



Plants
HD Genotyping and Sequencing

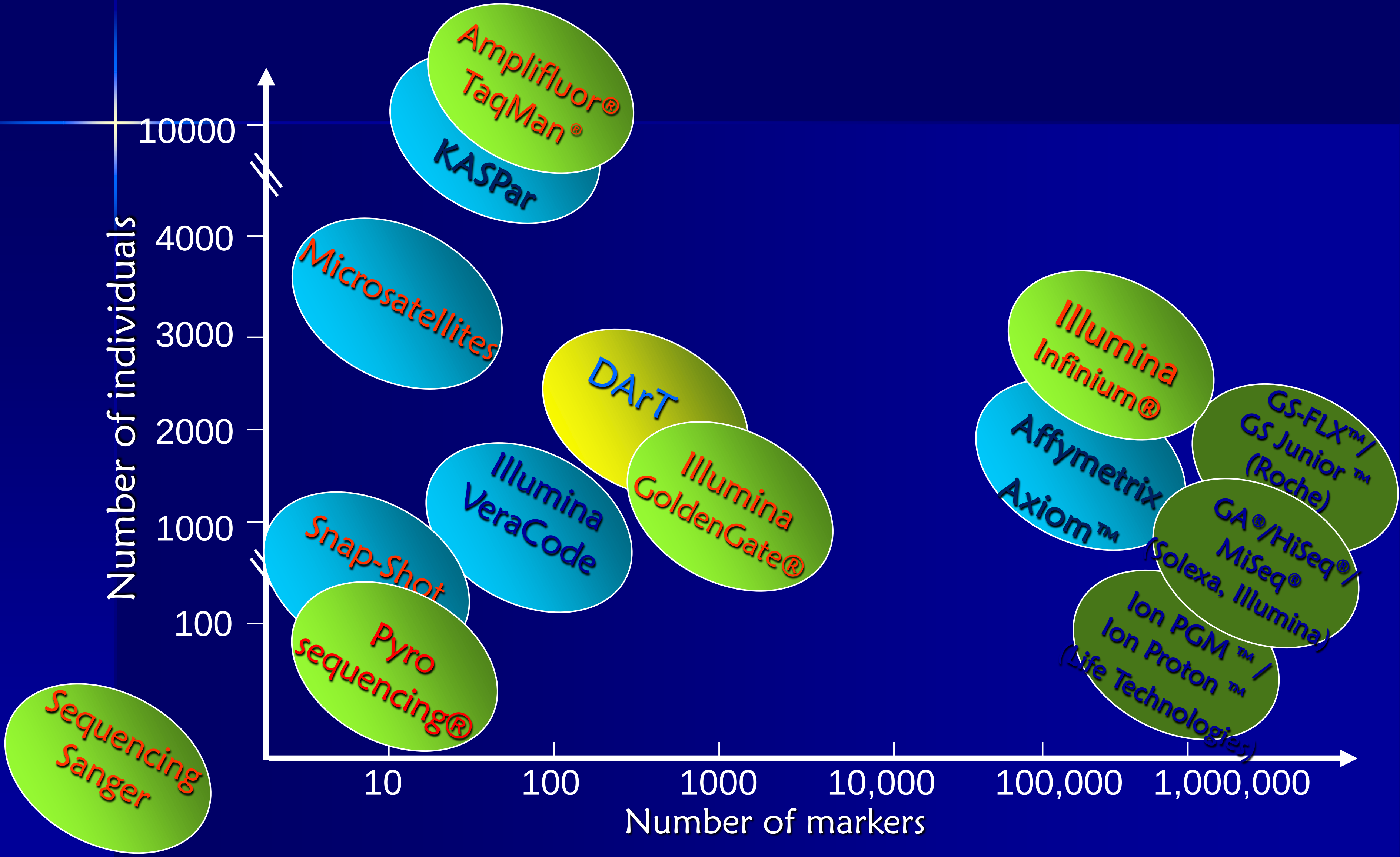
Which for what !





