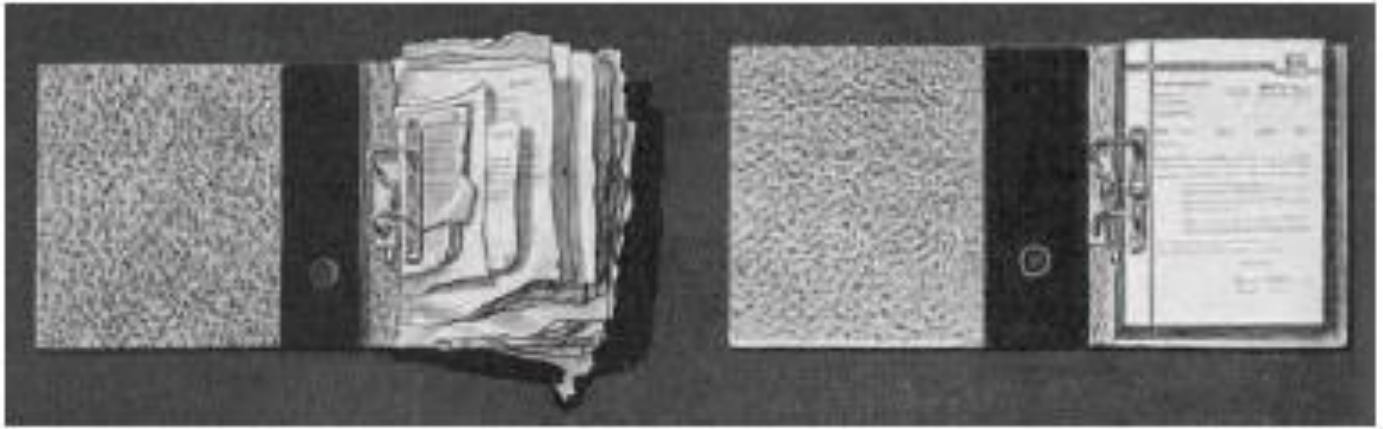
Standardization in Next-Generation Sequencing

