Ce-M-M-
Research Center for Molecular Medicine
of the Austrian Academy of Sciences

MEDICAL UNIVERSITY
OF VIENNA

Ludwig Boltzmann Institute
Rare and Undiagnosed Diseases

ÖAW

# New Tools for Personalized Medicine

## *Tools = Assays, Devices, Software

**Christoph Bock**

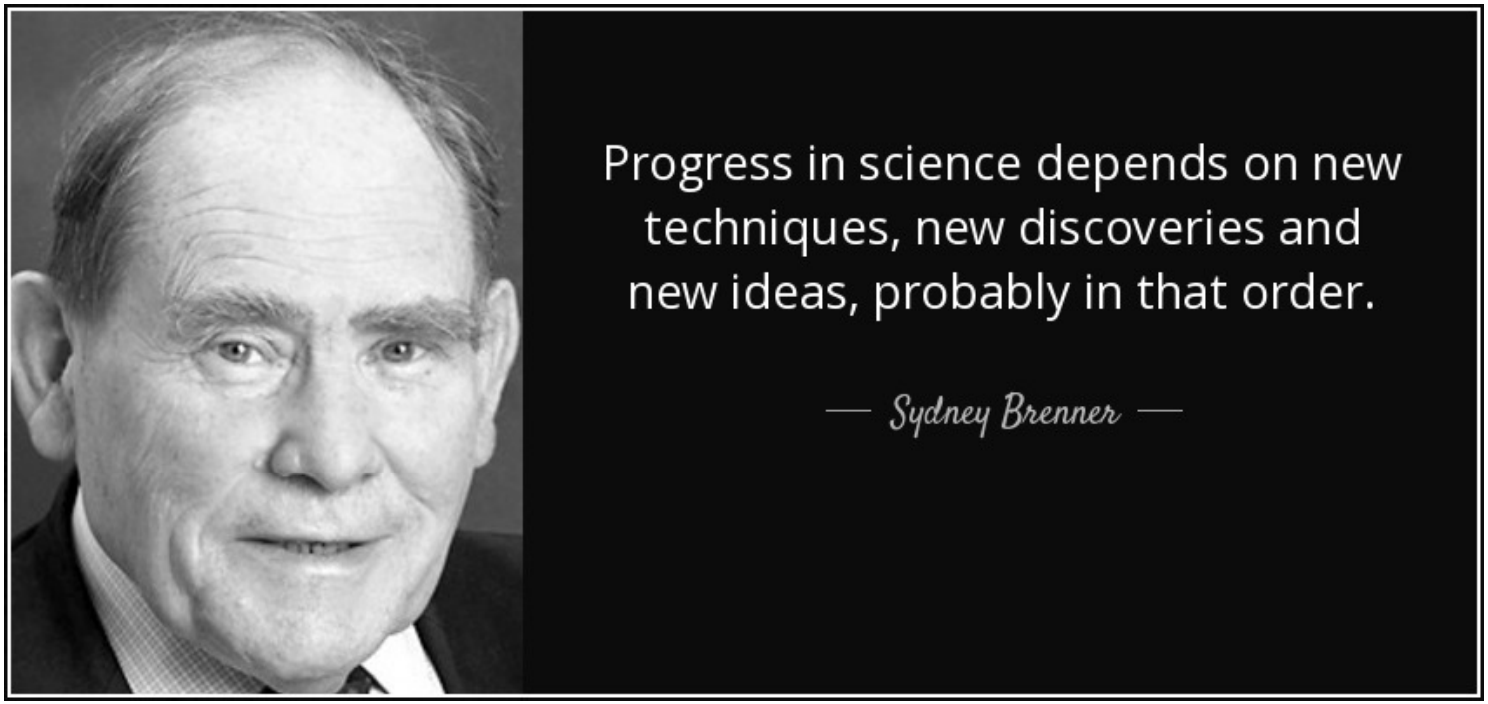
ICPerMed – First Research Workshop

Milano, 26 June 2017

http://epigenomics.cemm.oeaw.ac.at
http://biomedical-sequencing.at

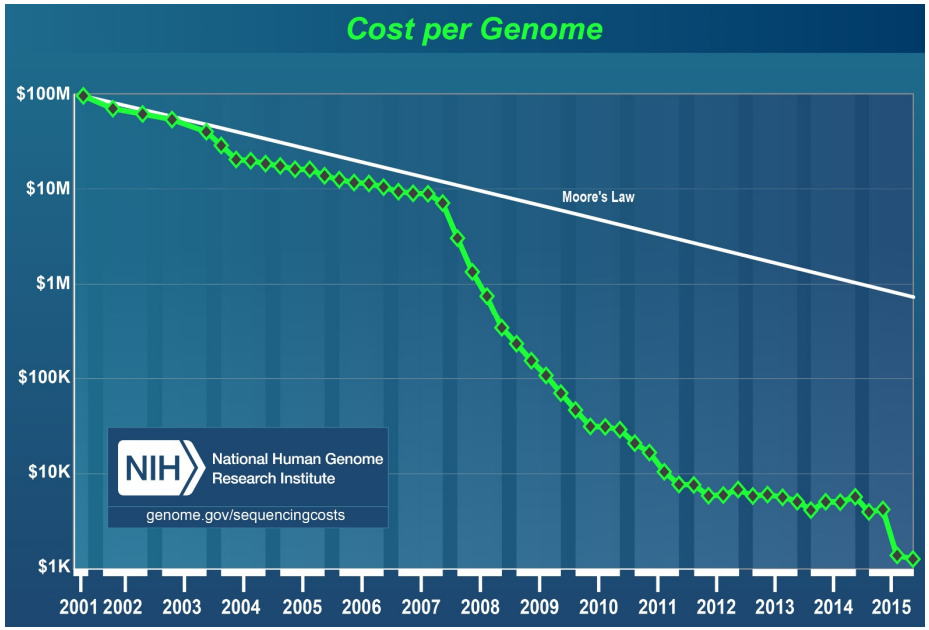
Research Laboratory
Sequencing Platform

# Development of new tools (techniques, technologies) has impact!




Sydney Brenner, 2002 Nobel Prize in Physiology or Medicine
http://www.ncbi.nlm.nih.gov/pmc/articles/PMC139404/

# Example 1: Next generation sequencing for genomic diagnostics


## Impact on Personalized Medicine

- **Cancer**: Disease stratification based on driver mutations
- **Rare diseases**: Most patients now receive a genetic diagnosis
- **Drugs**: Patient-specific prediction of efficacy and side effects

# Example 2: The CRISPR/Cas9 system for genome editing

RESEARCH ARTICLE

## A Programmable Dual-RNA–Guided DNA Endonuclease in Adaptive Bacterial Immunity

Martin Jinek,[1,2]* Krzysztof Chylinski,[3,4]* Ines Fonfara,[4] Michael Hauer,[2]†
Jennifer A. Doudna,[1,2,5,6]‡ Emmanuelle Charpentier[4]‡

Clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated (Cas) systems provide bacteria and archaea with adaptive immunity against viruses and plasmids by using CRISPR RNAs (crRNAs) to guide the silencing of invading nucleic acids. We show here that in a subset of these systems, the mature crRNA that is base-paired to trans-activating crRNA (tracrRNA) forms a two-RNA structure that directs the CRISPR-associated protein Cas9 to introduce double-stranded (ds) breaks in target DNA. At sites complementary to the crRNA-guide sequence, the Cas9 HNH nuclease domain cleaves the complementary strand, whereas the Cas9 RuvC-like domain cleaves the noncomplementary strand. The dual-tracrRNA:crRNA, when engineered as a single RNA chimera, also directs sequence-specific Cas9 dsDNA cleavage. Our study reveals a family of endonucleases that use dual-RNAs for site-specific DNA cleavage and highlights the potential to exploit the system for RNA-programmable genome editing.

