From yesterday to tomorrow: past, present and future of sequencing

The NGS revolution

Laurent FARINELLI

19 October 2017

ICCMg2 Second International Conference on Clinical Metagenomics

Geneva, Switzerland

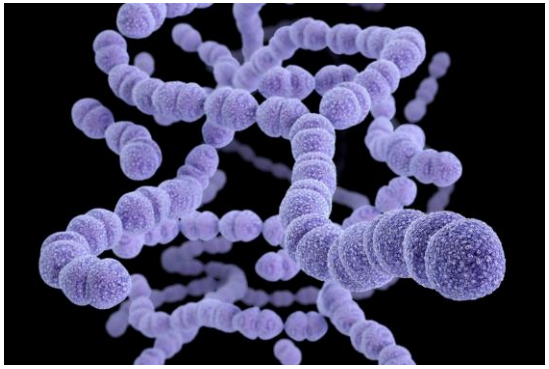
International Conference on Clinical Metagenomics

Be at the right place and time

1996

GlaxoWellcome

Geneva Biomedical Research Institute



Streptococcus pneumoniae

7 instruments ABI 377

Jonathan Knowles:

- Have fun in you work
- *"The right drug for the right patient"*

Bead Arrays

- September:

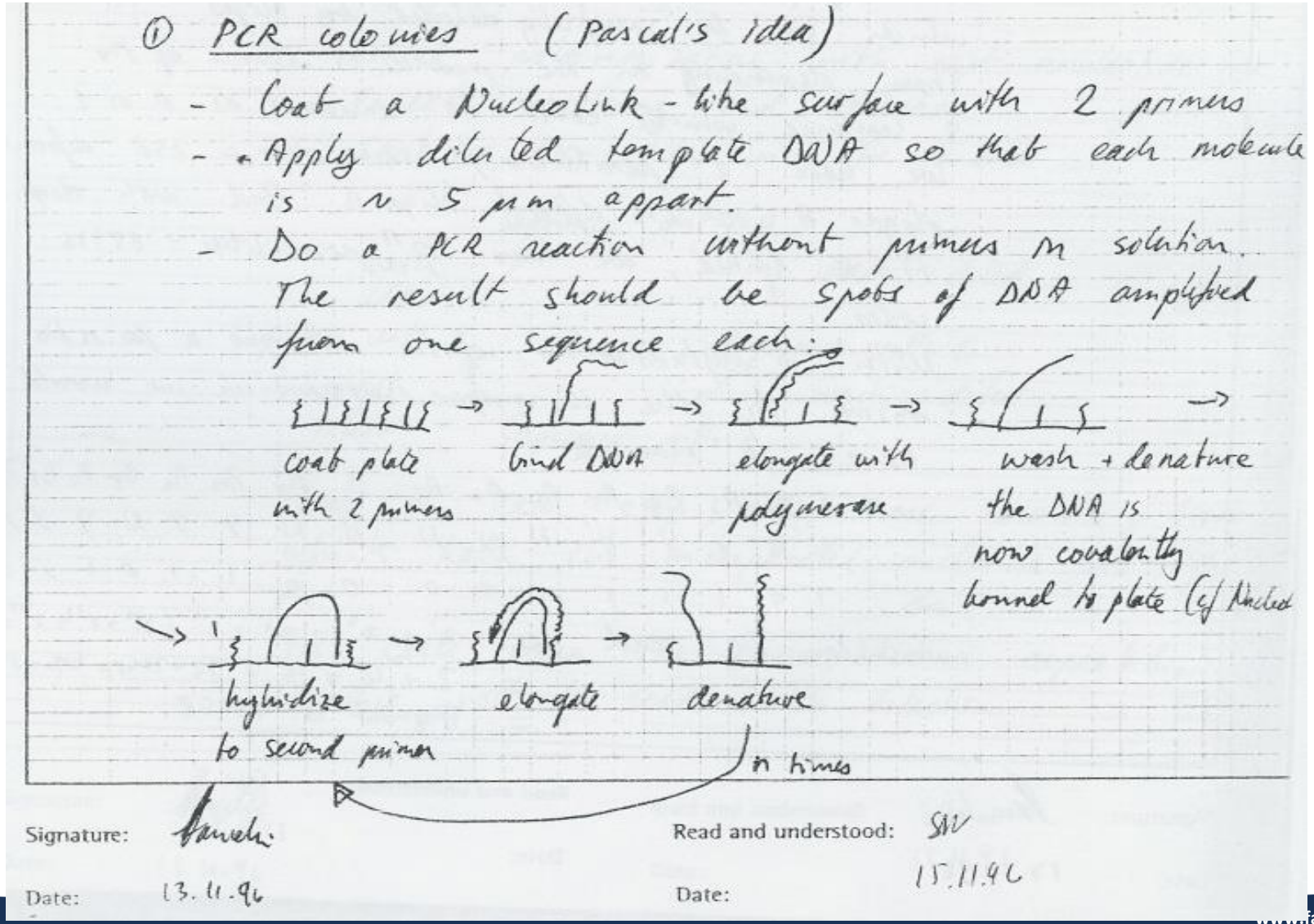
  I was asked, with Pascal Mayer, to develop faster sequencing

- October:

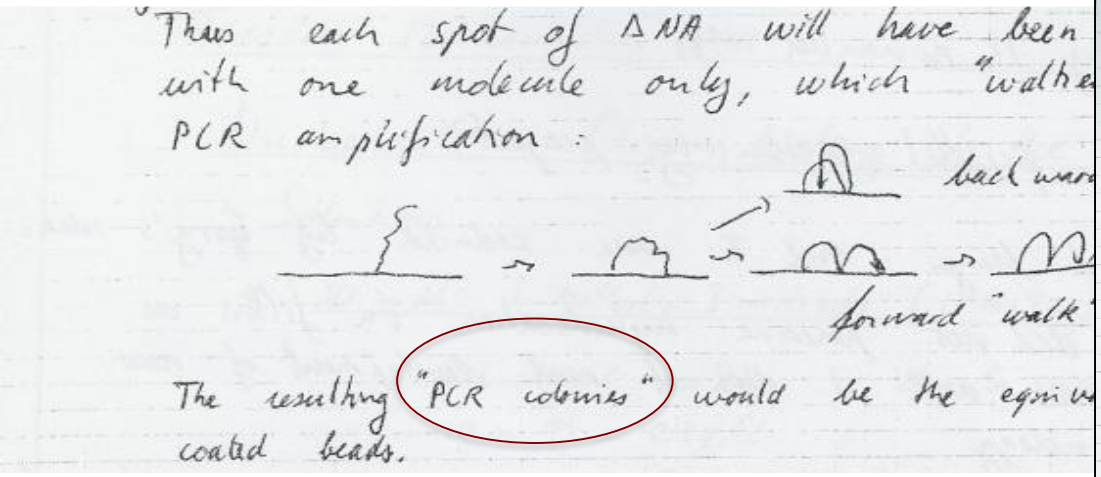
  Heard of Bead Arrays: brilliant, but too complicated

Lynx Therapeutics' MegaClone
Massively Parallel Signature Sequencing (**MPSS**)
Gene Expression Profiling
Brenner, S. *et al.*, Nature Biotechnology **18**:630-634 (2000)

