions V4 ou V6)



Métagénomique 16S: Applications

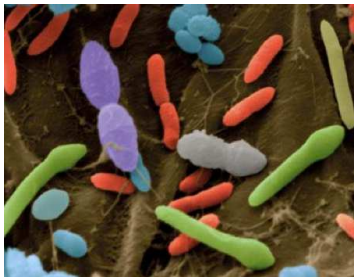


- Diversité microbienne d'échantillons environnementaux complexes :
  - Flore intestinale
  - Sol
  - Eau
  - Air …



- Epidémiologie :
  - Criblage haut débit et sans a priori de bactéries zoonotiques

Métagénomique 16S: Intérêts



- Epidémiologie :
  - Criblage haut débit et sans *a priori* de bactéries potentiellement zoonotiques:
    - Rongeurs: réservoirs majeurs d'agents zoonotiques
    - Tiques: vecteurs de maladies zoonotiques

Métagénomique 16S: Matériel & Méthodes

AEM Journals.ASM.org

Development of a Dual-Index Sequencing Strategy and Curation Pipeline for Analyzing Amplicon Sequence Data on the MiSeq Illumina Sequencing Platform

James J. Kozich,[a] Sarah L. Westcott,[a] Nielson T. Baxter,[a] Sarah K. Highlander,[b] Patrick D. Schloss[a]

APPLIED AND ENVIRONMENTAL MICROBIOLOGY, Dec. 2009, p. 7537–7541 Vol. 75, No. 23
0099-2240/09/$12.00 doi:10.1128/AEM.01541-09

Introducing mothur: Open-Source, Platform-Independent, Community-Supported Software for Describing and Comparing Microbial Communities[▽]

Patrick D. Schloss,[1,2]* Sarah L. Westcott,[1,2] Thomas Ryabin,[1] Justine R. Hall,[3] Martin Hartmann,[4] Emily B. Hollister,[5] Ryan A. Lesniewski,[6] Brian B. Oakley,[7] Donovan H. Parks,[8] Courtney J. Robinson,[2] Jason W. Sahl,[9] Blaz Stres,[10] Gerhard G. Thallinger,[11] David J. Van Horn,[2] and Carolyn F. Weber[12]

Métagénomique 16S: Matériel & Méthodes

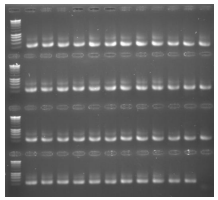
Sampling

96 DNA plate extraction

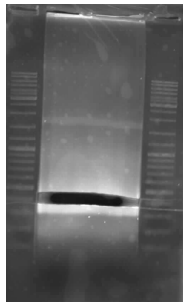
96 PCR plate: 16Sv4

Dual-indexing (*n*=864)

Agarose gel



Pooling & Gel excision

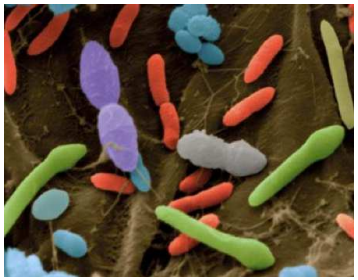


KAPA quantification (qPCR)

