

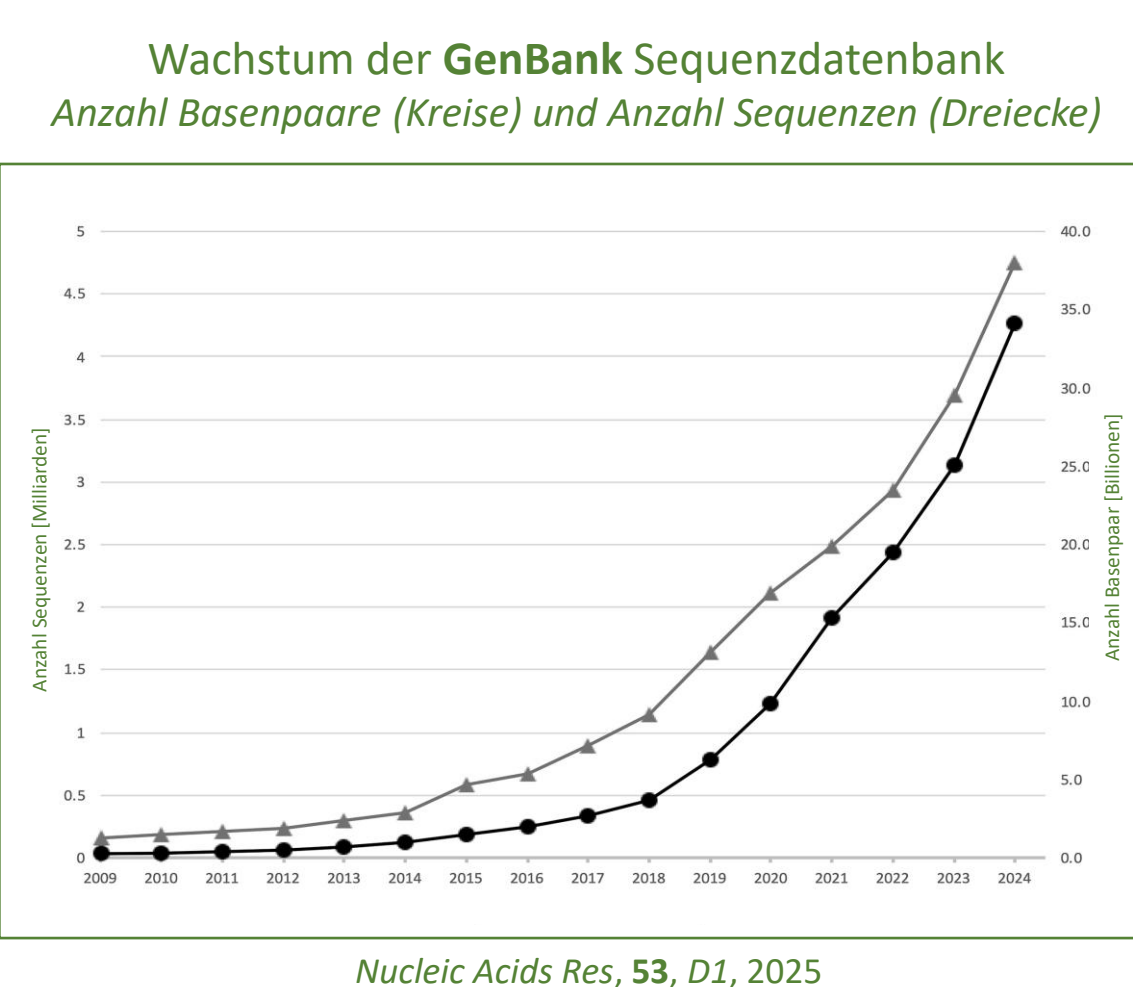
Energieeffiziente Analyse hochvolumiger Sequenzdaten mit rekonfigurierbaren Architekturen

SeqAn@FPGA - Energieoptimierte Genomsequenzierungs-bibliothek für energieeffiziente programmierbare Mikroelektronik in Rechenzentren

1 Projektinformation

Motivation:

- exponentielles Datenwachstum durch *Next-Gen Sequencing*
- künftig wachsende Bedarfe mit personalisierter Medizin
- CO₂-Einsparung im Hochdurchsatzrechnen



Ziel des Vorhabens

- Implementierung von Schlüsselfunktionen der **SeqAn** Software-Bibliothek für Analysen genomischer Sequenzen auf nutzerprogrammierbaren integrierten Schaltkreisen (*Field Programmable Gate Array*)