Invention of PCR Colonies



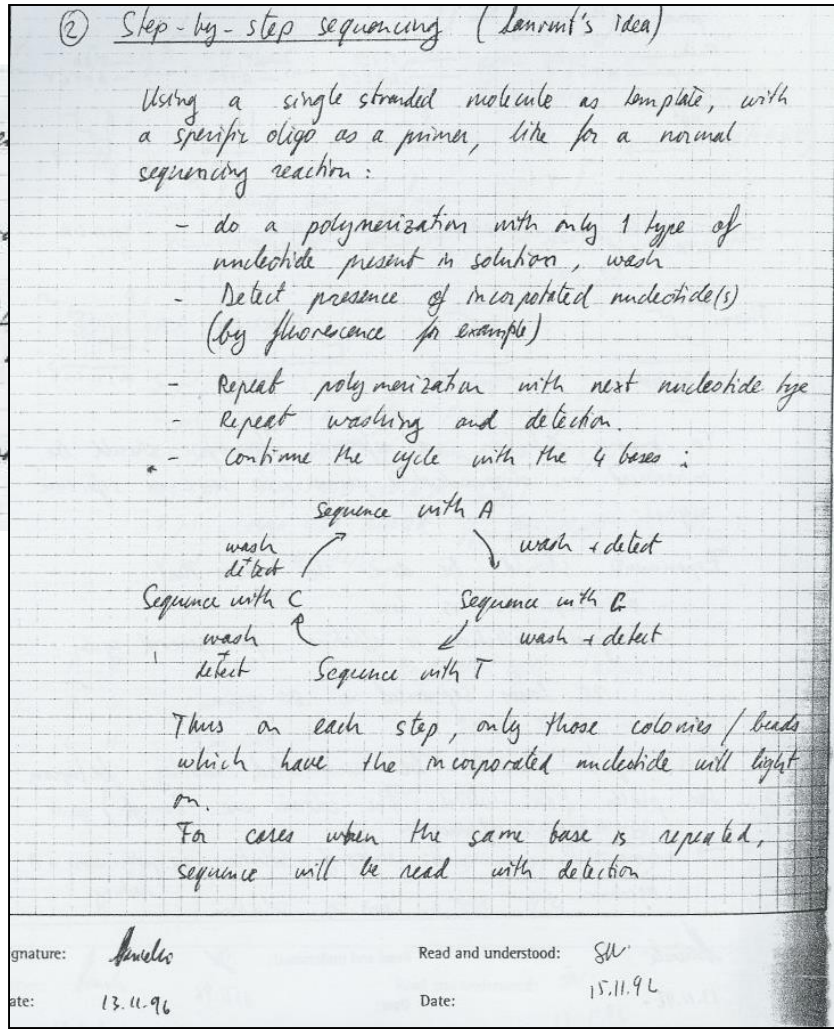
A special day


"Thus each spot of DNA will have been with one molecule only, which "walke PCR amplification :

back ward

forward "walk

The resulting "PCR colonies" would be the equiv coated beads.


② Step-by-step sequencing (Laurent's idea)

Using a single stranded molecule as template, with a specific oligo as a primer, like for a normal sequencing reaction :

- do a polymerization with only 1 type of nucleotide present in solution, wash
- Detect presence of incorporated nucleotide(s) (by fluorescence for example)
- Repeat polymerization with next nucleotide type
- Repeat washing and detection.
- Continue the cycle with the 4 bases :

Sequence with A
wash + detect
Sequence with C
wash + detect
Sequence with G
wash + detect
Sequence with T
wash + detect

Thus on each step, only those colonies / beads which have the incorporated nucleotide will light on.
For cases when the same base is repeated, sequence will be read with detection

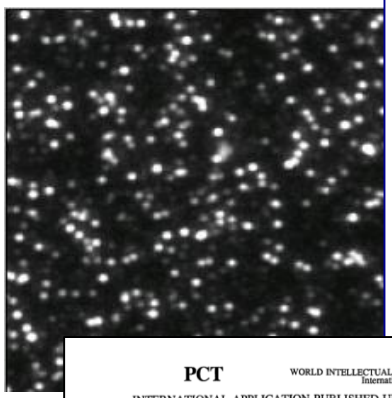
gnature:
ate: 13.11.96
Read and understood:
Date: 15.11.96

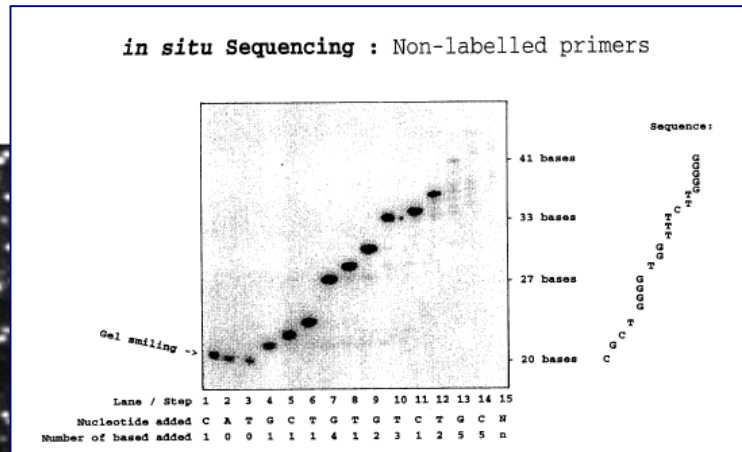
13 November 1996:

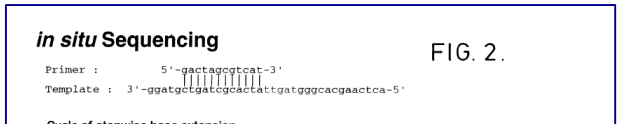
Invention of PCR Colonies and step-by-step sequencing

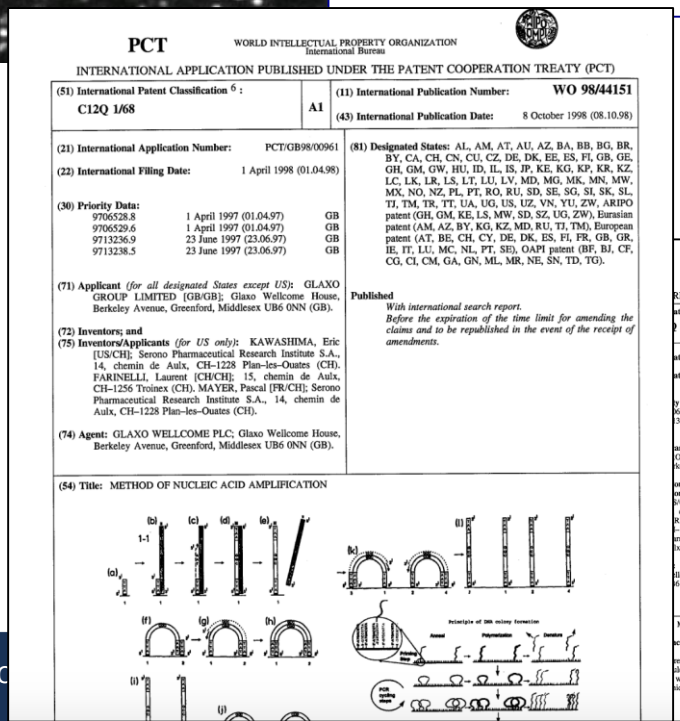
Two patents filed

- By X'Mas:
  - Proof-of-principles
- 1 April 1997:
  - Two Patents filed







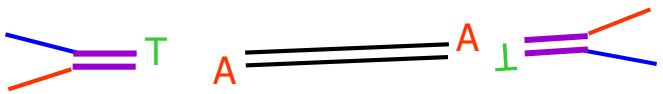


PCR Colonies project

- 1996-1997 GlaxoWellcome
- 1998-2000 Serono
- 2001-2003 GenInEx / **Manteia Predictive Medicine**

With your genome on CD you consult our database to predict response to treatment or risk of disease

  - Library preparation, including Y-shaped adapter (Laurent) and P5, P7 (Magne)
  - Instrument (Microscope and flow-cell)
  - Reverse terminator sequencing chemistry
  - Software for real-time DNA colonies detection and base-calling