MiSeq paired-end sequencing

Taxonomic identification

Métagénomique 16S: Contrôles

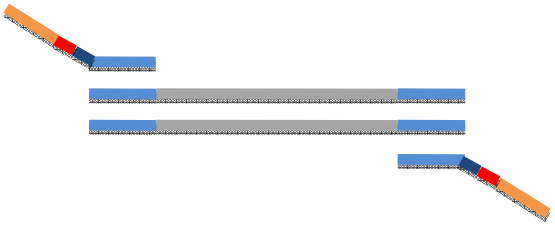


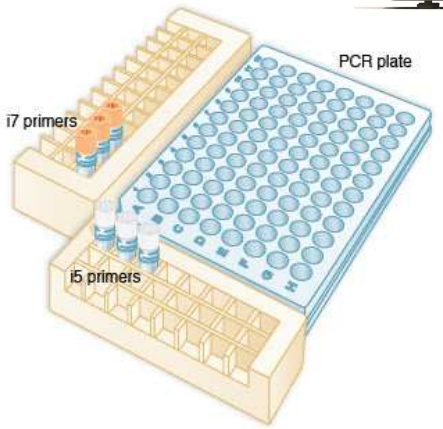
- Contrôles négatifs:
  - Dissection
  - Extraction
  - PCR
  - Primers indexés « réels (puit vide)
  - Primers indexés « virtuels »
- Contrôles positifs (souches):
  - *Borrelia*, *Bartonella*, *Rickettsia* …
  - *Mycoplasma putrefaciens* & *mycoides* (ruminant)
- Echantillons dupliqués:
  - 2 PCR indépendantes (répétabilité)

- 1 run MiSeq = 864 PCRs = 10 millions de reads 2 x251 bp

Métagénomique 16S: Une PCR = une librairie indexée

PCR to amplify region of interest





Amorces avec Adapter Illumina, Index, Pad et Primer 16S V4 universel:

SA501 5'-
AATGATACGGCGACCACCGAGATCTACACATCGTACGTATGGTAATTGTGTGCCAGCMGCCGCGGT
AA-3'

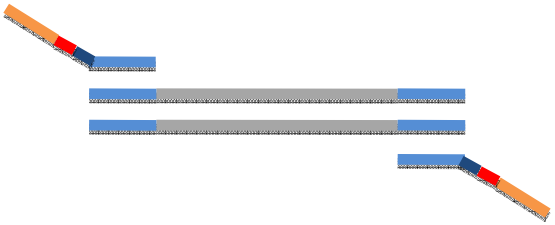
SA502 5'-
AATGATACGGCGACCACCGAGATCTACACACTATCTGTATGGTAATTGTGTGCCAGCMGCCGCGGT
AA-3'

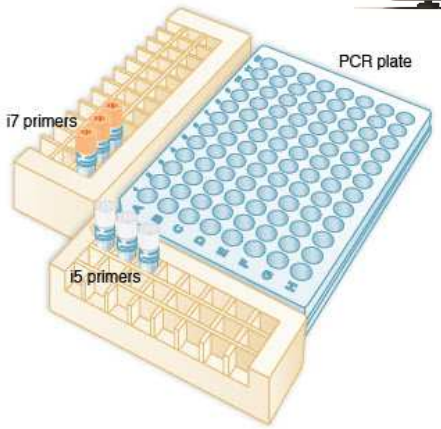
SA503 5'-
AATGATACGGCGACCACCGAGATCTACACTAGCGAGTTATGGTAATTGTGTGCCAGCMGCCGCGGT
AA-3'