Combined CHARME, EMBnet and NETTAB Workshop

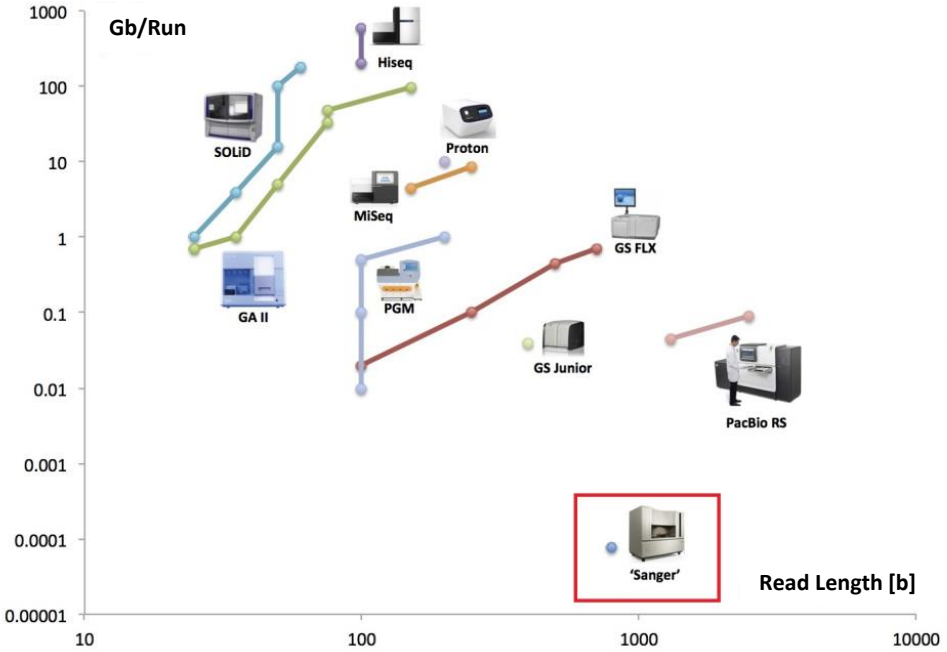
Christoph Endrullat
cen@katece.de

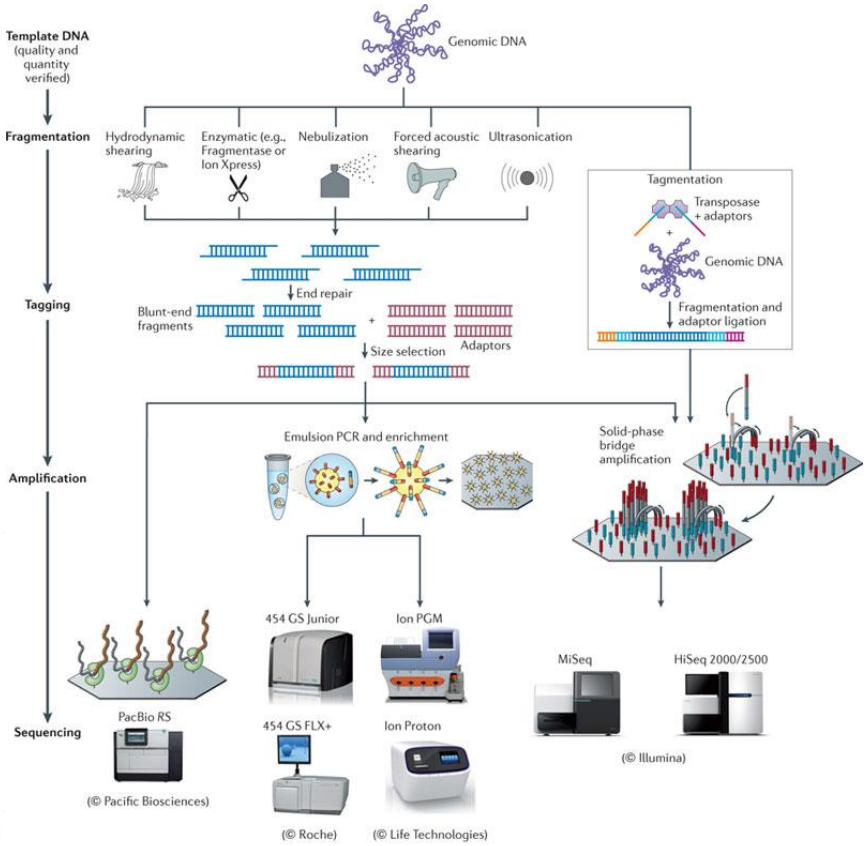
25 October 2016

The Cost Action CA15110 is supported by the EU Framework Programme H2020



NGS Technologies

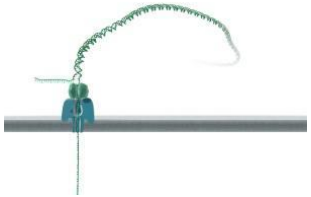


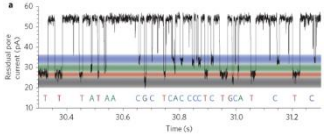


NGS Technologies

IBM

GenapSys





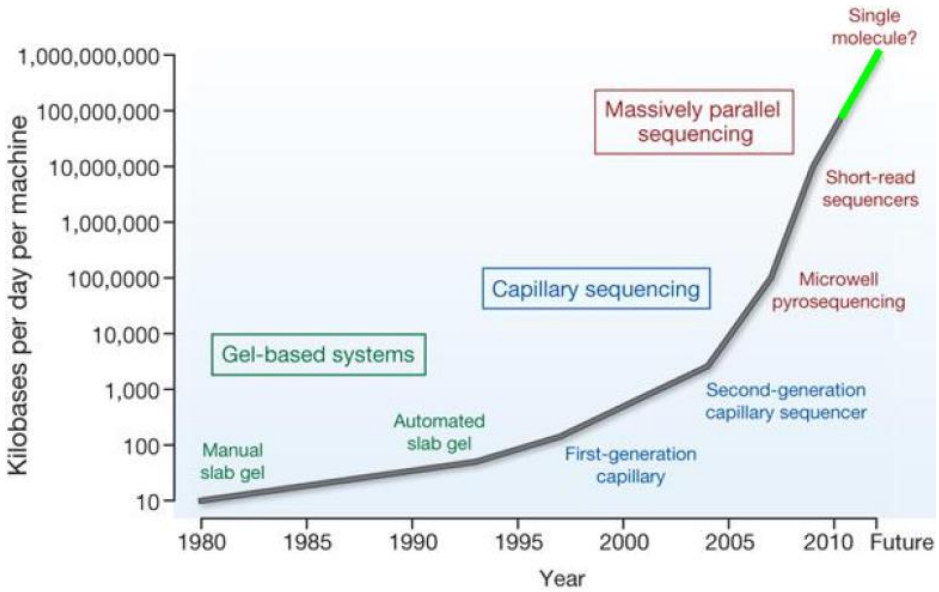
Nabsys

genia

gnubio

Problems in NGS Standardization

- Fast development of new platforms
- Simultaneously obsolescence of former systems
  - 454 and SOLiD: No development of new systems and ceased support
  - Expensive workflows and decreased cost of competitor platforms led to economic pressure