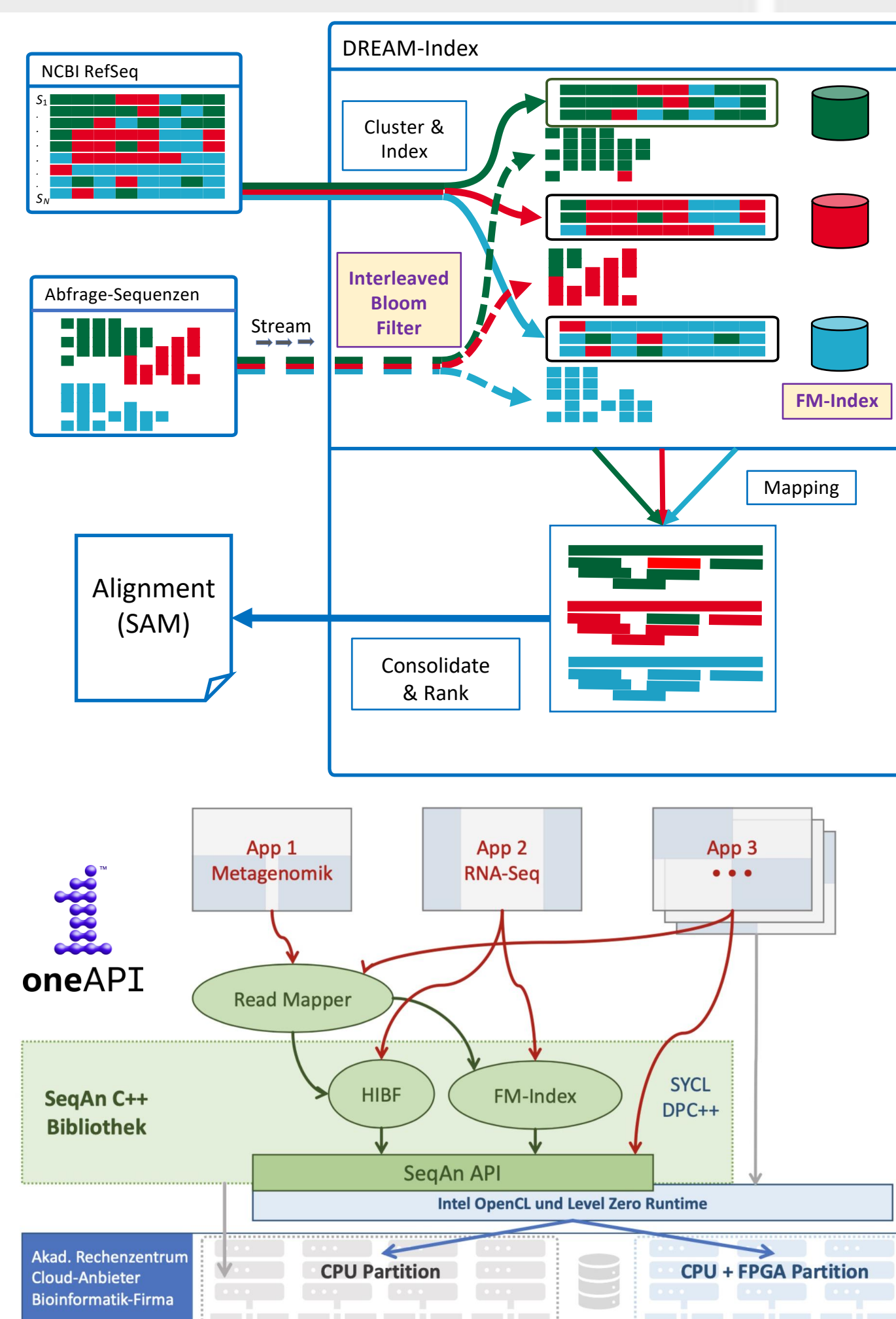
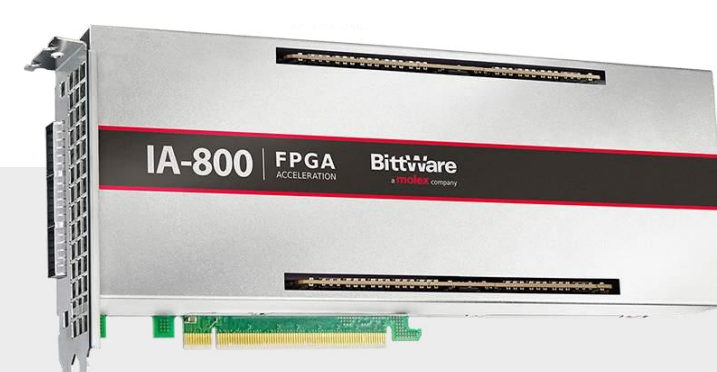
2 Technologie & Anwendungsfelder