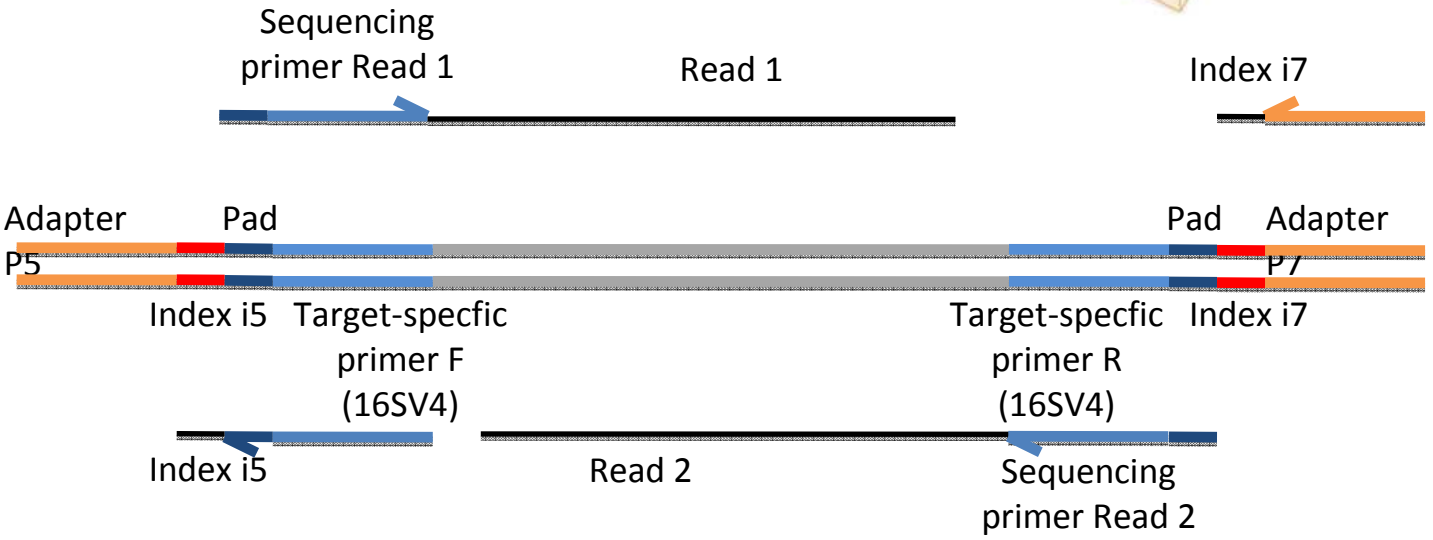
(x36)

Métagénomique 16S: séquençage MiSeq (Illumina)