Emmanuelle Charpentier

Jennifer A. Doudna

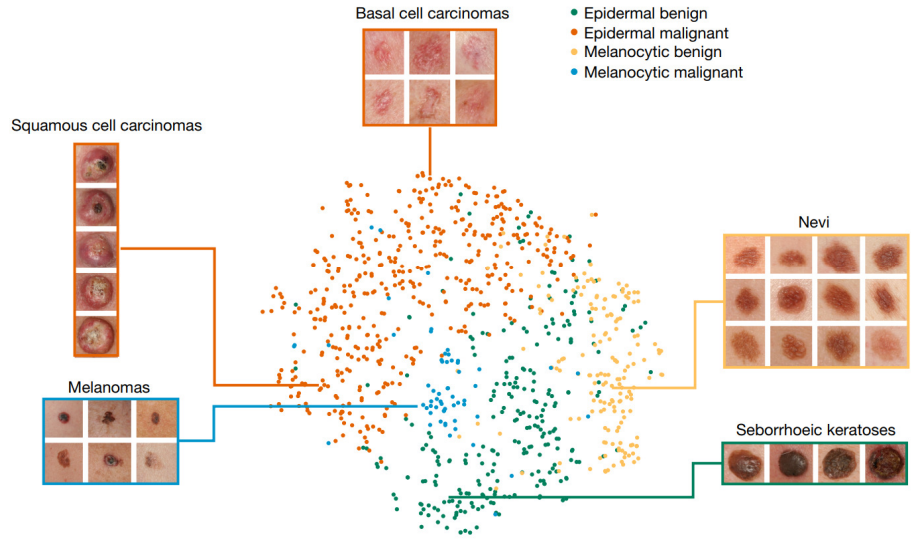
## Impact on Personalized Medicine

- **Biomedical research**: Faster target discovery and validation
- **Somatic gene therapy**: Better control and (hopefully) lower cost
- **Regenerative medicine**: Tissue engineering for transplantation

# Example 3: Machine learning makes expert knowledge scalable


## Dermatologist-level classification of skin cancer with deep neural networks

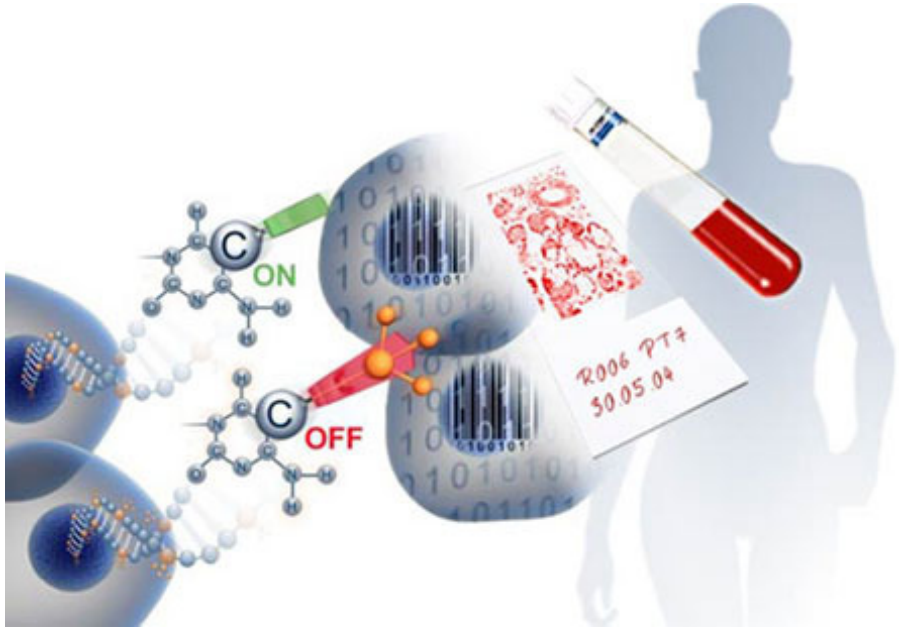
Andre Esteva[1]*, Brett Kuprel[1]*, Roberto A. Novoa[2,3], Justin Ko[2], Susan M. Swetter[2,4], Helen M. Blau[5] & Sebastian Thrun[6]




## Impact on Personalized Medicine

- **Computer vision**: Classify pictures in dermatology, radiology, etc.
- **Natural language processing**: Annotating free text documents
- **Data mining**: Identifying hidden patterns in large clinical datasets

# Example 4: Epigenetics as the interface to the environment




Quantitative comparison of DNA methylation assays for biomarker development and clinical applications