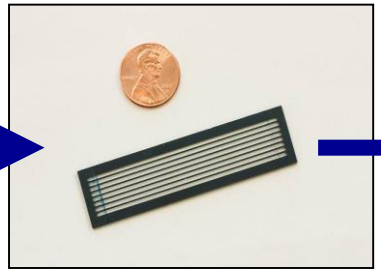
- 2004-2006 Solexa & Lynx Therapeutics ($4 M)
- 2007- illumina ($600 M)

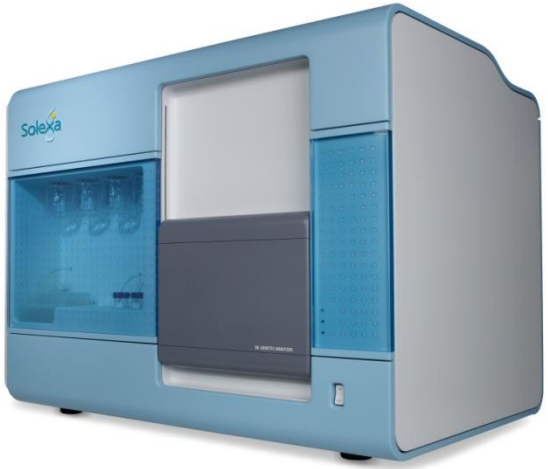
2003, Fasteris

Everything started in a chalet...

NGS Pioneers

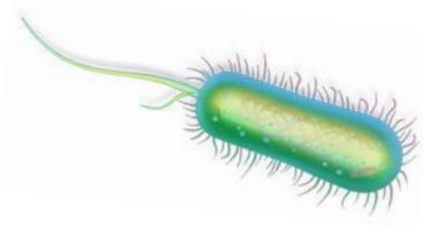
Solexa Cluster Station

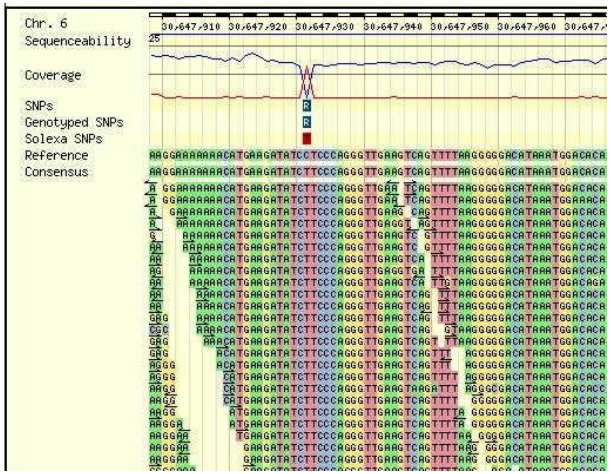




Solexa Genome Analyzer

Library Preparation Kit

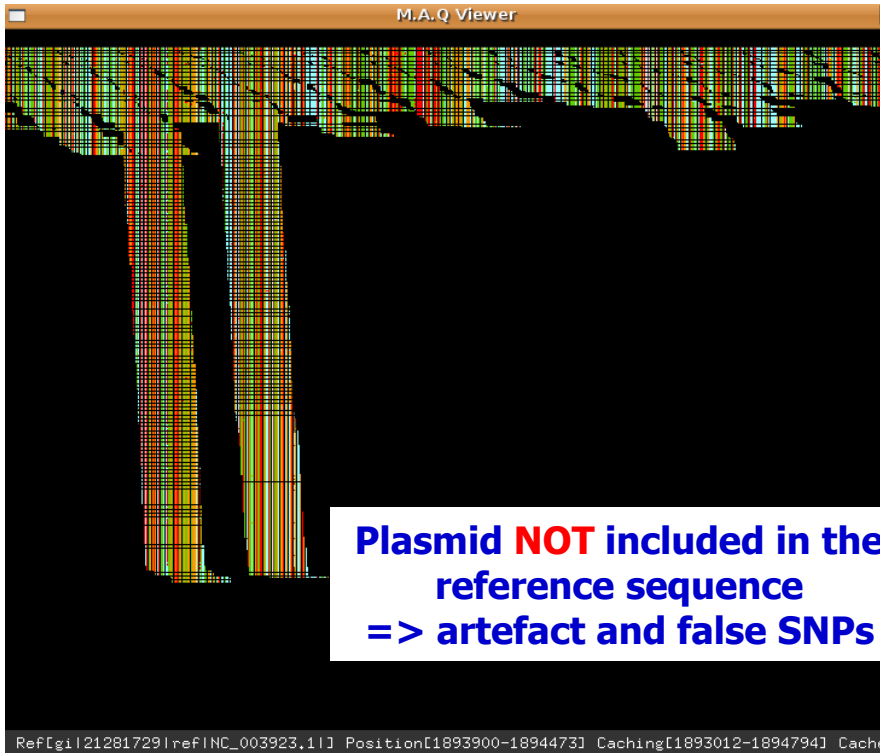




Data analysis

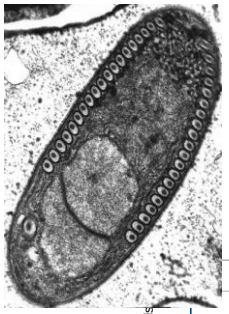
2007 installation run

- *Staphylococcus aureus* for HUG Genomic Laboratory
- 8 lanes of 500’000 reads of 26 bases
- Difficult to analyze data:
  I had to use *mySQL* database to sort reads
- I could easier assemble genomes using EDENA than map on reference

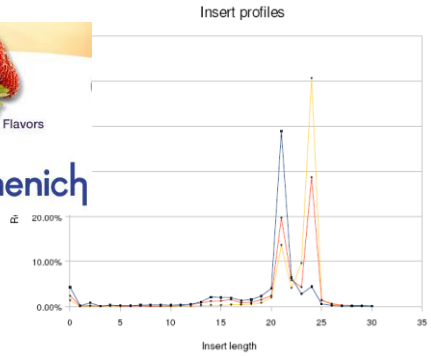


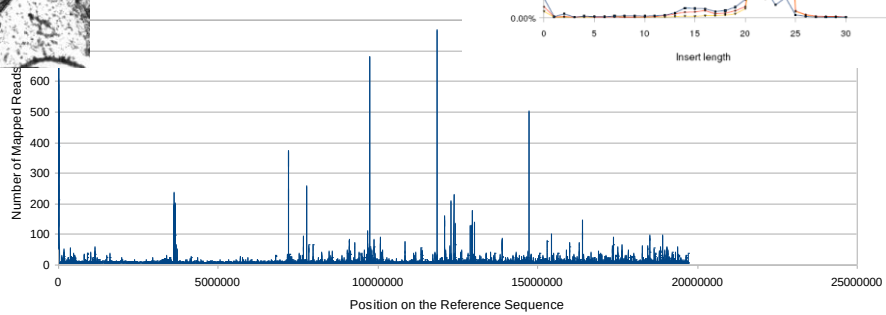
NGS Pioneers

- We developed:
  - Genomes
  - *de novo* assembly of transcriptome
  - ChIP-SEQ
  - Bar-coded small RNAs



Microsporidia





- At the end of 2007, proud to be considered by illumina
  - First service lab to buy the machine
  - Smallest lab to use the technology
  - The lab that developed the broadest range of applications

Today applications

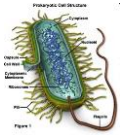
Medical diagnostics

- Under Swiss law for genetics

Bestsellers

- Batches every 1-2 weeks
- Stable protocols, short TAT, lower cost, higher volumes