Problems in NGS Standardization

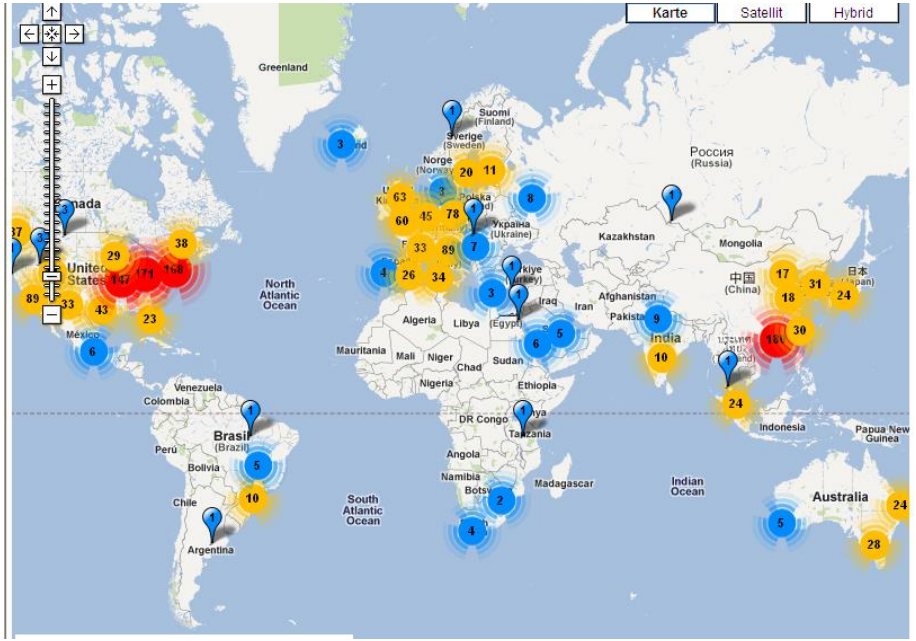
- Big market established in relatively short time
  - Sequencing service providers
  - Sequencing platform manufacturers
- Market leaders like Illumina are not willing to be forced by a general standard
  - 71 % market share
  - 90 % of all generated sequences by Illumina
  - Want to establish own standards
  - Considerable standardization efforts by Illumina

Problems in NGS Standardization

- Validity of standards across NGS applications
  - Biggest effort in clinical diagnostics
  - Oncology/precision medicine as a promising field
  - Same standards are not applicable or reasonable in other applications and vice versa