Christoph Bock[1–3], Florian Halbritter[1], Francisco J Carmona[4], Sascha Tierling[5], Paul Datlinger[1], Yassen Assenov[6], María Berdasco[4], Anke K Bergmann[7,8], Keith Booher[9], Florence Busato[10], Mihaela Campan[11], Christina Dahl[12], Christina M Dahmcke[12], Dinh Diep[13], Agustín F Fernández[14–16], Clarissa Gerhauser[6], Andrea Haake[7], Katharina Heilmann[6], Thomas Holcomb[17], Dianna Hussmann[18], Mitsuteru Ito[19], Ruth Kläver[20], Martin Kreutz[20], Marta Kulis[21], Virginia Lopez[14,15,16], Shalima S Nair[22,23], Dirk S Paul[24], Nongluk Plongthongkum[13], Wenjia Qu[22], Ana C Queirós[21], Frank Reinicke[20], Guido Sauter[25], Thorsten Schlomm[25], Aaron Statham[22], Clare Stirzaker[22,23], Ruslan Strogantsev[19], Rocío G Urdinguio[14,15,16], Kimberly Walter[17], Dieter Weichenhan[6], Daniel J Weisenberger[11], Stephan Beck[24], Susan J Clark[22,23], Manel Esteller[4,26,27], Anne C Ferguson-Smith[19], Mario F Fraga[14–16], Per Guldberg[12], Lise Lotte Hansen[18], Peter W Laird[11,28], José I Martín-Subero[21], Anders O H Nygren[29], Ralf Peist[20], Christoph Plass[6], David S Shames[17], Reiner Siebert[7,30], Xueguang Sun[9], Jörg Tost[10], Jörn Walter[5] & Kun Zhang[13]for the BLUEPRINT consortium

## Impact on Personalized Medicine

- **Risk prediction**: Epigenetic memory of environmental exposures
- **Liquid biopsy**: Determining the cell-of-origin of circulating tumor DNA
- **Treatment monitoring**: Measuring the effect of epigenetic drugs

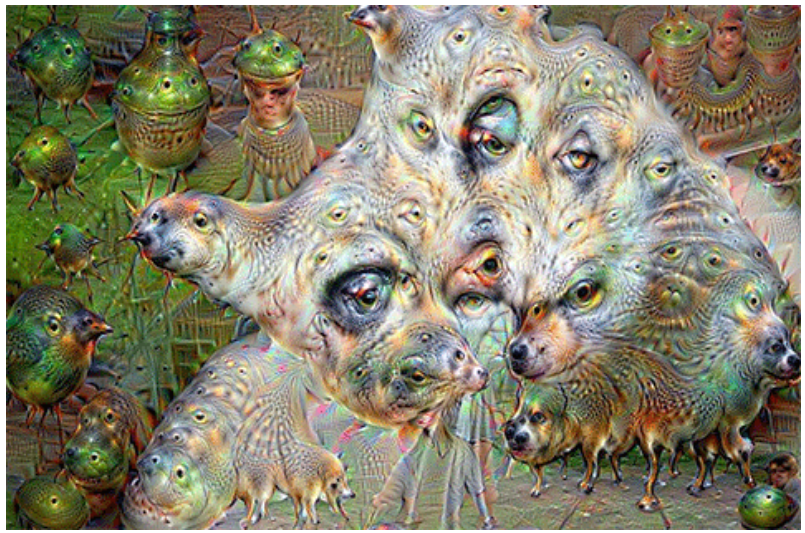

http://dx.doi.org/10.1038/nbt.3605

# The power of modularity

## Computer science is all about building reusable tools (algorithms/software)

| TCP/IP model | Protocols and services | OSI model |
|---|---|---|
| Application | HTTP, FTTP, Telnet, NTP, DHCP, PING | Application |
| | | Presentation |
| | | Session |
| Transport | TCP, UDP | Transport |
| Network | IP, ARP, ICMP, IGMP | Network |
| Network Interface | Ethernet | Data Link |
| | | Physical |


## Much of the creativity in IT comes from smart combinations of such tools

# Technological progress can be fast

## High-performance computing

1979

today

**Who has a computer?**

- 1960s: Major research institutes
- 1970s: University departments
- 1980s: Companies and schools
- 2017: Almost everybody & always

## Genome sequencing


2006

today

**Whose genome has been sequenced?**

- 1996: First bacterium (E. coli)
- 2001: Human reference genome
- 2007: First personal genomes
- 2017: Many thousand personal genomes

## GOAL

To map the tools that will contribute to personalized medicine

To make concrete recommendations on tool research, development, and implementation

