
Seminarbaustein: Next Generation Sequencing Variant Calls mit Galaxy

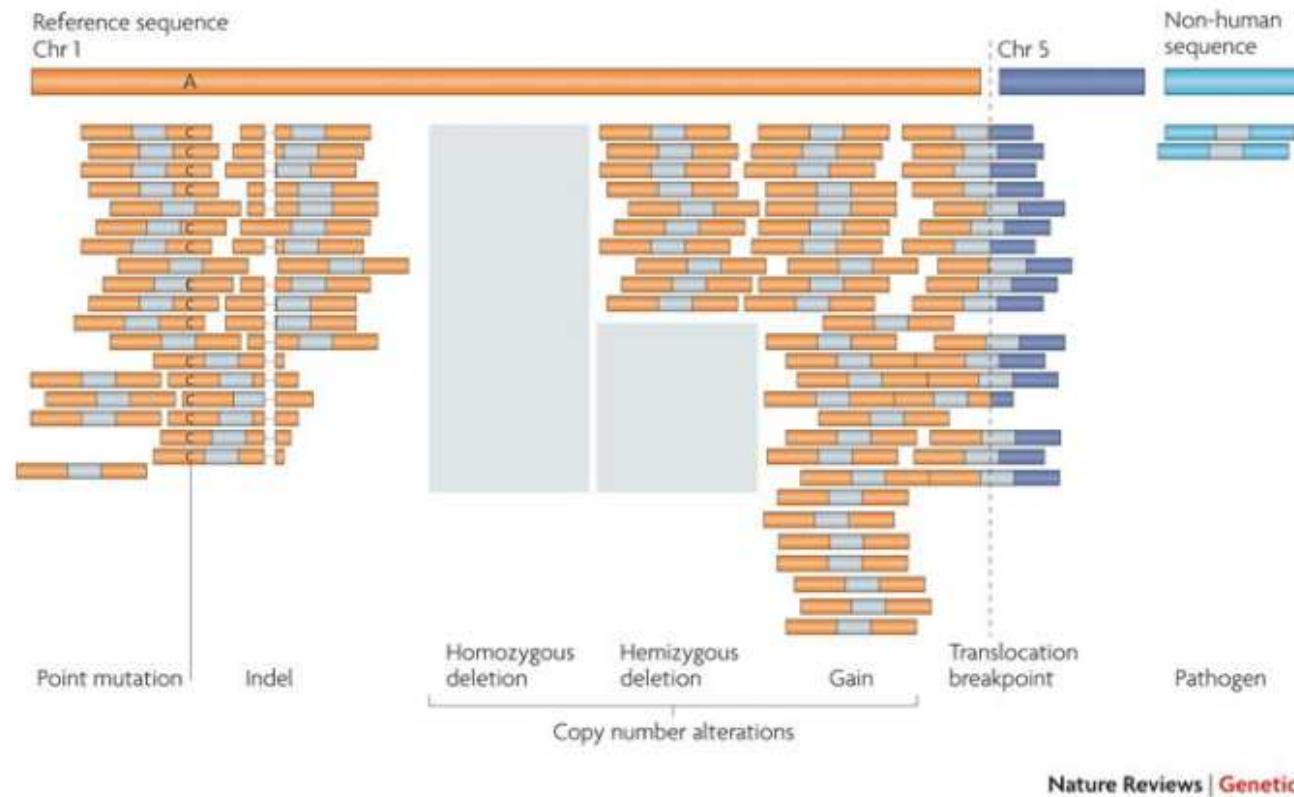
Hochkomplexe Analysegeräte wie NGS produzieren in kürzester Zeit umfangreiche Datenmengen, die für Forschung, Entwicklung und Produktion zeitnah auszuwerten sind. Das stellt besondere Anforderungen an das Fachpersonal. Im Seminar erlernen Sie die Anwendung und erhalten praktische Hilfe.