Personalized



- 1-2 batch per month
- Many options, partnership with researchers

Research

- Developing new protocols

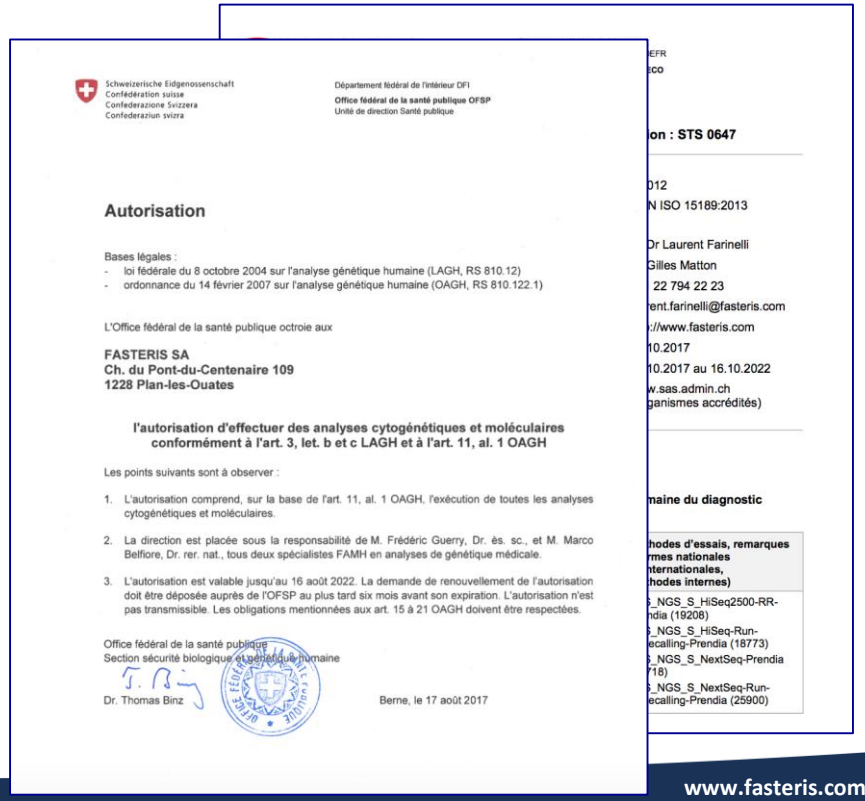
Medical applications

- Non-invasive prenatal test (NIPT)

prendia EXPERT — 2013: Whole genome

prendia START — 2015: Reimbursed T13, 18, 21

- Somatic Cancer Panels
- Germline Cancer Panels
- Preimplantation Diagnostic (PGD-A)
- 2017: Fasteris accreditation and genetics authorization



Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Département fédéral de l'intérieur DFI
Office fédéral de la santé publique OFSP
Unité de direction Santé publique

Autorisation

Bases légales :
- loi fédérale du 8 octobre 2004 sur l'analyse génétique humaine (LAGH, RS 810.12)
- ordonnance du 14 février 2007 sur l'analyse génétique humaine (OAGH, RS 810.122.1)

L'Office fédéral de la santé publique octroie aux

FASTERIS SA
Ch. du Pont-du-Centenaire 109
1228 Plan-les-Ouates

l'autorisation d'effectuer des analyses cytogénétiques et moléculaires conformément à l'art. 3, let. b et c LAGH et à l'art. 11, al. 1 OAGH

Les points suivants sont à observer :

1. L'autorisation comprend, sur la base de l'art. 11, al. 1 OAGH, l'exécution de toutes les analyses cytogénétiques et moléculaires.
2. La direction est placée sous la responsabilité de M. Frédéric Guerry, Dr. ès. sc., et M. Marco Belfiore, Dr. rer. nat., tous deux spécialistes FAMH en analyses de génétique médicale.
3. L'autorisation est valable jusqu'au 16 août 2022. La demande de renouvellement de l'autorisation doit être déposée auprès de l'OFSP au plus tard six mois avant son expiration. L'autorisation n'est pas transmissible. Les obligations mentionnées aux art. 15 à 21 OAGH doivent être respectées.

Office fédéral de la santé publique
Section sécurité biologique et génétique humaine

Dr. Thomas Binz

Berne, le 17 août 2017

Bestsellers applications

- Genomes
- Exomes
- Transcriptomes
- Small RNA

PNAS

Proceedi

The ge

invicta

Yannick Wu

Miguel Cor

Laurent Fal

The NEW ENGLAND JOURNAL of MEDICINE

HOME ART

nature International weekly journal of science

Archive Volume 479 Issue 7374 Articles Article

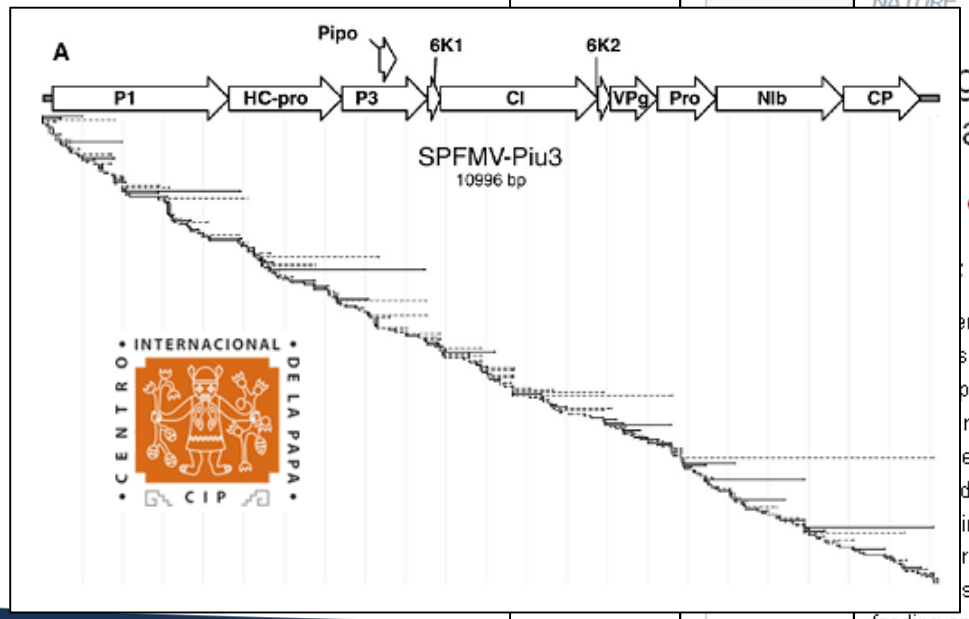
NATURE | ARTICLE OPEN previous

genome of *Tetranychus urticae* reveals h

adaptations

Grbić, Thomas Van Leeuwen, Richard M. Clark, Stephane Rombauts,

r mite *Tetranychus urticae* is a cosmopolitan agricultural pest with
sive host plant range and an extreme record of pesticide resistance.
present the completely sequenced and annotated spider mite
representing the first complete chelicerate genome. At 90
es *T. urticae* has the smallest sequenced arthropod genome.
d with other arthropods, the spider mite genome shows unique
in the hormonal environment and organization of the Hox complex,
reveals evolutionary innovation of silk production. We find strong
s of polyphagy and detoxification in gene families associated with
feeding on different hosts and in new gene families acquired by lateral gene
transfer. Deep transcriptome analysis of mites feeding on different plants
shows how this pest responds to a changing host environment. The *T. urticae*