# Research workshop: Goals and anticipated outcomes

## Mapping emerging tools with major impact on personalized medicine

- **Time dimension**: Predicting realistic timescales, identifying interdependencies
- **Geographical dimension**: Defining the context for research/implementation
- **Systems effects**: Anticipating change to the personalized medicine ecosystem

## Concrete recommendations for tool-driven research in personalized medicine

- Example 1: Which tools to prioritize in upcoming ERA-NET etc. calls?
- Example 2: Best practices for national personalized medicine initiatives
- Example 3: A checklist for planning personalized medicine infrastructure

# Research workshop: The time dimension

**By which time will a tool start having major impact for personalized medicine?**

**NOW** →
- How to maximize its productive use and patient impact?
- How to monitor and improve cost effectiveness?

**1-5 years** →
- How to effectively integrate research and development?
- How to create a viable ecosystem for the emerging tool?

**5-10 years** →
- How to prioritize the various areas of promising research?
- How to create critical mass without losing out on diversity?

# Research workshop: The geographical dimension

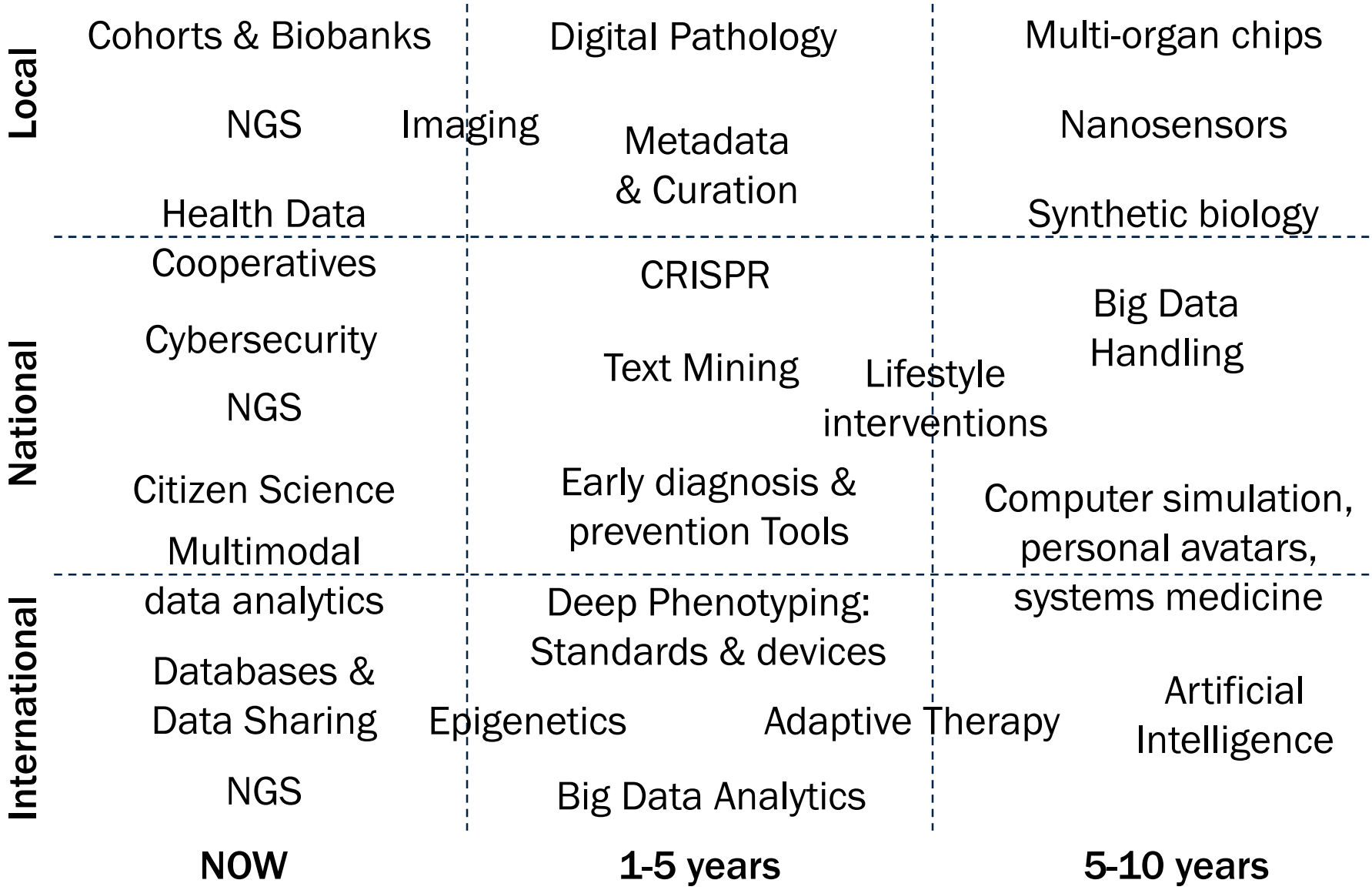
**What is the appropriate geographical level to study/implement a given tool?**