Hier sehen Sie ein Umsetzungsbeispiel.

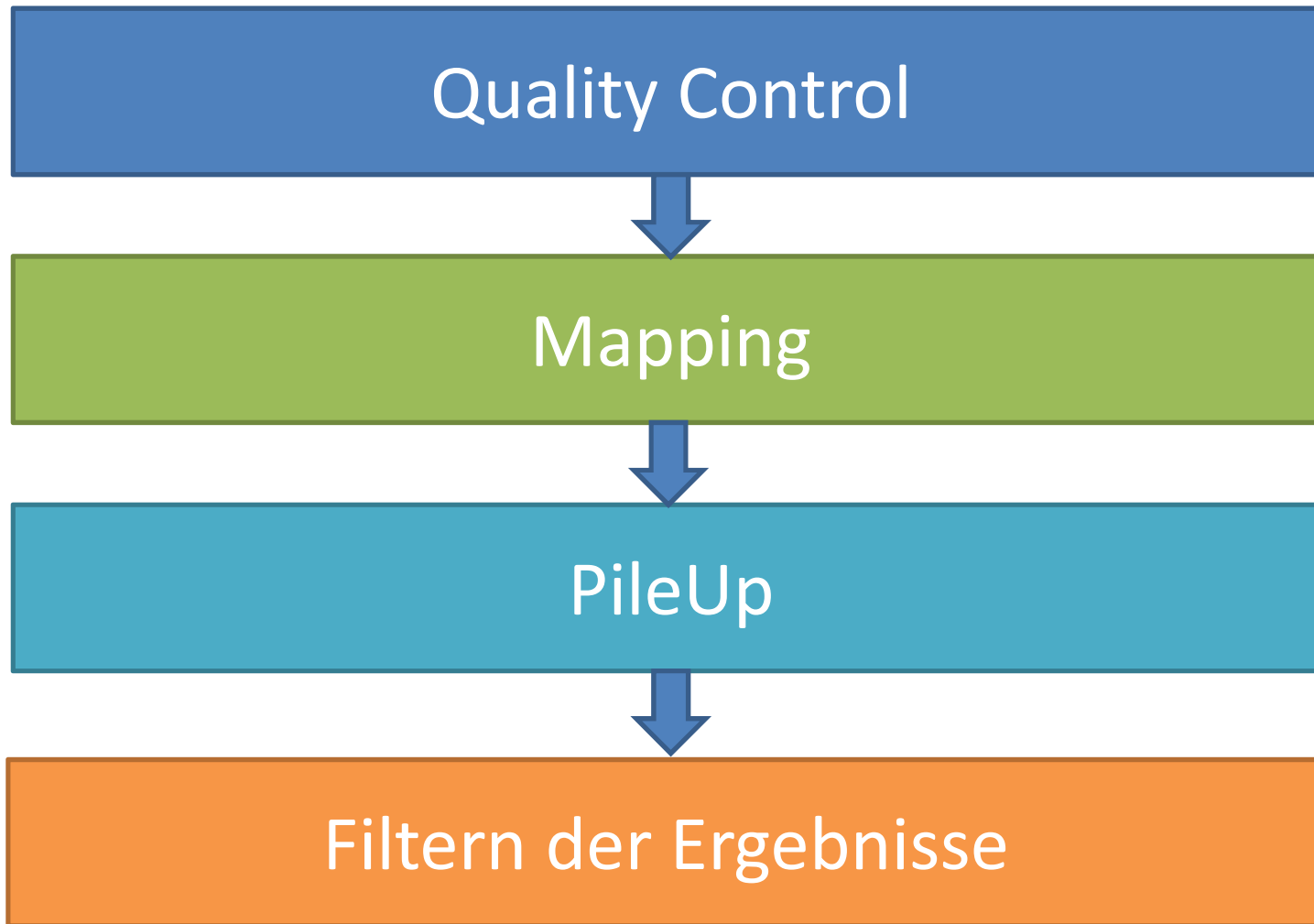
Dozent: M. Sc. Benjamin Böhmer



Somatic Mutations



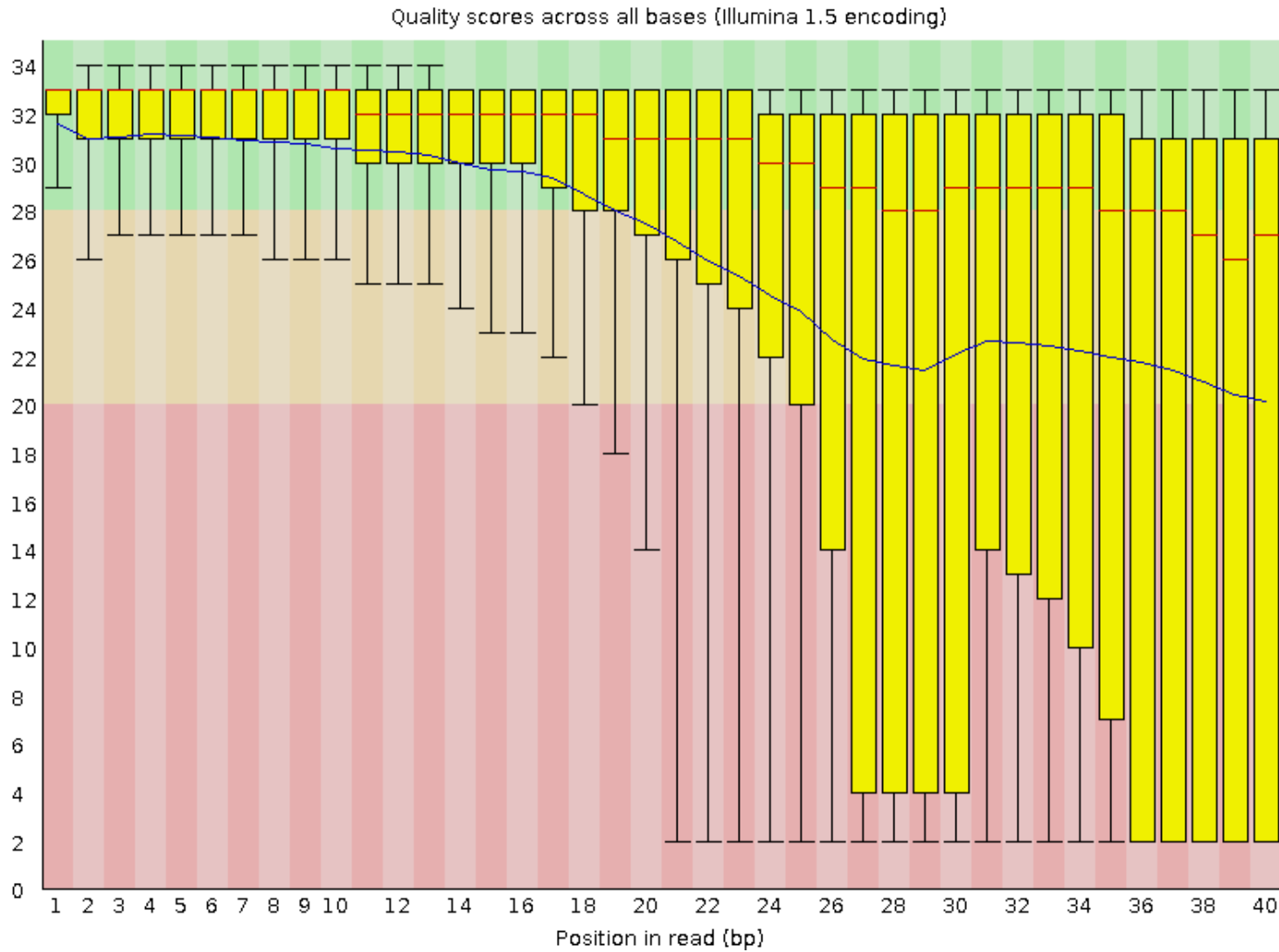
Meyerson M et al. Advances in understanding cancer genomes through second-generation sequencing. Nature Reviews Genetics (2010) 11, 685-696