PCR to amplify region of interest

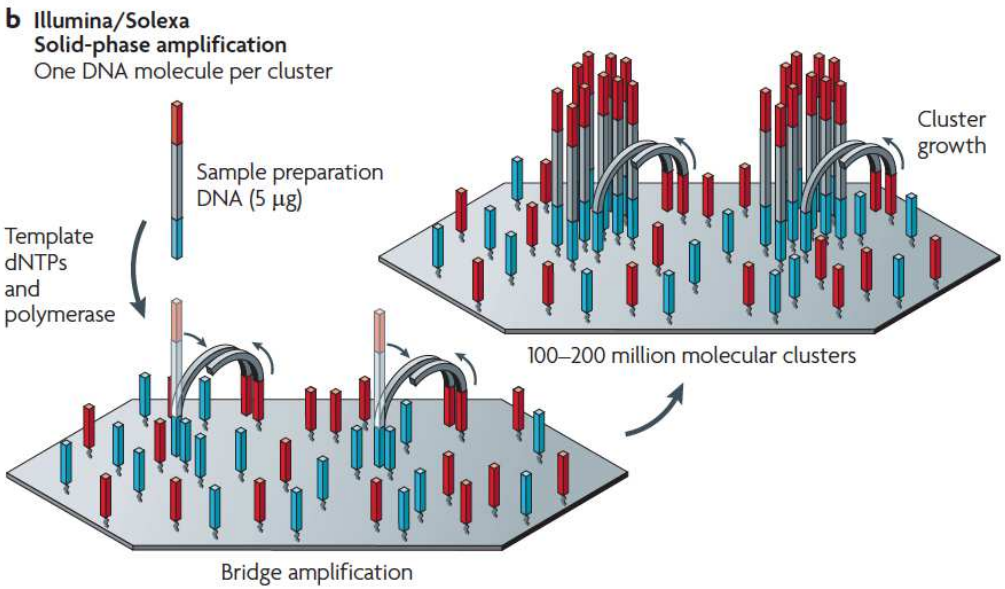


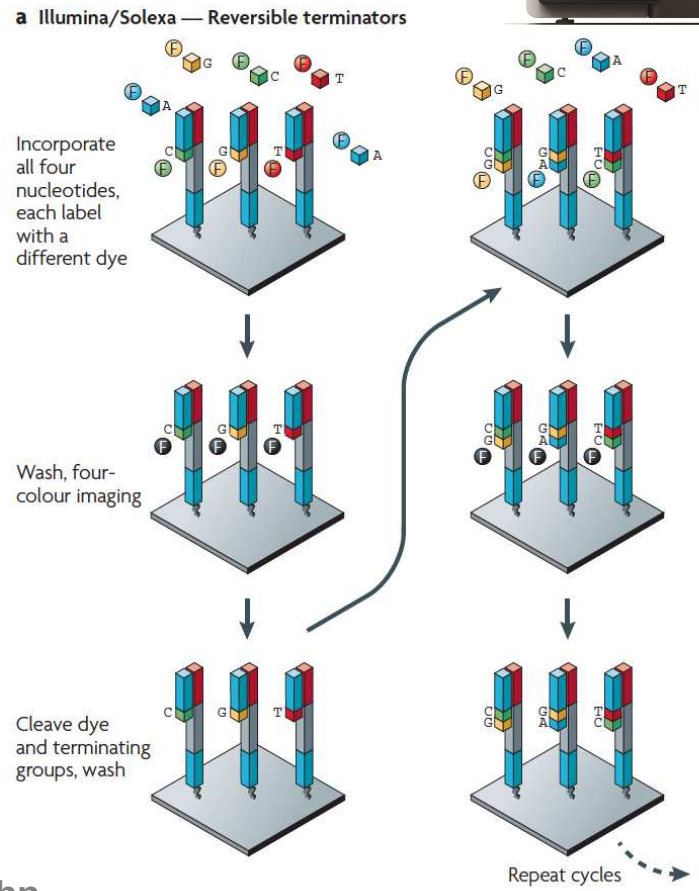




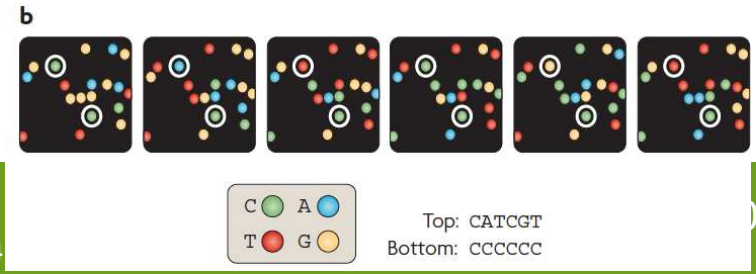
Métagénomique 16S: séquençage MiSeq (Illumina)

Illumina sequencing





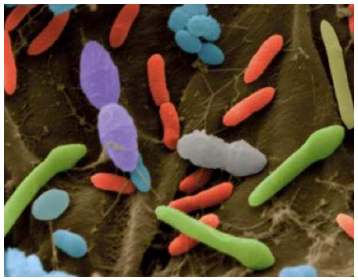
- 1 run MiSeq = 864 PCRs = 10 millions de reads 2 x 251 bp



Métagénomique 16S: Comparaison 454 vs. MiSeq

Plateforme	454 GS-FLX (Roche)	MiSeq (Illumina)
Méthode	Claesson et al 2009 Galan et al 2010	Kozich et al 2013
Barcode 16V4	207bp	251bp
Prix/run	3 400€	1 250€
Pipeline	Galaxy (Maria Bernard)	Mothur (MiSeq SOP)
Rongeurs Ardennes	62.866 reads (*n*=288)	6.502.582 x 2 reads (*n*=288)
Rongeurs Asie	32.067 reads (*n*=190)	6.879.560 x 2 reads (*n*=384) 8.221.268 x 2 reads (*n*=384)
Tiques Ardennes	226.398 reads (*n*=190)	1.136.328 x 2 reads (*n*=95)
TOTAL	321.331 reads (*n*=380)	22.739.738 reads (*n*=1151)
Moyenne/échantillon	342 reads	19.756 reads