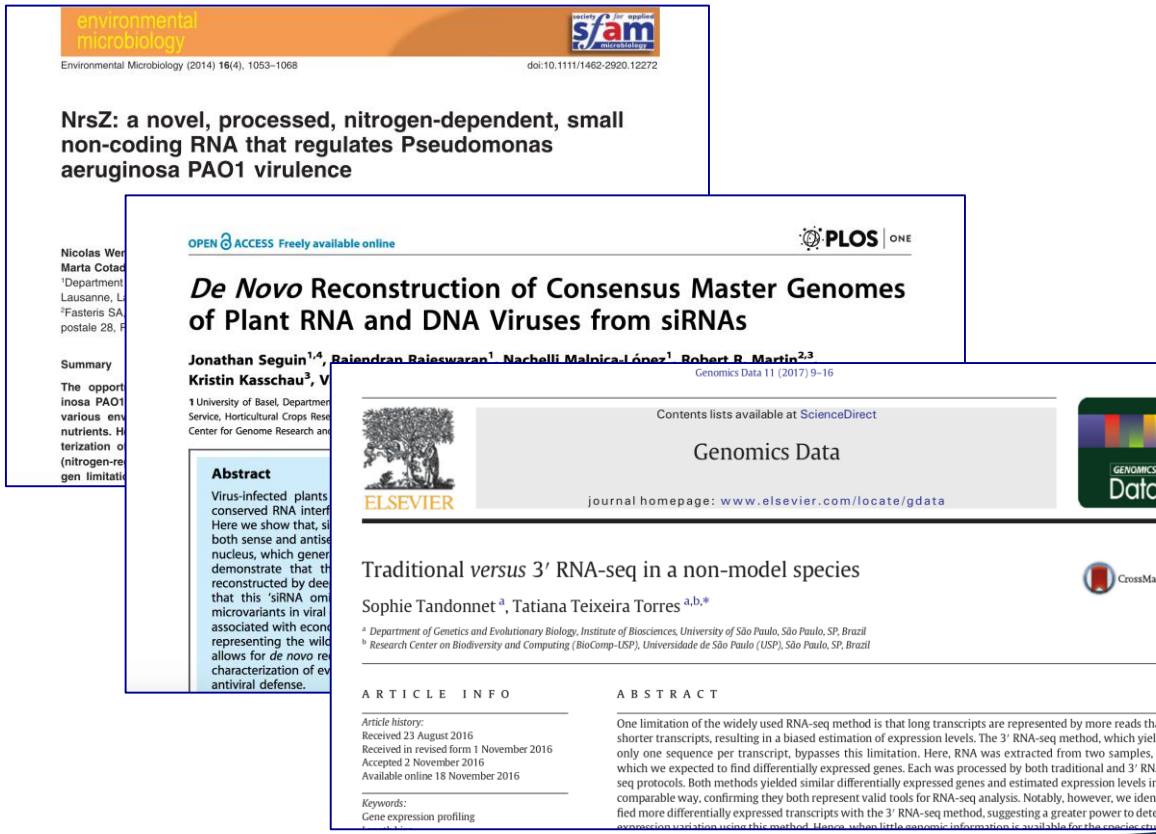
Hairy-cell
whose unc

Personalized applications

- Many different protocols
- Several options, *e.g.* Ribozero to eliminate host and bacterial rRNA
- MetaFAST

Research

- Custom protocols



2008 NGS Metagenomics

- First metagenomics paper using illumina NGS?
- Genomic Research Laboratory University Hospital of Geneva
- Oral samples, amplicons of the 16S rRNA V5 variable region

Journal of Microbiological Methods 79 (2009) 266–271

Contents lists available at ScienceDirect

Journal of Microbiological Methods

journal homepage: www.elsevier.com/locate/jmicmeth

Metagenomic study of the oral microbiota by Illumina high-throughput sequencing

Vladimir Lazarevic [a,*,1], Katrine Whiteson [a,*,1], Susan Huse [b], David Hernandez [a], Laurent Farinelli [c], Magne Østerås [c], Jacques Schrenzel [a], Patrice François [a]

[a] Genomic Research Laboratory, Geneva University Hospitals, Rue Gabrielle-Perret-Gentil 4, CH-1211 Geneva 14, Switzerland
[b] Josephine Bay Paul Center, Marine Biological Laboratory, Woods Hole, MA 02543, USA
[c] Fasteris, Ch. du Pont-du-Centenaire 109, Case postale 28, CH-1228 Plan-les-Ouates, Switzerland

ARTICLE INFO

Article history:
Received 11 September 2009
Accepted 14 September 2009
Available online 29 September 2009

Keywords:
Metagenomics
Oral cavity
Flora composition
Microbiome
High-throughput sequencer

ABSTRACT

To date, metagenomic studies have relied on the utilization and analysis of reads obtained using 454 pyrosequencing to replace conventional Sanger sequencing. After extensively scanning the 16S ribosomal RNA (rRNA) gene, we identified the V5 hypervariable region as a short region providing reliable identification of bacterial sequences available in public databases such as the Human Oral Microbiome Database. We amplified samples from the oral cavity of three healthy individuals using primers covering an ~82-base segment of the V5 loop, and sequenced using the Illumina technology in a single orientation. We identified 135 genera or higher taxonomic ranks from the resulting 1,373,824 sequences. While the abundances of the most common phyla (*Firmicutes*, *Proteobacteria*, *Actinobacteria*, *Fusobacteria* and *TM7*) are largely comparable to previous studies, *Bacteroidetes* were less present. Potential sources for this difference include classification bias in this region of the 16S rRNA gene, human sample variation, sample preparation and primer bias. Using an Illumina sequencing approach, we achieved a much greater depth of coverage than previous oral microbiota studies, allowing us to identify several taxa not yet discovered in these types of samples, and to assess that at least 30,000 additional reads would be required to identify only one additional

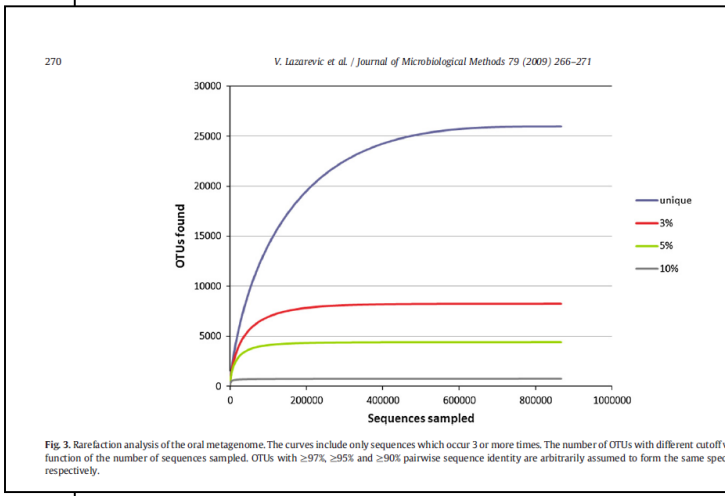


Fig. 3. Rarefaction analysis of the oral metagenome. The curves include only sequences which occur 3 or more times. The number of OTUs with different cutoff v function of the number of sequences sampled. OTUs with ≥97%, ≥95% and ≥90% pairwise sequence identity are arbitrarily assumed to form the same spec respectively.