- Zielplattformen: **Rekonfigurierbare Architekturen** mit **Field-Programmable Gate Arrays (FPGA)**
- Hochdurchsatzanalyse biologischer Sequenzdaten
- Workflows für personalisierte Medizin
- SeqAn** für Patientenversorgung & Forschung

3 Projektziele und Methodik

- signifikante **Energiereduktion** für Analysen von Genomsequenzen
- Demonstrator: *Read Mapper mit DREAM-Index*
- Implementierung der Datenstrukturen *Hierarchical Interleaved Bloom Filter (HIBF)* und *FM-Index* aus **SeqAn** auf **FPGA**

- Nutzung der Vorteile von FPGAs:
 - Algorithmus definiert Hardware-Design
 - optimale Repräsentation von Sequenzdaten (z.B. DNA mit 4 Bit je Element)
 - effiziente Bit-/String-Operationen
- Nutzung des *High-Level Synthesis (HLS)* Ansatzes für Produktivität („Programmierung“ eines FPGA) → Implementierungen mit **SYCL** (Intel oneAPI)
- FPGA-Entwicklungsplattform: BittWare IA-840f mit Intel Agilex 7 FPGA

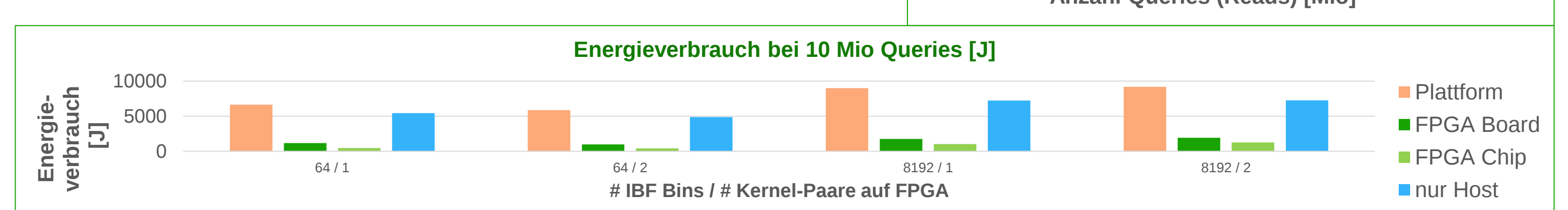
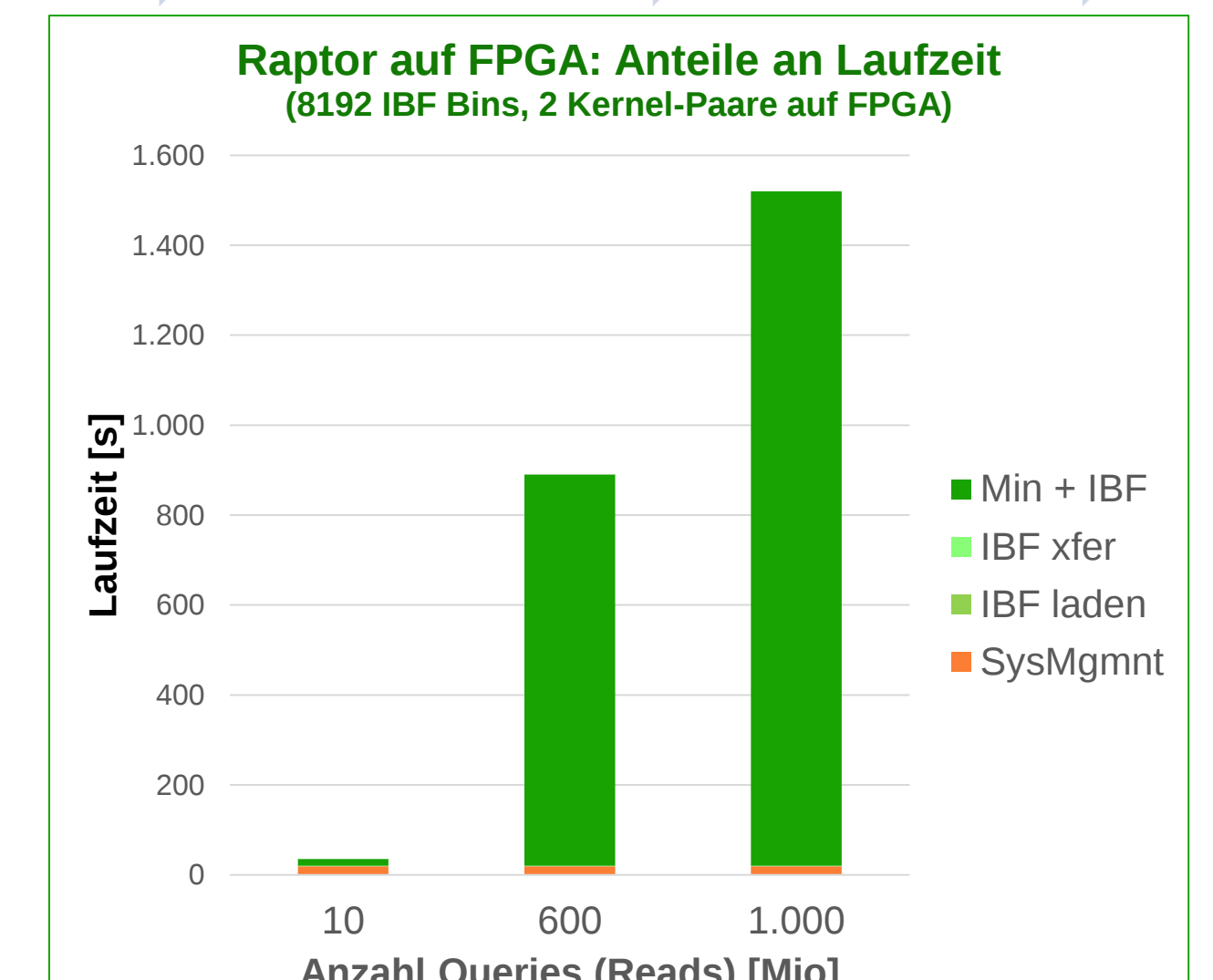
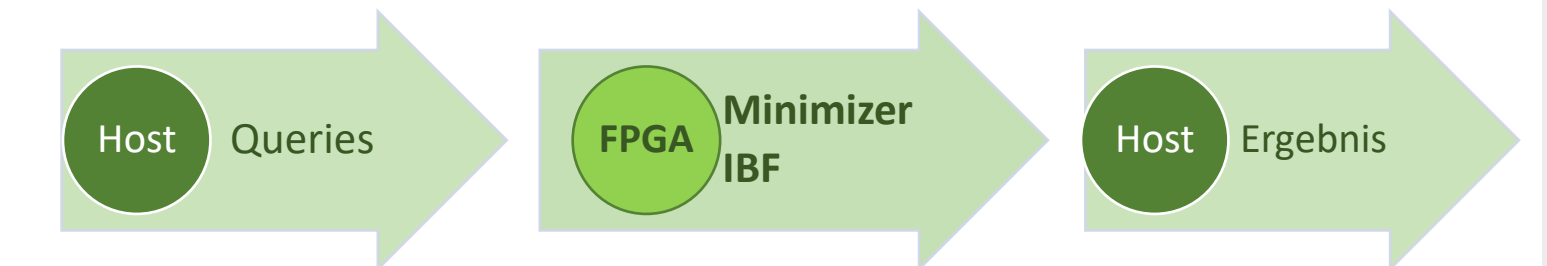
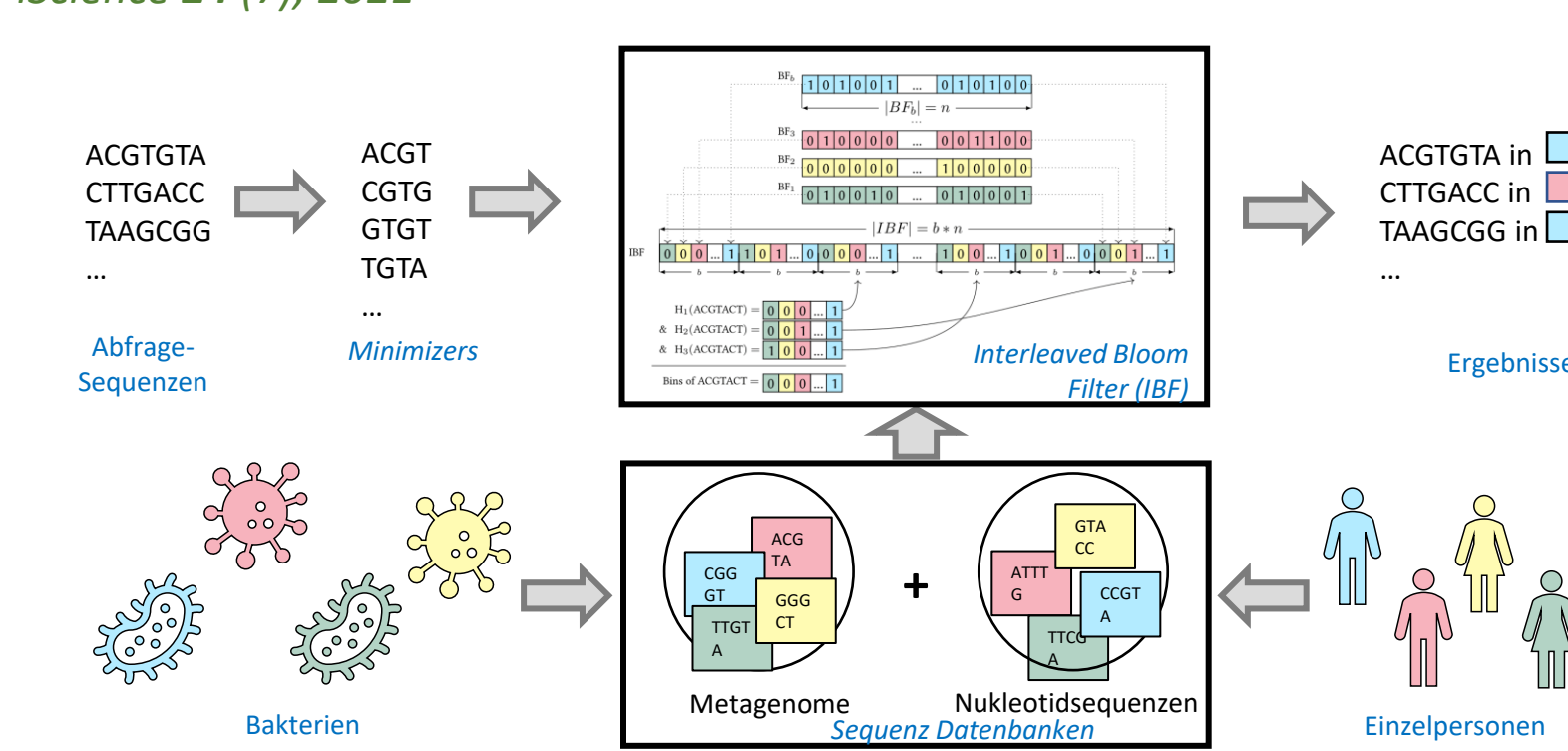