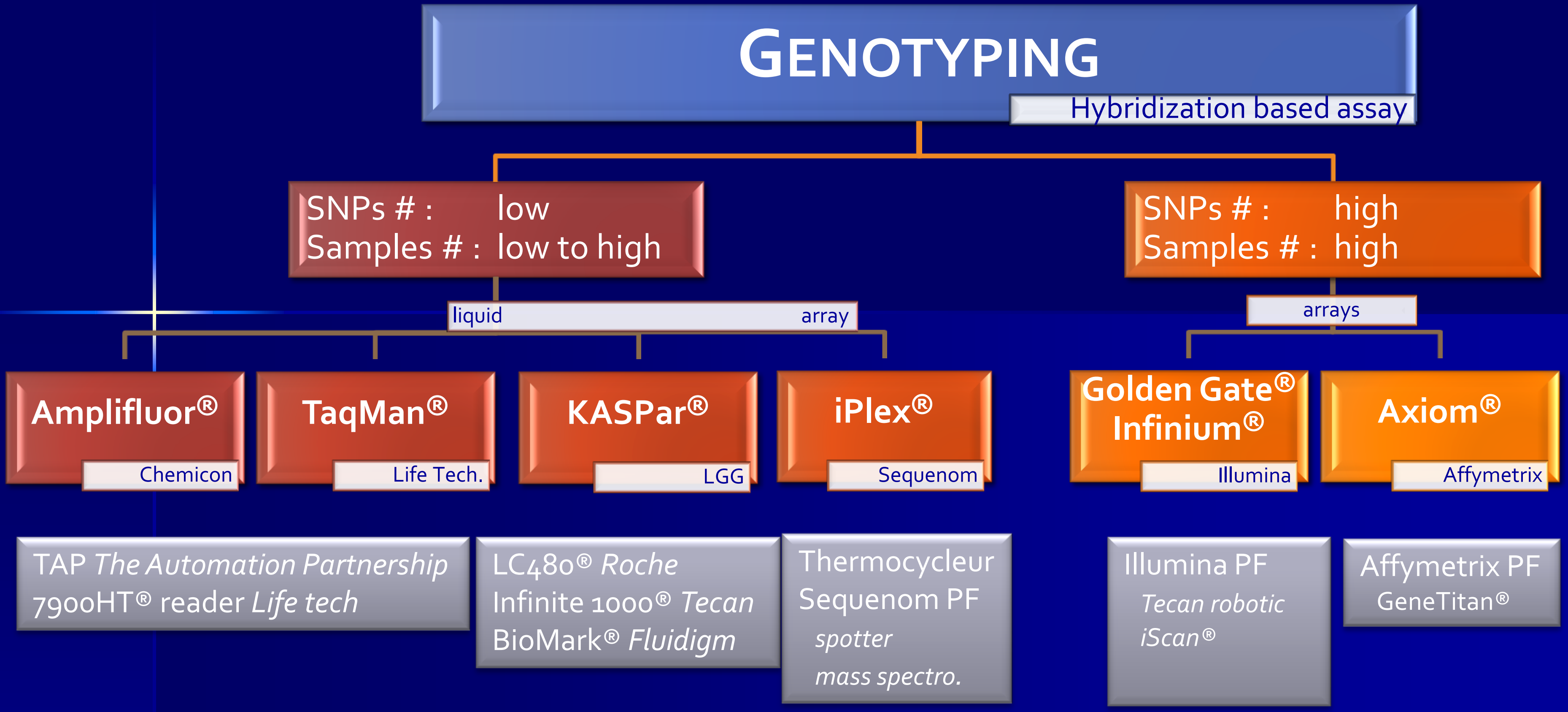
Marie-Christine Le Paslier, INRA_EPGV

Genotyping & Sequencing technologies 2013-2014



GENOTYPING

Hybridization based assay



SNPs # : low
Samples # : low to high

liquid array

Amplifluor® Chemicon

TaqMan® Life Tech.

KASPar® LGG

iPlex® Sequenom

TAP *The Automation Partnership*
7900HT® reader *Life tech*

LC480® *Roche*
Infinite 1000® *Tecan*
BioMark® *Fluidigm*

Thermocycleur
Sequenom PF
spotter
mass spectro.