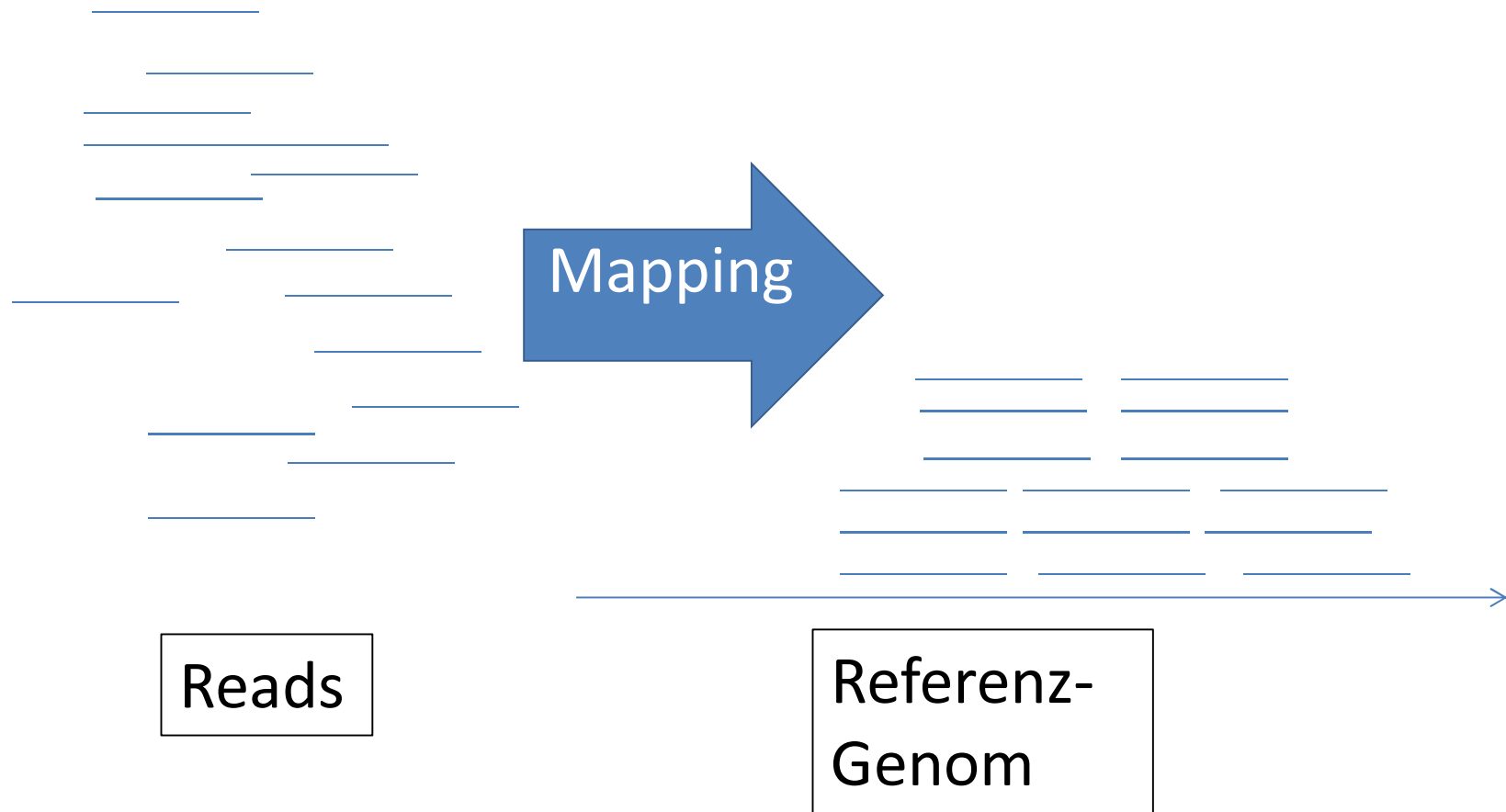


- Alle Sequenzen in einer Datei
- Qualitätsscore für jede Base

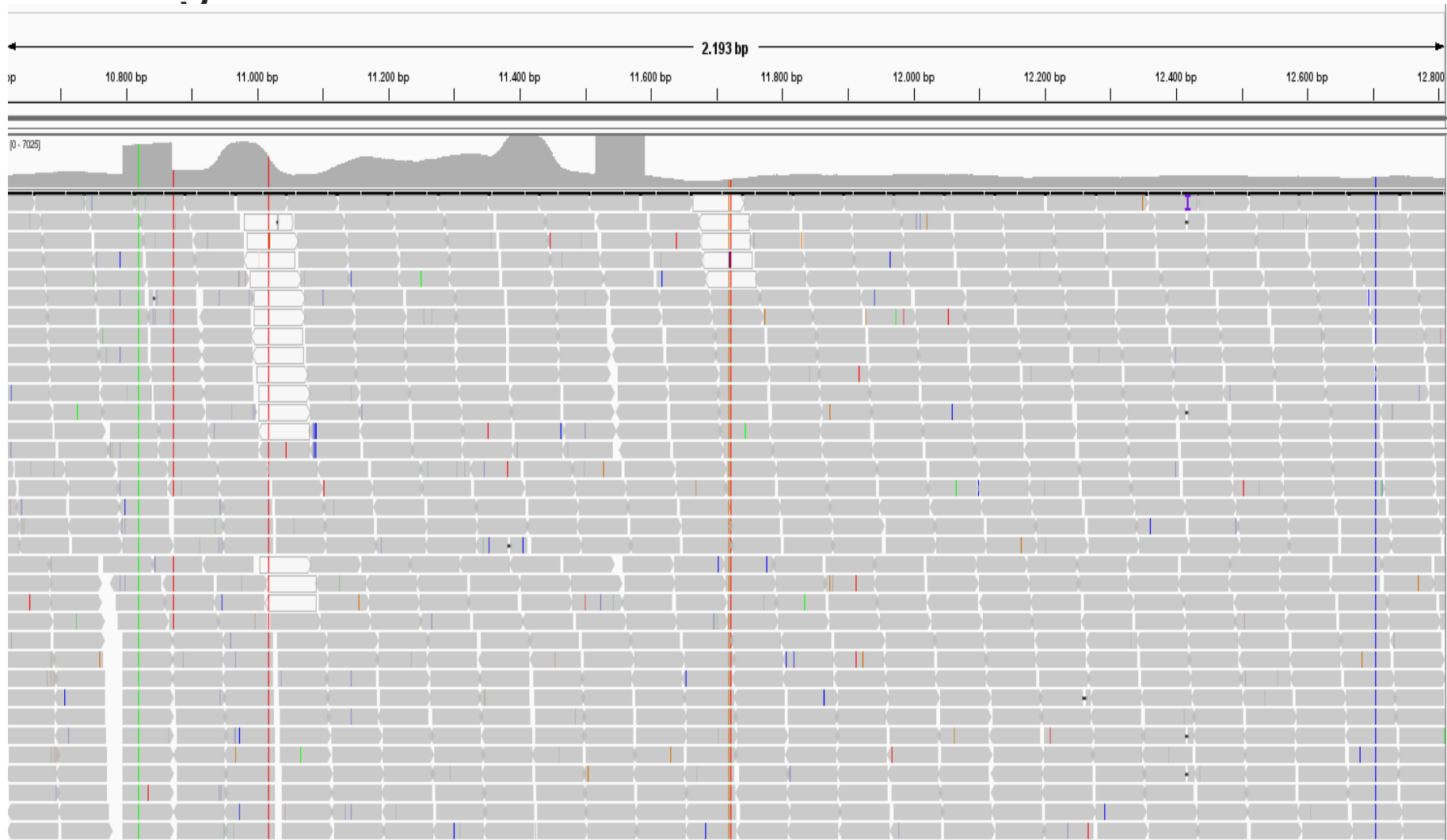
```
@M00805:5:000000000-A0VLL:1:1101:16473:1320 1:N:0:1
NTTGTCATCAGCTGAAGATGAAATAGGATGTAATCAGACGACACAGGAAGCAGATTTTGCTAAT
TTGGAAGTAGGTCAGCTGAAGATCCTGTGAGCGAAGTTCGGCAGTGTCACAGCAC
+
#55<<?BBDBDDDDDDFFFFFFHHHHHFHHAFHHHHHHHHHBHHHHHFFHHHHHHHHHDGDGHC
AFHFHHHHHHFGHDDHFBFHDFHFFHHHFFA=@BEEEEED)@<B?BE3==?EEEE
@M00805:5:000000000-A0VLL:1:1101:15023:1321 1:N:0:1
NAGAAATCACAGACATACAAAGCAGTCTGTGTCCTTAGGTCCTGAGCAGCCTCCAGCACATTCT
AGCATCTGCCGTCACATTGTTCTGCACACACCGTCCTTGTCACTGCAGAAGACAGA
+
#55???BBDEDDDDDGGGGGGIIIIIIIIIIIIIIIIIIIIIIHIIHIIHFGHHHIIIIIIIIHIIII
HHHHHHHHHHHHHHHHHHHHHHHHHHGGFGEGGGGGGGGGGGGGGGGGEGGGGCEGG>
@M00805:5:000000000-A0VLL:1:1101:14046:1321 1:N:0:1
NTTTCGTGGAAGTGGGTTACCTGACAGTGTGCACGCCCCAGCAGGTTACAATATTCTCGTGG
ACATGAGTGCCTCTCTTTCAGAGCTGTCTGCTTTTTCTGTCAAAGAAAGGAGCATT
```