Clinical Metagenomics

FASTERIS member of the sonic healthcare network

SONIC HEALTHCARE

Collaboration with HUG Hôpitaux Universitaires Genève

- Impact on survival after cystic fibrosis lung transplant
- Looking for viruses in glioblastoma
- Shotgun approach to find viruses and fungi

Collaboration with ortho-analytic Labor für Human- und Umweltmedizin · Mitglied des Medisupport Netzwerkes

biovis DIAGNOSTIK

ETH zürich

WILEY INTERNATIONAL JOURNAL OF CANCER. JOURNAL INTERNATIONAL DU CANCER

This Article | For Authors | Learn More | Submit

Int J Cancer. 2014 Sep 15; 135(6): 1381–1389. PMCID: PMC4235296
Published online 2014 May 8. doi: 10.1002/ijc.28670

Comprehensive metagenomic analysis of glioblastoma reveals absence of known virus despite antiviral-like type I interferon gene response

Érika Cosset,[1,*] Tom J Petty,[2,3,*] Valérie Dutoit,[4] Samuel Cordey,[5] Ismael Padioleau,[2,3] Patricia Otten-Hernandez,[6] Laurent Farinelli,[6] Laurent Kaiser,[5] Pascale Bruyère-Cerdan,[1] Diderik Tirefort,[1] Soraya Amar El-Dusouqui,[1] Zeynab Nayernia,[7] Karl-Heinz Krause,[7] Evgeny M Zdobnov,[2,3] Pierre-Yves Dietrich,[4] Emmanuel Rigal,[1] and Olivier Preynat-Seauve[1]

Author information ▶ Article notes ▶ Copyright and License information ▶

This article has been

Abstract

Glioblastoma is a
to be caused by a
such as cytomega
glioblastoma rema
validated a robust
traditional clinical
potentially discov
viruses were dete
signature and unc
methodology that

Keywords: gliob

19 Octo

ELSEVIER

Journal of Microbiological Methods

Volume 79, Issue 3, December 2009, Pages 266-271

Metagenomic study of the oral microbiota by Illumina high-throughput sequencing

Vladimir Lazarevic [a] [1] ✉, Katrine Whiteson [a] [1] ✉, Susan Huse [b], David Hernandez [a], Laurent Farinelli [c], Magne Østerås [c], Jacques Schrenzel [a], Patrice François [a]

ers
robiology

ORIGIN
published: 07
doi: 10.3389/fm

Microbial Communities of Conducting and Respiratory Zon of Lung-Transplanted Patients

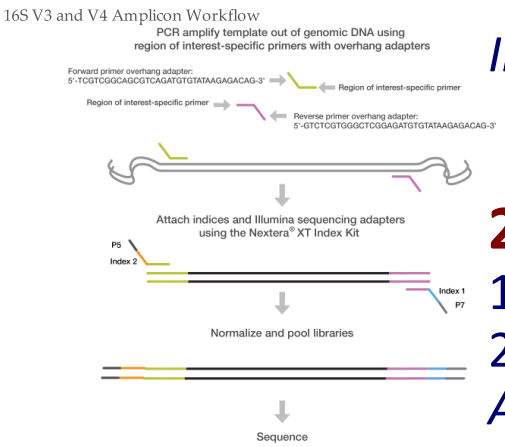
Marie Beaume[1], Vladimir Lazarevic[2†], Thilo Köhler[1†], Nadia Gaïa[2], Oriol Manuel John-David Aubert[4], Loïc Baerlocher[5], Laurent Farinelli[5], Paola Gasche[6], Jacques Schrenzel[2], Christian van Delden[1] and the Swiss Transplant Cohort S*

[1] Transplant Infectious Diseases Unit, Geneva University Hospitals and Department of Microbiology and Mole University of Geneva, Geneva, Switzerland, [2] Genomic Research Laboratory, Geneva University Hospitals, Ge Switzerland, [3] Infectious Diseases Service and Transplantation Center, Lausanne University Hospital Center, L Switzerland, [4] Division of Pulmonary Diseases, Lausanne University Hospital Center, Lausanne, Switzerland, Plan-les-Ouates, Switzerland, [6] Division of Pulmonary Diseases, Geneva University Hospitals, Geneva, Switze

Metagenomics protocols

16S Metagenomic Sequencing Library Preparation

Preparing 16S Ribosomal RNA Gene Amplicons for the Illumina MiSeq System



1-step PCR

Using primers with **long tags** (>60 nucleotides)

Advantage: - Add the entire illumina adapter sequence in one step.

Inconvenient: - Long primers

Expensive and significant impact on amplification efficiency

2-steps PCR

1st step with primers with medium tag (30 nt)

2nd step with illumina primers

Advantage: - Only one primer pair needed for 1st step.

Inconvenient: - 2 PCR reactions needed

Libraries with low complexity

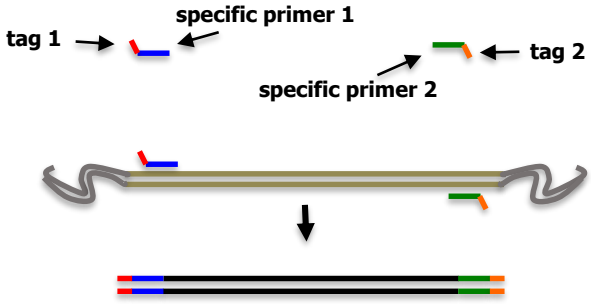
All inserts in the same orientation with similar sequences

=> Impact on sequence quality

MetaFast

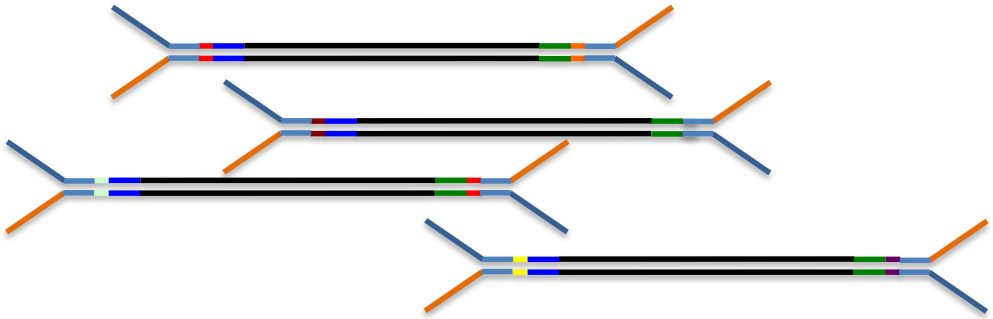
Developed by Fasteris

Step 1: PCR using primers with short tags



Step 2: Normalization and pooling

Step 3: End-repair, A-tailing and TruSeq adapter ligation



	Tags
1	NNAACCGT
2	NNCGATGT
3	NNNGTGAGC
4	NNNTCCGCT
5	NNNNACAGTG
6	NNNNCTTGTA
7	NNGACTCT
8	NNTGACCA
9	NNNATCACG
10	NNNCCGGAG
11	NNNNGGCTAC
12	NNNNTAGCTT
13	NNACTTGA
14	NNCAGATC
15	NNNGGTATA
16	NNNTTTAGG
17	NNNNAGAAGA
18	NNNNGATCAG
19	NNTTGTTC
20	NNATTCGC
21	NNNCGTTAA
22	NNNGAGGAT
23	NNNNTTAGGC
24	NNNNGCCAAT

Improvements by MetaFast

- Standard library preparation (Nano) ≈ 10% chimera
- Standard without PCR ≈ 1% chimera
- MetaFast (modified end-repair) ≈ 0.1% chimera

Samples 1 – 8, Fwd Tag

Rev Tags Found

	TagF1		TagF2		TagF3		TagF4		TagF5		TagF6		TagF7		TagF8	
TagR1	90.44	99.76	1.61	0.04	0.89	0.02	1.37	0.06	1.79	0.02	1.49	0.03	1.1	0.02	1.32	0.05
TagR2	1.24	0.03	90.51	99.8	1.1	0.02	1.44	0.05	2.08	0.03	1.21	0.03	0.96	0.01	1.45	0.03
TagR3	1.13	0.03	1.62	0.02	90.16	99.82	1.61	0.04	1.82	0.02	1.24	0.04	0.96	0.02	1.46	0.02
TagR4	0.83	0.01	1	0.01	0.83	0	93.39	99.92	1.42	0.03	0.85	0.01	0.68	0	0.99	0.01
TagR5	1.03	0.03	1.47	0.02	0.92	0.01	1.42	0.04	91.69	99.83	1.05	0.02	0.83	0.02	1.6	0.03
TagR6	1.47	0.04	1.28	0.02	0.87	0.02	1.32	0.07	1.64	0.04	90.88	99.71	1.2	0.07	1.32	0.04
TagR7	1.72	0.03	1.7	0.02	1.22	0.02	1.59	0.05	2.06	0.03	1.78	0.04	88.01	99.75	1.94	0.06
TagR8	1.16	0.04	1.64	0.03	0.96	0.02	1.54	0.04	2.36	0.02	1.32	0.03	1.19	0.01	89.84	99.82