SNPs # : high
Samples # : high

arrays

Golden Gate®
Infinium® Illumina

Axiom® Affymetrix

Illumina PF
Tecan robotic
iScan®

Affymetrix PF
GeneTitan®

		☑		GENTYANE PLATE-FORME de GENOTYPAGE et SÉQUENÇAGE en AUVERGNE		☑
☑	☑		☑	INRA BAP EPGV US 1279	☑	

GENOTYPING

Hybridization based assay

SNPs # : low
Samples # : low to high

liquid array

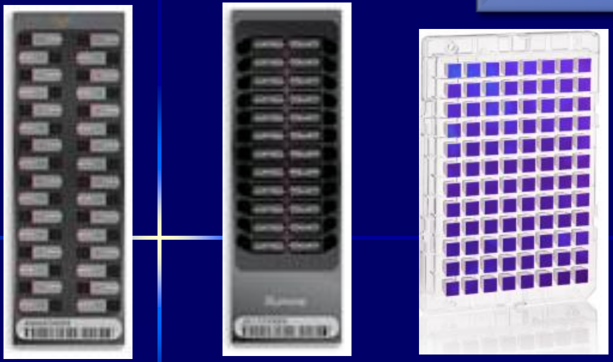
Amplifluor® Chemicon	**TaqMan®** Life Tech.	**KASPar®** LGG	**iPlex®** Sequenom	
*ASO – LSO_PCR *Hairpin UniPrimer™ fluorescein	*ASO-LSO_PCR *5' exonuclease activity of PCR-DNApol	*ASO-LSO _PCR *improved Amplifluor®	*LSO _PCR *ASO-SBE *Mass spectometry	allelic discrimination method
non	non 2014-> duplex	non	~ 32	multiplexing capacity
☺ max 0,5 € depending on samples nb	☺ ↗ 0,38 to 1,3 € depending on samples nb	☺ 0,08 to 0,16 € Biomark96:16€/96 SNPs	☺ max 0,24 max 7,2 €	data point cost (€) sample cost for X SNPs
☺	☺	☺ ☹ ↘ Biomark (48*48/96*96)	☺	flexibility
96 to N 10^3	96 to N 10^3	48-96 to N 10^3	96 to N 10^3	samples throughput

GENOTYPING

Hybridization based assay

SNPs # : high
Samples # : high

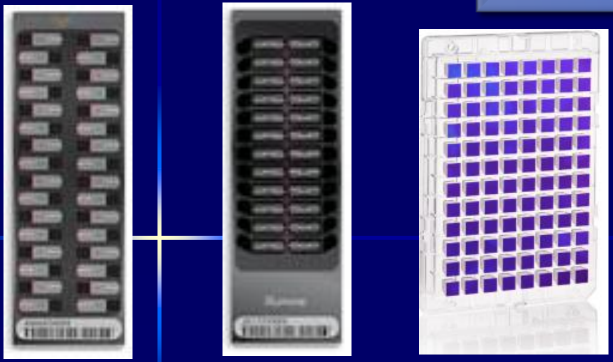
arrays



	Golden Gate®	Infinium®	Axiom®
	Illumina	Illumina	Affymetrix
allelic discrimination method	ASO/LSO hybridization ASO SBE -> LSO Ligation PCR Tag labeled	ASO hybridization ASO SBE Labeled nucleotides	ASO hybridization Labeled probes
Samples format Samples 1st order	**32** 480	**HD 24** /12 /4 Minimum in the sample range	**96/384** 480
multiplexing capacity	**96 - 3072**	**3 072 - 90 000** 90 000 – 250 000 250 000 – 1M	**1 500 - 675 000**
flexibility	☹	☹ ☺	☺
sample cost (€) data point cost (cts €)	☹ 55 (960 sples-1536 SNPs) 3,5	☺☺	☺☺☺
coût initial (K€)	☹ ↗	☹☹ ↗↗	☹ ↗

GENOTYPING

Hybridization based assay



SNPs # : high
Samples # : high

arrays

Infinium®

Illumina

Axiom®

Affymetrix

SNPs #	3-6 K	9-12K	12-15K	15-20K	15-20K consortium 2012	50K maïze 2010	1,5K	675 K
Samples #	1152/6048	1152/3024/6048	1152/6048	1152/3024/6048	48 to N	48 to N	480	480
coût initial *(K€)	53/239	61/154/284	73/320	87,5/213/330			37,5	85,5
sample cost *(€)	46/39,5	94 (2012) 53/51/47	79 (2013) 63/53	118 (2011) 76/70,5/54,5	51	75,5	50+28	150+28
data point cost *(cts €) for upper limit	0,77/0,66	0,78 0,44/0,42/0,39	0,52 0,42/0,35	0,6 0,38/0,58/0,45	0,42	0,15	5,2	0,026

*if design =100%_true >=80%

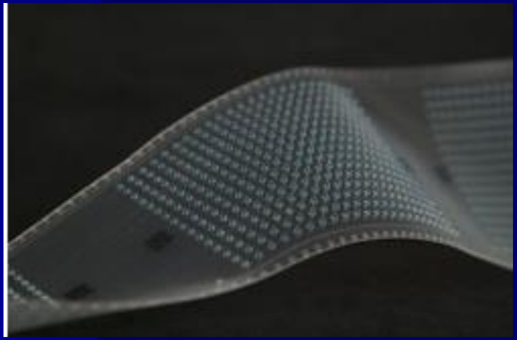
End of services or no more support

- Veracode : Golden Gate on BeadXpress
- Other ?