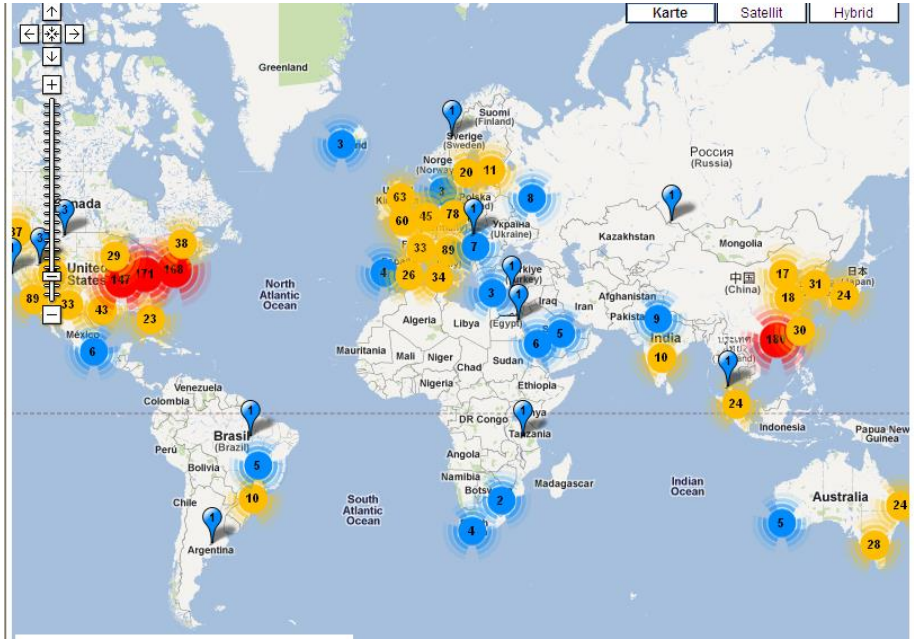
Problems in NGS Standardization

- Distribution of NGS systems worldwide
- Highest amount in industrial countries



Problems in NGS Standardization

- Distribution of NGS standardization efforts
  - NGS originated in US
  - Majority of approaches are located overseas
  - No verifiable or published approaches currently in other countries
  - Initiatives should be found and encouraged to participate in the field (CEN)



Standardization Efforts

- NGS guidelines for somatic genetic variant detection by the New York State Department of Health – CLIA guidelines
  - Validation parameters/CLIA performance characteristics
    - Accuracy, robustness, precision, repeatability, reproducibility, analytical sensitivity and specificity, reportable and reference range
  - QC RM
    - NTC, - control, +/sensitivity control

Standardization Efforts

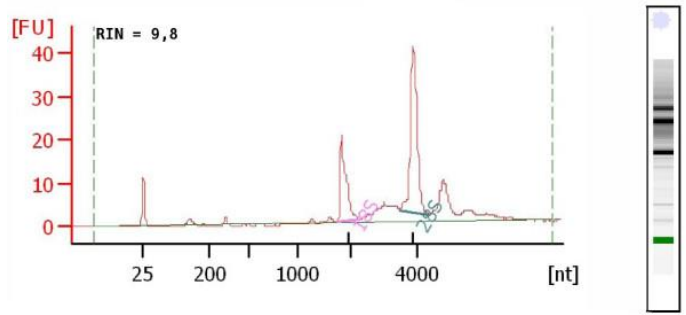
- NIST-GIAB
  - Investigation of reference data, methods and standards for NGS
  - Well-characterized whole human genomes as RM
  - Methods for use of RM for understanding NGS performance
- ABRF-NGS
  - Identification of optimal methods and strategies for NGS projects
  - Performance evaluation of different NGS platforms
- CAP-MOL
  - Standards for documentation, validation, QA, confirmatory testing, exception logs, monitoring of upgrades, variant interpretation and reporting, incidental findings, data storage, version traceability and data transfer confidentiality
- ERCC
  - Development of RNA spike-in controls

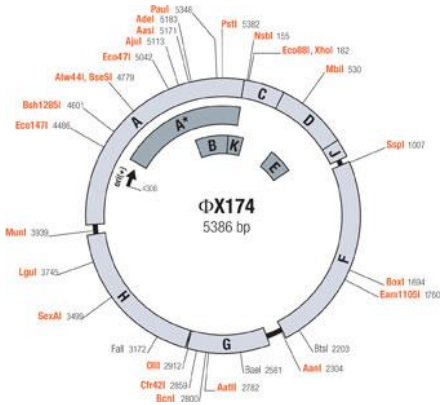
Standardization Efforts

- GSC-MIGS
  - Remedy lack of incomplete genome descriptions for data submission
  - Depth of coverage, overall quality, taxonomy, trophic level, propagation
- MAQC-III/SEQC and MAQC-IV/SEQC2
  - Evaluation of technical performance between different NGS platforms by establishing benchmarks with reference samples
- HGP-Bermuda Standards
  - Standards for sequence fidelity
  - Q40 or 99.99% accuracy (1 error/10,000 bp)
- Nex-StoCT
  - Developed principles, guidelines, standards as well as recommendations for the implementation of NGS into diagnostic laboratories
  - Novel NGS PT system
    - Enables error identification (FMEA)
    - Indication of QC problems
    - Verification of test performance in laboratories

Possible Standards and *de facto* Standards

Prepare sequencing library	Prepare and enrich template	Sequencing	Data analysis
(1) Control input Checkpoint – spectrophotometer, capillary gel electrophoresis (section 4.2., 4.3., 4.5.) (2) Fragmentation and end-repair Checkpoint - capillary gel electrophoresis (section 4.4.) (3) Adapter ligation and nick repair Checkpoint - capillary gel electrophoresis (section 4.4.) (4) Size selection Checkpoint - capillary gel electrophoresis (section 4.4.) (5) Library normalization / quantification Checkpoint - fluorometer, qPCR, ddPCR (section 4.4.)	(1) Prepare template (clonally amplified DNA on surface or beads) Checkpoint – fluorometer (section 4.4.) (2) Enrich template	(1) Create a run (2) Clean and initialize the sequencer (3) Start sequencing	(1) Data quality check and analysis Checkpoint – (section 5.)