**Local** →
- How to create critical mass and avoid duplication?
- How to maximize synergy and collaboration?

**National** →
- How to coordinate all relevant national stakeholders?
- How to reach adequate visibility among policy makers?

**International** →
- How to connect and coordinate very diverse partners?
- How to balance speed, quality, and inclusion?

# Tools for personalized medicine: When, where, and how



# Lead questions for the “New Tools Impact” working group

- Are there already best practice examples for new tools in personalized medicine existing?

-> rare diseases (IRDiRC, European Reference Networks), cancer (TCGA, ICGC -> impact of data sharing; MAPPs: http://efpiamapps.eu/), genomic medicine (Genomics England), hepatitis C in Spain (40k patients in 2 years, mandatory genotyping, driven by patient pressure), INCa breast cancer screening

- What are the major lessons learned so far?

-> implementing personalized medicine approach is almost always complex (in part due to complexities of the healthcare system), political commitment is a major success factor, joint production of data and standards by international consortia, need to integrate diverse stakeholders, need for standardization of clinical protocols, rapid development of tools requires fast and flexible regulatory policy, (some of) the tools are there – we need to use them in better/smarter ways for clinical impact, we learnt a lot of (disease) biology on the way, actionability problem: diagnosis doesn’t always mean therapy, bioinformatics has become the single biggest bottleneck

# Lead questions for the “New Tools Impact” working group

- Which are the crucial inputs by e.g. medical informatics and ICT so far and in the future?

-> Medical informatics, bioinformatics, and ICT provide the enabler and “glue” between data production, data analysis, medical decisions, etc.; ICT needs to be better integrated into European Reference Networks; basic science and technology development in bioinformatics, medical informatics, ICT, genomics, molecular biology, phenotyping & lifestyle profiling etc.

- How could research benefit from such tools?

-> Discovery of new biology; reality check for biological understanding, new technologies, etc.; new challenges for research and development, large-scale databases available for re-analysis and hypothesis generation/testing, resource for massive-scale data mining

# Lead questions for the “New Tools Impact” working group

- Which could be the best approaches to support health providers and the health system with new tools?

-> Access to epidemiological databases, monitoring tools for healthcare quality, disparities, etc. (e.g., implemented in the form of Health Data Cooperatives), facilitate pilot studies for personalized medicine, systematic incorporation of representative patient feedback (Responsible Research & Innovation tools, consensus conference, citizen forum, etc.)

- How can genomic markers for predicting antimicrobial resistance be identified, validated and implemented for routine analysis ?

-> Antimicrobial resistance is an important field of application for various tools developed to advance personalized medicine, this including next generation sequencing, personal microbiome, metagenomics, and metabolome profiling, machine learning, international data exchange, and economic modeling

- How is the validity of using subset of resistors on new diagnostics?

-> We did not understand the question

# Topics to develop into concrete recommendations

## Tools with direct relevance to personalized medicine

1. Biomarker-driven medicine: multi-omics, IT, validation, reproducibility, clinical utility
2. Genomics data interpretation, plus phenotypes
3. Artificial Intelligence, Machine Learning, Simulation (Personal Avatar)
4. Citizen Science, Biobanks, Health Data Cooperatives
5. European infrastructures for personalized medicine (e.g. open science cloud)

## Additional topics

- Education and communication for healthcare workers and citizens/patients
- Economic modeling & cost effectiveness research