Out of BAP INRA and News

- SNPline service/LGC Genomics
- OpenArray/Life Tech
- Ultra High Throughput/Douglas Scientific

- *our revolutionary consumable, Array-Tape™*: low cost, flexible microplate
  - < 800nL/well



  - Soellex® simultaneously thermocycles up to 600 arrays / 150 block heaters 4 plates!!!!

Take-home Messages

- Genotyping / hybrization -> specificity
- Technology and Analysis constraints
  - DNA quality
  - Time consuming/Friendly analysis
- Synthesis success, ordering/re-order and delivering constrainsts
- Costs
  - Oligos
    - Librairies
    - Arrays production

-> the critical factor is samples number

- More is better!
- But need genotyping capacity

- Scientific projects to gather
  - Consortium for a specie project
  - Multi-projects?

SEQUENCING

NGS

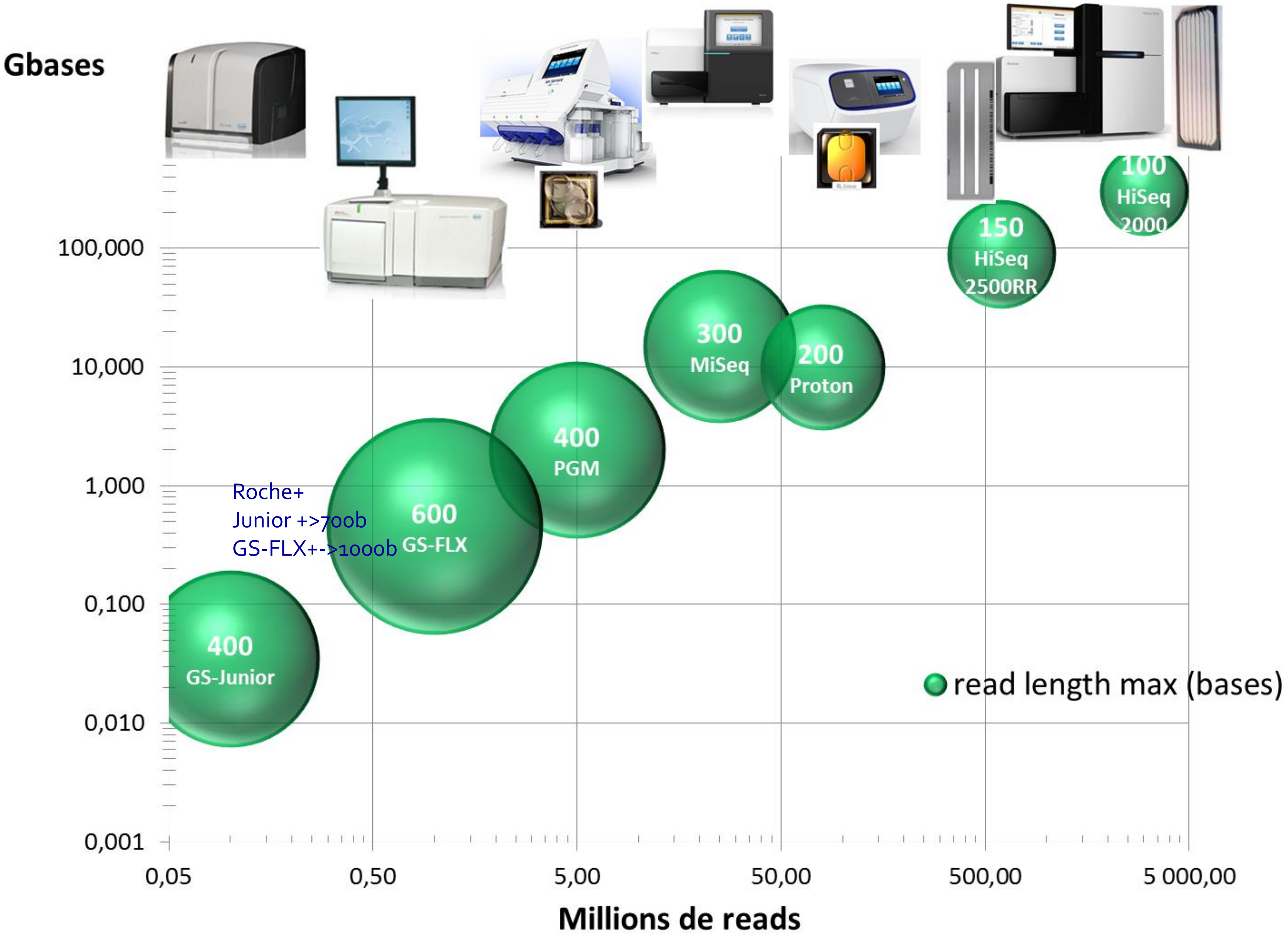
- Benchtop instrument
  - Low to Medium Throughput
    - **GS Junior®** Roche
    - **PGM®** Ion Torrent/Life Tech.
    - **PROTON®** Ion Torrent/Life Tech
    - **MiSeq®** Sequenom
- Center facilities
  - Medium to High Throughput
    - **GS-Flex®** Roche
    - **HiSeq®** Illumina