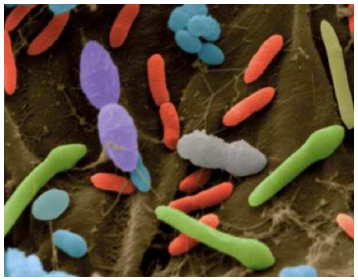
Métagénomique 16S: Analyse mothur (MiSeq SOP)



maxime — galan@r01: ~/16S_ENEMI — ssh — 80×30

```
mothur v.1.32.1
Last updated: 10/16/2013

by
Patrick D. Schloss

Department of Microbiology & Immunology
University of Michigan
pschloss@umich.edu
http://www.mothur.org

When using, please cite:
Schloss, P.D., et al., Introducing mothur: Open-source, platform-independent, co
mmunity-supported software for describing and comparing microbial communities. A
ppl Environ Microbiol, 2009. 75(23):7537-41.

Distributed under the GNU General Public License

Type 'help()' for information on the commands that are available

Type 'quit()' to exit program


mothur > classify.seqs(fasta=stability.trim.contigs.good.unique.good.filter.uniq
ue.precluster.pick.fasta, count=stability.trim.contigs.good.unique.good.filter.u
nique.precluster.uchime.pick.count_table, reference=silva.bacteria.fasta, taxono
my=silva.bacteria.silva.tax, cutoff=80)
```

Métagénomique 16S: Analyse mothur (MiSeq SOP)



- 1) Contig read 1 + read 2
- 2) Filtrage reads > 275 bp (taille attendue: 251bp)
- 3) Alignement sur séquences de référence unique SILVA (14.956 séquences de *Bacteria*)
- 4) Filtrage reads mal alignées / reads chimériques / reads eukaryotiques (18S)
- 5) Regroupement en OTUs (Operational Taxonomic Units): < 3% de divergence
- 6) Classification taxonomique bayesienne (Wang et al 2007) avec bootstrap >80%
- 7) Regroupement en phylotypes (OTUs appartenant aux mêmes taxons)

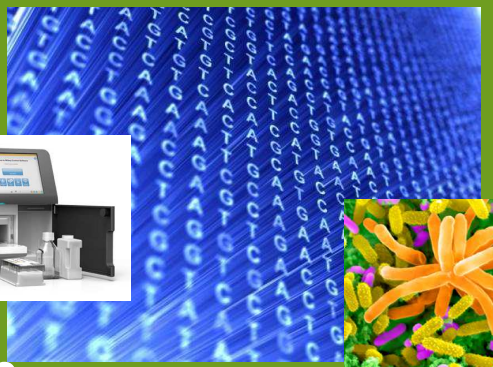
ENEMI team







MERCI!!!