4 Ergebnisse und Erkenntnisse

Raptor mit IBF auf FPGA

- Raptor**: ein schneller und speicherplatzsparender Filter für sehr große Sequenzdaten

Enrico Seiler, Svenja Mehringer, Mitra Darvish, Etienne Turc, Knut Reinert;
iScience 24 (7), 2021



IBF Bins	Energiereduktion Kernel auf Device: FPGA vs. CPU	Energiereduktion Server Plattform: FPGA- vs. nur-CPU-Plattform	Energieverbrauch vom Host typisch 80%!
64	7,7 x	6,2 x	→ <i>Host-less</i> Infrastrukturen
8192	5,3 x	2,4 x	

FPGA: Intel Agilex 7 auf BittWare IA-840

CPU: 2x AMD EPYC 9654 (Genoa) auf Lenovo ThinkSystem SD665 V3



5 Innovation & Ausblick

- hohes Energieeinsparpotenzial mit FPGA
- SYCL-Implementierung in SeqAn erlaubt transparenten Zugriff auf FPGAs
- SeqAn ermöglicht sehr hohe Breitenwirksamkeit
- Herausforderung HLS-Ökosystem **intel**
- alternative Datenfluss-Architekturen **NEXTSILICON**

6 Kontakt

Prof. Dr. Knut Reinert
Freie Universität Berlin
Knut.Reinert@fu-berlin.de

Dr. Thomas Steinke
Zuse-Institut Berlin
steinke@zib.de

Gefördert durch:



Knut Reinert, Enrico Seiler (FUB)
Tobias Baumann, Mahnoor Maghroori, Thomas Steinke (ZIB)
Christian Färber (Intel / Altera)

MAX PLANCK INSTITUTE
FOR MOLECULAR GENETICS



intel altera



Förderkennzeichen: 16ME0588K / 16ME0589