# Recommendations (page 1)

## Biomarker-driven medicine: multi-omics, IT, reproducibility, clinical utility

- Move beyond single-gene biomarkers -> multi-modal, network type assays / biomarkers to increase robustness and capture wider disease-relevant biology
- Need for more and smarter replication, model-based selection of biomarkers for validation (but avoid to get locked into outdated, substandard assay technology)
- Better connect technology development, data analytics, and clinical validation
- Make biomarker research future-proof by collecting cohorts that can be re-used as resource for future biomarker studies (e.g. as validation or control cohort)
- Emerging dimension: dense timelines, n-of-one studies/trials, personal utility
- Interaction with regulatory bodies on suitable standards and procedures

# Recommendations (page 2)

## Genomics data interpretation (including phenotypes)

- Create an infrastructure and political commitment to make sure that all information needed for a robust genetic diagnosis are in the public domain (avoid privatization of e.g. allele frequency information)
- Standardization of phenotype information across borders and language barriers with tool-supported ontologies (HPO etc.)
- Contribute to implementing data sharing in line with recommendations of the Global Alliance for Genomics and Health (strengthening Europe's representation)
- We need high-throughput tools for connecting genotype to (cellular) phenotype / molecular biological functions
- Better connect clinical genetics and molecular biological studies

# Recommendations (page 3)

## Artificial Intelligence, Machine Learning, Simulation (Personal Avatar)

- Many loosely related fields contribute in complementary ways: Machine learning, medical statistics, computer vision, network medicine, Bayesian statistics, etc.
- Collect and aggregate massive datasets in a way that makes them accessible to computational analysis
  (ethical, legal, social, economic, policy, lifestyle, competitive, technical etc. limitations)
- Organize large-scale continuous validation/replication/benchmarking initiatives
- 'Reproducible Research' -> re-runnable analyses with all relevant data / code / model parameters available to others in the field (open source, open data, etc.)
- New / better methods for multi-scale modeling (molecule – cell – organ – patient)
- Training, education, attracting talent to overcome the bioinformatics bottleneck, career perspectives for bioinformatics
- Algorithm provider accountability / review committees that monitor the ethical and social dimension of artificial intelligence

# Recommendations (page 4)

## Citizen Science, Biobanks, Health Data Cooperatives

- Pilot projects that seek to combine aspects of biobanking, citizen science, epidemiology, health data cooperatives (high citizen/patient involvement)
- New ways of obtaining and updating consent: e-consent, mobile devices, broad consent vs. dynamic consent (reconnecting on an as-needed basis)
- Connecting digital and social innovation with healthcare to create broader citizen engagement
- Monitoring the incentive structures of citizens and other stakeholders

# Recommendations (page 5)

## European infrastructures for personalized medicine (e.g. open science cloud)

- Robust, reproducible, scalable, validated pipelines for data processing in the clinic
- Easily accessible, connected databases with suitable governance models (as open as possible, while accounting for patient privacy etc.)
- Easy-to-use visualization, exploration, and analysis tools accessible to non-bioinformaticians
- Put European supercomputing infrastructures and initiatives at the service of life sciences research (and clinical applications)
- International standardization and shared infrastructure for technologies that are going to be the backbone of personalized medicine (NGS, omics, imaging, etc.)
- Standardization and integration of clinical, social data (repositories, ontologies, etc.)
- Connecting to ongoing developments of the Internet, mobile infrastructures, European open science cloud etc.

# Recommendations (page 6)

## Economic modeling & cost effectiveness research

- Innovate methods for economic modeling that are tailored to the specific requirements of personalized medicine
- Augment cost effectiveness research with emerging methods such as behavioral economics, game theory etc.

## Education and communication for healthcare workers and citizens/patients

- Emphasize tool-related skills in the education of healthcare workers (similar to computer literacy)
- Integrate IT and data science into the education of all researchers and clinicians
- Create ‘genetic literacy’ in the general population (many successful pilot studies, ready for broad, coordinated implementation)

# Topics to develop into concrete recommendations

## Tools with direct relevance to personalized medicine

1. Biomarker-driven medicine: multi-omics, IT, validation, reproducibility, clinical utility
2. Genomics data interpretation, plus phenotypes
3. Artificial Intelligence, Machine Learning, Simulation (Personal Avatar)
4. Citizen Science, Biobanks, Health Data Cooperatives
5. European infrastructures for personalized medicine (e.g. open science cloud)

## Additional topics

- Education and communication for healthcare workers and citizens/patients
- Economic modeling & cost effectiveness research