With 5 cycles PCR

MetaFast (no PCR)

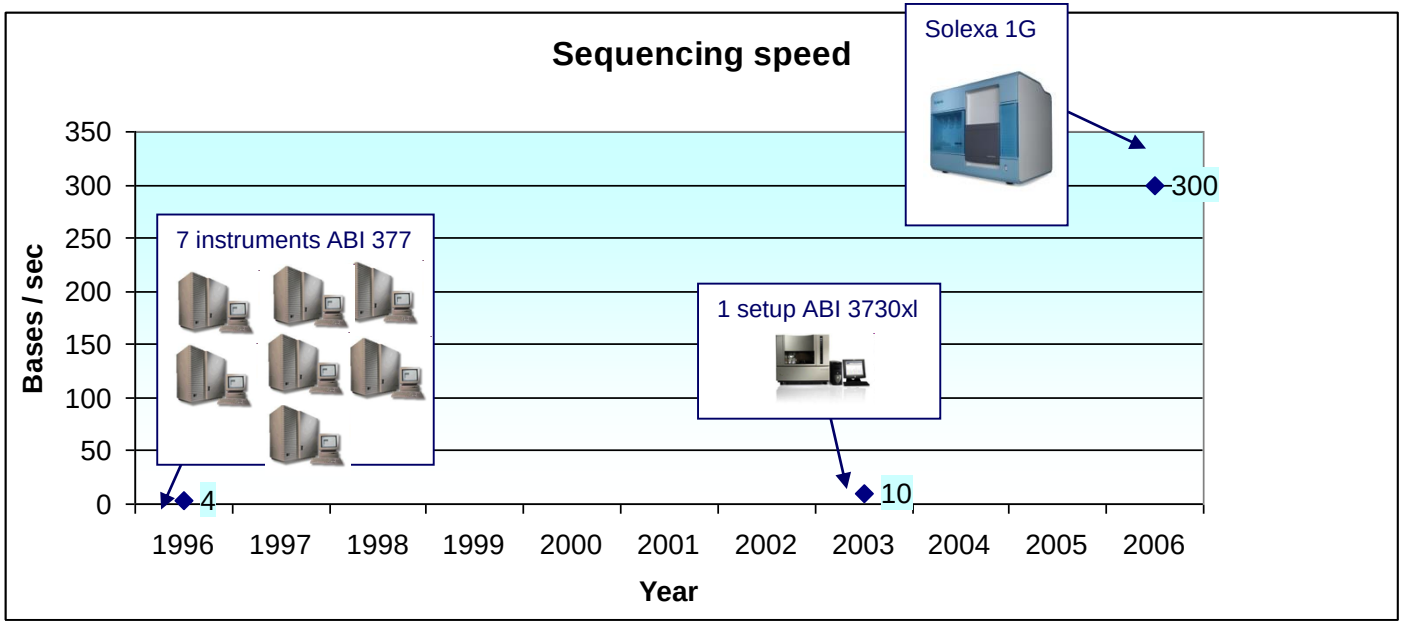
Advantages of MetaFast

- Optimized pipeline for **large number of samples**
  Large number of PCRs can be pooled to prepare a single illumina library
- Improved **quality**
  - Short tags on primers:
    Very low incidence on amplification efficiency
  - Insert present in both forward and reverse orientation:
    Increases sequence variability
  - Reduced chimera
- Lower **cost**

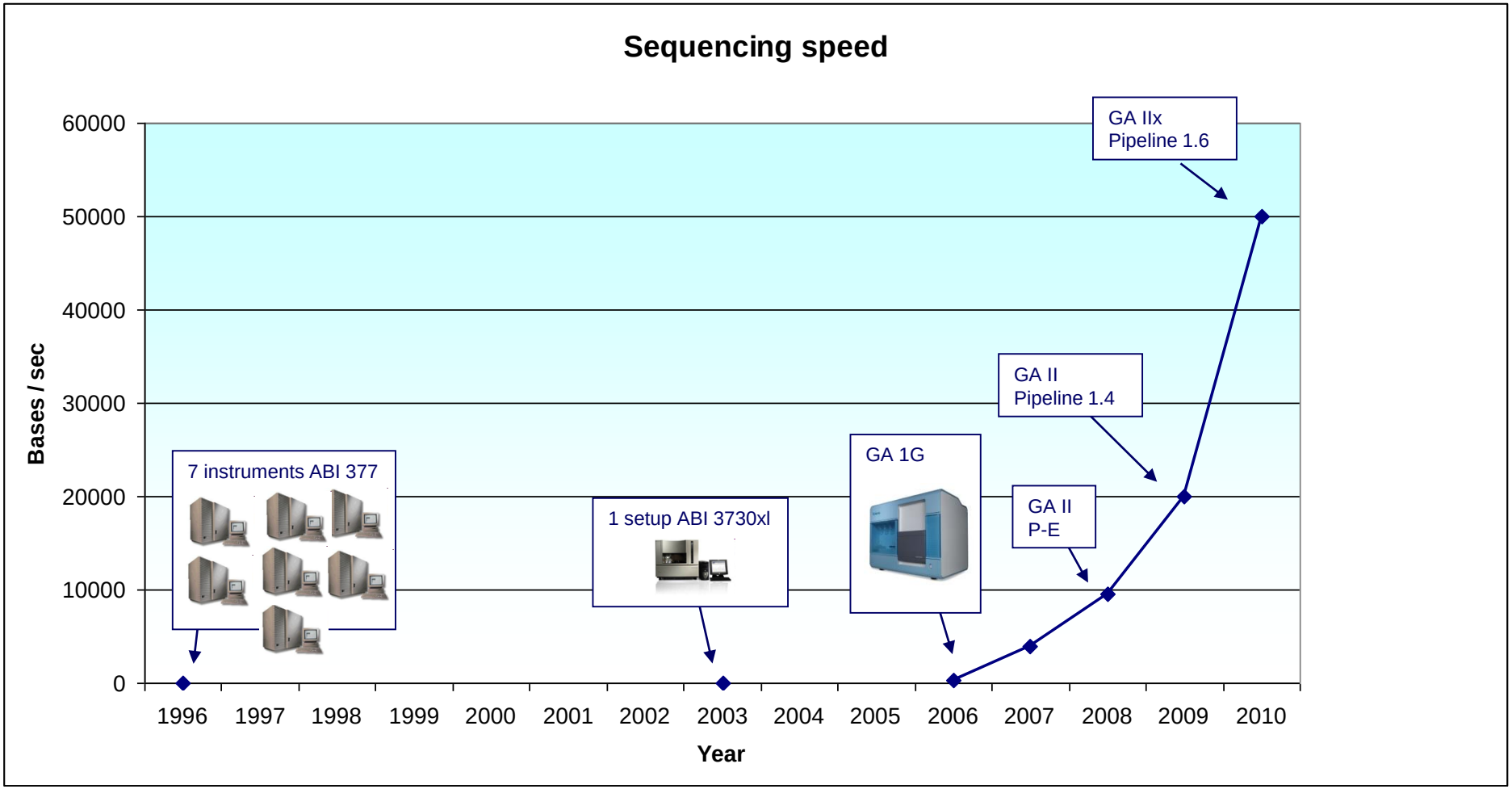
Future of NGS

- When looking at the last 10 years, we can clearly speak of a revolution
- Speed and cost