SR 400	SR 100 to 400	SR 100 to 200	PE 25 to 300	**read length**	SR 400 to 1000	PE 25 to 150

85 €	**lib**		85 €
1300 €/PE300 800 € /PE150	**run**		1300 €/PE100 3000 €/PE150

			☑	GENTYANE		
☑	☑	☑	☑	EPGV	☑	☑

NGS Sequencer : Gbases/Millions de reads



Roche+
Junior +>700b
GS-FLX+->1000b

Partners applications

- Whole genome re-sequencing
- GBS-Cornell
- Amplicon sequencing
  - Amplicons librairies performed on 48*48 Acess-Array® (Fluidigm)
- Bacterial or viral sequencing

End of services or no more support

- Roche for GS-Junior
- Other ?

Out of BAP INRA and News

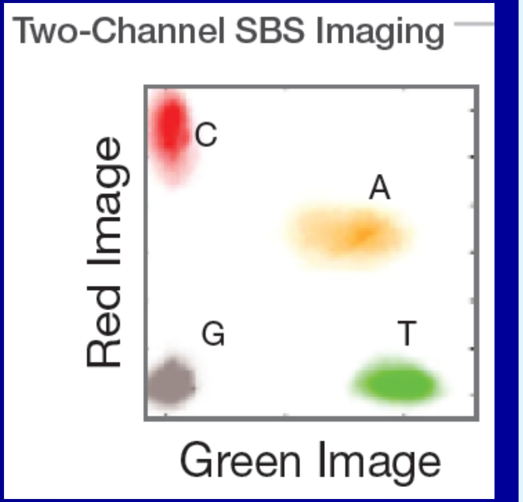
ILLUMINA

- **NextSeq 500 Sequencing System**

“The First High-Throughput Desktop Sequencer”… so exiting!

  - Flexibility
  - Evolution of SBS chemistry : 2 channels
  - Reducing cycle and data processing