PAthoID team



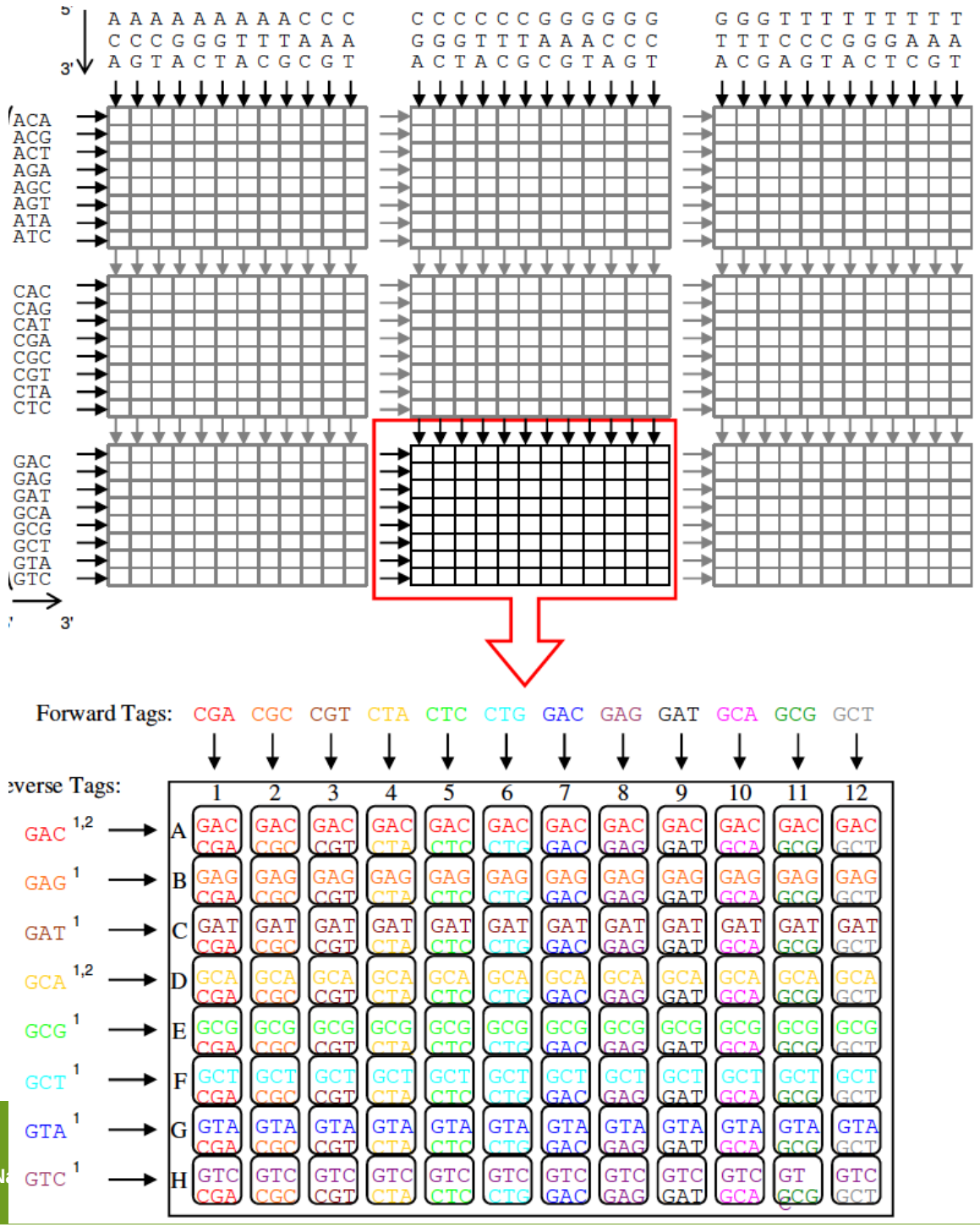



Spécial dédicace « mothur »: Marie Pagès

Les « *Maximettes* »


Hélène « MiSeq » Vignes

Alex « Cluster » Dehne


5'
3'
AAAAAAAAACCC
CCCGGGTTTAAA
AGTACTACGCGT
CCCCCCGGGGGG
GGGTTTAAACCC
ACTACGCGTAGT
GGGTTTTTTTTT
TTTCCCGGGAAA
ACGAGTACTCGT
ACA
ACG
ACT
AGA
AGC
AGT
ATA
ATC
CAC
CAG
CAT
CGA
CGC
CGT
CTA
CTC
GAC
GAG
GAT
GCA
GCG
GCT
GTA
GTC
3'
Forward Tags: CGA CGC CGT CTA CTC CTG GAC GAG GAT GCA GCG GCT
everse Tags:
1 2 3 4 5 6 7 8 9 10 11 12
GAC 1,2
GAG 1
GAT 1
GCA 1,2
GCG 1
GCT 1
GTA 1
GTC 1
A
B
C
D
E
F
G
H