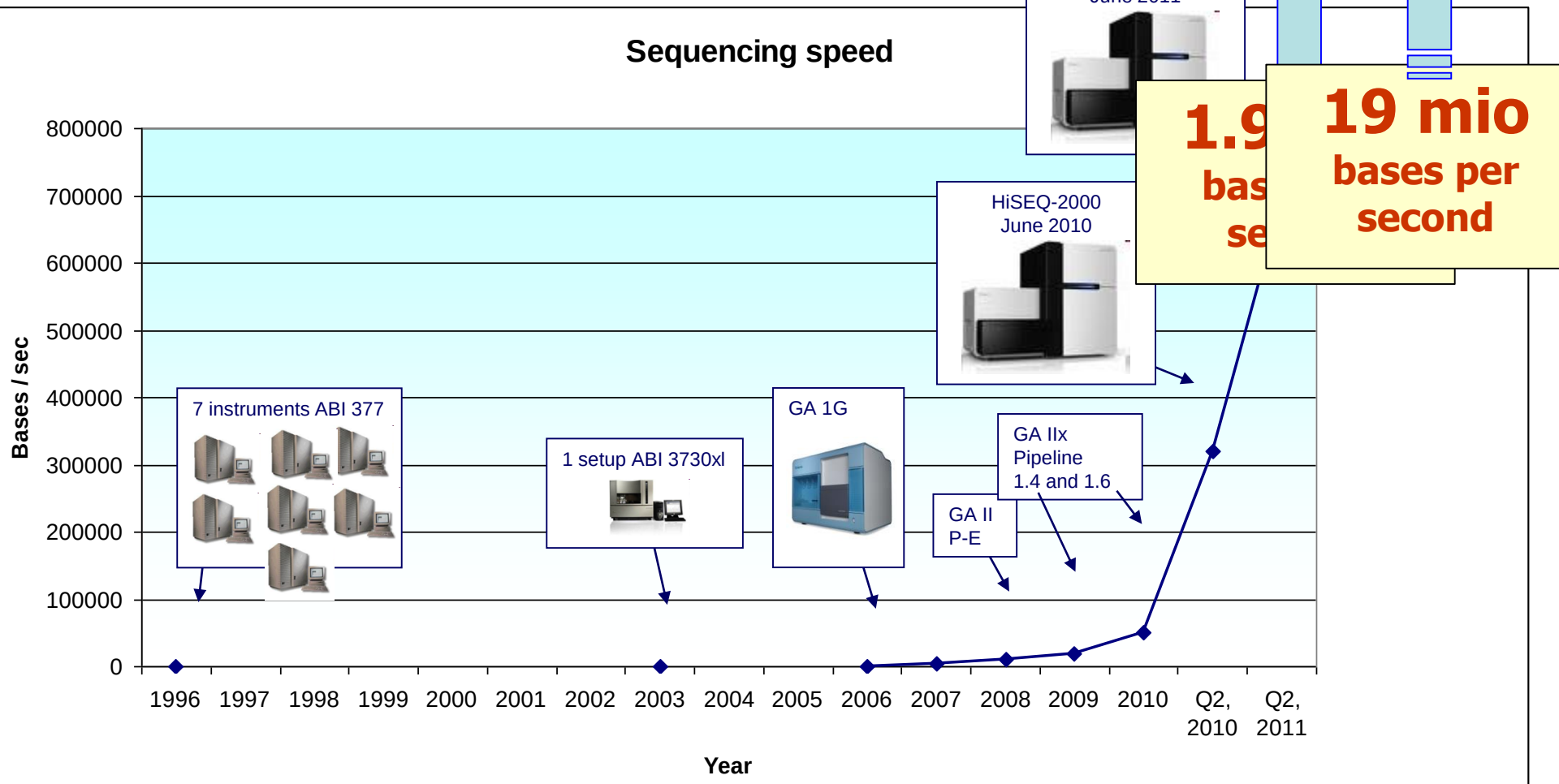
Future of NGS and Metagenomics



Tremendous speed increase



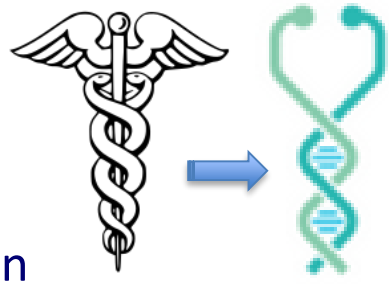
Tremendous speed incre



Future of NGS

- When looking at the last 10 years, we can clearly speak of a revolution
- Speed and cost

- It’s becoming cheaper to generate data than to store it
- We see now incredible growth in medicine

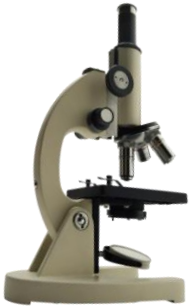


- Challenges as well, in particular for
  - Quality and Validation of data, CE-IVD, accreditation
  - Trend for automatized “For Diagnostics” systems
    => Importance of analysis

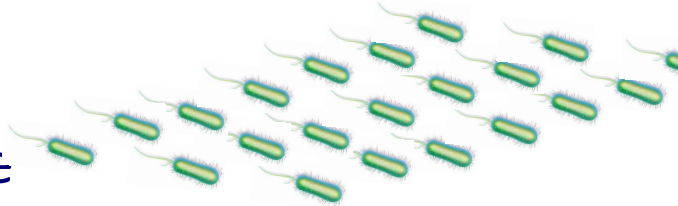
Future of NGS

- Changing role for doctors, who cannot know everything
  - Partnership with patients
  - Importance of Human experience
- Screening everywhere?
  => What about the right of **NOT** knowing?
- It is our responsibility to apply this technology for the good of humankind and not for profit

Future of Metagenomics

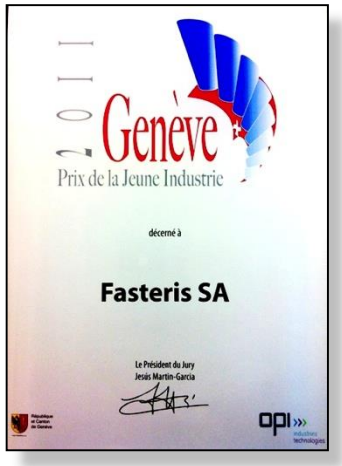


- We are discovering new universes
- We are discovering how bacteria manipulate us
- It is just the beginning, many surprises can be expected
- More data and longer reads, in complement to other information, will help us to understand microbial communities, to improve
  - Environment, Food, Well-being, ..
  - ..and Health, through the clinical applications that we'll hear about during ICCMg2 !
- When will we be able to manipulate sustainably microbial communities?



Fasteris team

FASTERIS
member of the sonic healthcare network
SONIC HEALTHCARE



Diagnostics / FAMH
Frédéric GUERRY (AUR)
Marco BELFIORE (MCL)
Fabien MURISIER (FER)

Sanger Sequencing
Anne SCHOENDORF
Chloé HOT
Christophe BUSER

NGS Production
Cécile DELUEN SAGNE
Christelle BARRAS
Elisabeth DIETERLE
Cristel BUSCA
Franz PELEGRIN
Sonia PERREIRA
Vianney FRIGARD
Fanny JOLY

Development
Marta COTADO
Pauline CHARRUAU

Research
Nadine VINCENT
Magne OSTERAS

Bioinformatics & IT
Loïc BAERLOCHER
Gérald BLANC
Manuele CASTELNUOVO
Nicolas GONZALEZ
Patricia OTTEN

Sales & Marketing
Axel STRITTMATTER
François ALLER
Svetlana SKARUPELOVA
Aneliya ETROPLOSKA

Quality
Gilles MATTON
Benoît WINIGER

Administration
Laurent FARINELLI
Isabelle LOROLE

Founders
Laurent FARINELLI
Magne OSTERAS