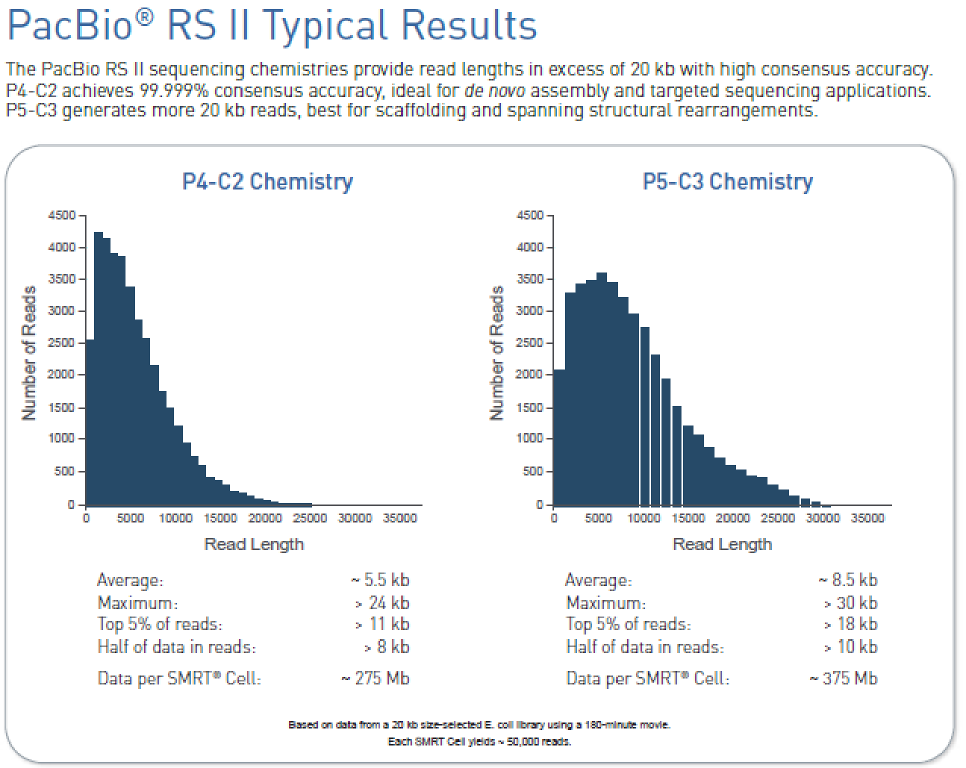
FC config NextSeq 500	**Read Length**	**Output *Gb***	**Run Time *hours***
High output *Up to 800M PE reads*	2*150	120	29
	2*75	50-60	18
	1*75	25-30	11h
Mid output *Up to 260 M PE reads*	2*150	32-39	26
	2*75	16-19	15h



- **HiSeq Xten** *The First $1000 Genome*
  - 18,000 genomes per year
  - Each HiSeq X->
    - 5 whole human genomes (30X) /day
    - or generate 1.8 terabases of sequence /3 da
  - New “patterned”FC w billions of ordonned wells

Out of BAP INRA and News

- SMRT Sequencing w PACBIO RSII / Pacific Biosciences
  - *The PacBio RS II finishes microbial genomes and improves assembly of larger organisms with multi-kilobase reads and unbiased coverage regardless of GC content. No amplification is required*.



PacBio® RS II Typical Results

The PacBio RS II sequencing chemistries provide read lengths in excess of 20 kb with high consensus accuracy. P4-C2 achieves 99.999% consensus accuracy, ideal for *de novo* assembly and targeted sequencing applications. P5-C3 generates more 20 kb reads, best for scaffolding and spanning structural rearrangements.

	P4-C2 Chemistry	P5-C3 Chemistry
Average:	~ 5.5 kb	~ 8.5 kb
Maximum:	> 24 kb	> 30 kb
Top 5% of reads:	> 11 kb	> 18 kb
Half of data in reads:	> 8 kb	> 10 kb
Data per SMRT® Cell:	~ 275 Mb	~ 375 Mb

Based on data from a 20 kb size-selected E. coli library using a 180-minute movie.
Each SMRT Cell yields ~ 50,000 reads.

Out of BAP INRA and News

From Bela Hausmann tweet

- MiniIon™/Oxford Nanopore Technology
  - **Exonuclease reaction**. *Hundreds of nanopores/device*
  - *The MinION™ device is a small instrument that is compatible with consumable flow cells containing the proprietary sensor chip, Application-Specific Integrated Circuit (ASIC) and nanopores that are needed to perform a complete **single-molecule sensing experiment**. **Plugging directly into a laptop or desktop computer through a USB port**, it is a self-contained device **to deliver real-time experimental data**.*
  - **What read lengths does the system give?**
    The system processes the lengths of DNA that are presented to it rather than 'generating' a read length. It is designed to process very long read lengths; tens of kb in a single read consisting of both strands of a template molecule, using a hairpin sample prep.
  - **MinION Access programm** (MAP) **in progress**
    - CEA-IG

Take-home Messages

- Evolutive technology
- Technology constraints
- Rapid or long sequencing run
- Costs
    - Kits/Reagents
        - libraries construction
        - libraries QC
- Bio-informatic processing to be considered

- Thanks for your attention

EPGV : équipe Bio

Réalisation Technique & BioAnalyse

Gestion et management des Projets

Séquençage

- NGS DNA library production et QC
- Validation des données

Génotypage

- Production
- Validation des données
- Analyse

Accueil et Formation

EPGV : équipe Info/Bio-Info

Support informatique

Base de Données

Processing des données NGS

- QC, Triming, Mapping

BioAnalyse des données

- SNP détection, Assemblage

Stockage des données