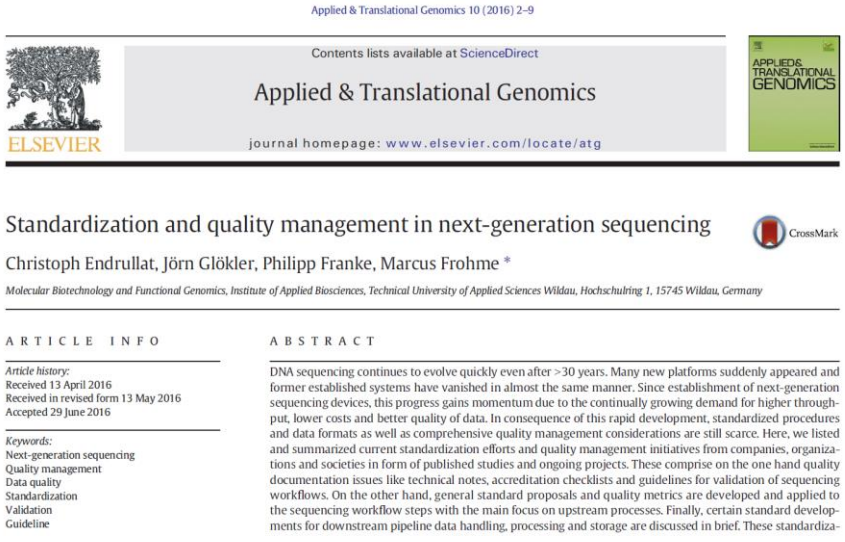


Possible Standards and *de facto* Standards

- Lack of documentation
- Development of procedure-, operating- and inspection instructions is required (SOP)
- Establishment of quality records respectively verification documents (quality proof of data)
- Quality records are suitable for providing quality certificates to customers

Remarks

- Enhancement of transparency, traceability and reproducibility of results
- Generation of reliable data
- Enhancement of financial and time efficiency
- Expansion of services (companies)
- Establishment of NGS into forensics and diagnostics, especially in consideration of emerging personalized medicine
- Acceleration of innovation process



Applied & Translational Genomics 10 (2016) 2–9

Contents lists available at ScienceDirect

Applied & Translational Genomics

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

Standardization and quality management in next-generation sequencing

Christoph Endrullat, Jörn Glökler, Philipp Franke, Marcus Frohme *

Molecular Biotechnology and Functional Genomics, Institute of Applied Biosciences, Technical University of Applied Sciences Wildau, Hochschulring 1, 15745 Wildau, Germany

ARTICLE INFO

Article history:
Received 13 April 2016
Received in revised form 13 May 2016
Accepted 29 June 2016

Keywords:
Next-generation sequencing
Quality management
Data quality
Standardization
Validation
Guideline

ABSTRACT

DNA sequencing continues to evolve quickly even after >30 years. Many new platforms suddenly appeared and former established systems have vanished in almost the same manner. Since establishment of next-generation sequencing devices, this progress gains momentum due to the continually growing demand for higher throughput, lower costs and better quality of data. In consequence of this rapid development, standardized procedures and data formats as well as comprehensive quality management considerations are still scarce. Here, we listed and summarized current standardization efforts and quality management initiatives from companies, organizations and societies in form of published studies and ongoing projects. These comprise on the one hand quality documentation issues like technical notes, accreditation checklists and guidelines for validation of sequencing workflows. On the other hand, general standard proposals and quality metrics are developed and applied to the sequencing workflow steps with the main focus on upstream processes. Finally, certain standard developments for downstream pipeline data handling, processing and storage are discussed in brief. These standardiza-

DIN

ISO

Acknowledgements

Technische Hochschule Wildau [FH] Technical University of Applied Sciences

molecular biotechnology functional genomics

cost EUROPEAN COOPERATION IN SCIENCE AND TECHNOLOGY

CHARME

The Cost Action CA15110 is supported by the EU Framework Programme H2020

Thanks for attention!