❌ Per base sequence quality

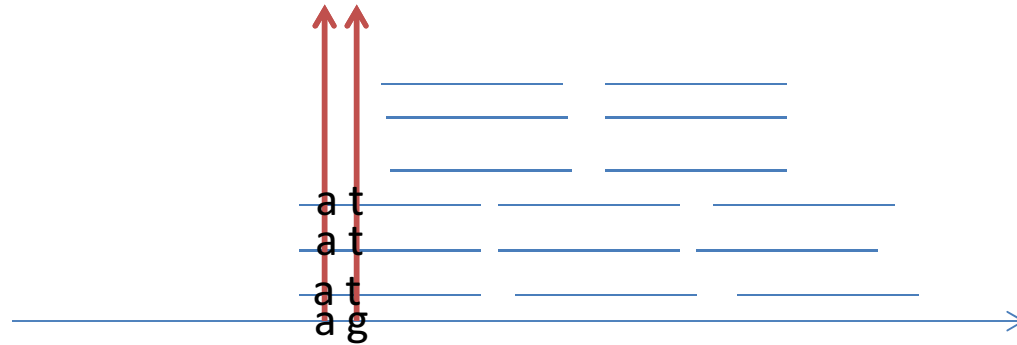




- Integrative Genome Viewer



Information pro Position: PileUp



Referenz-
Genom

Position	Referenz Nukleotid	Alignierte Nukleotide
1	a	aaa
2	g	ttt

← ↑ →
VARIANZ



- Vorteile: Mitogenom ist klein (ca 16k bp)
 - ▶ Schnelle Durchführung des Mappings

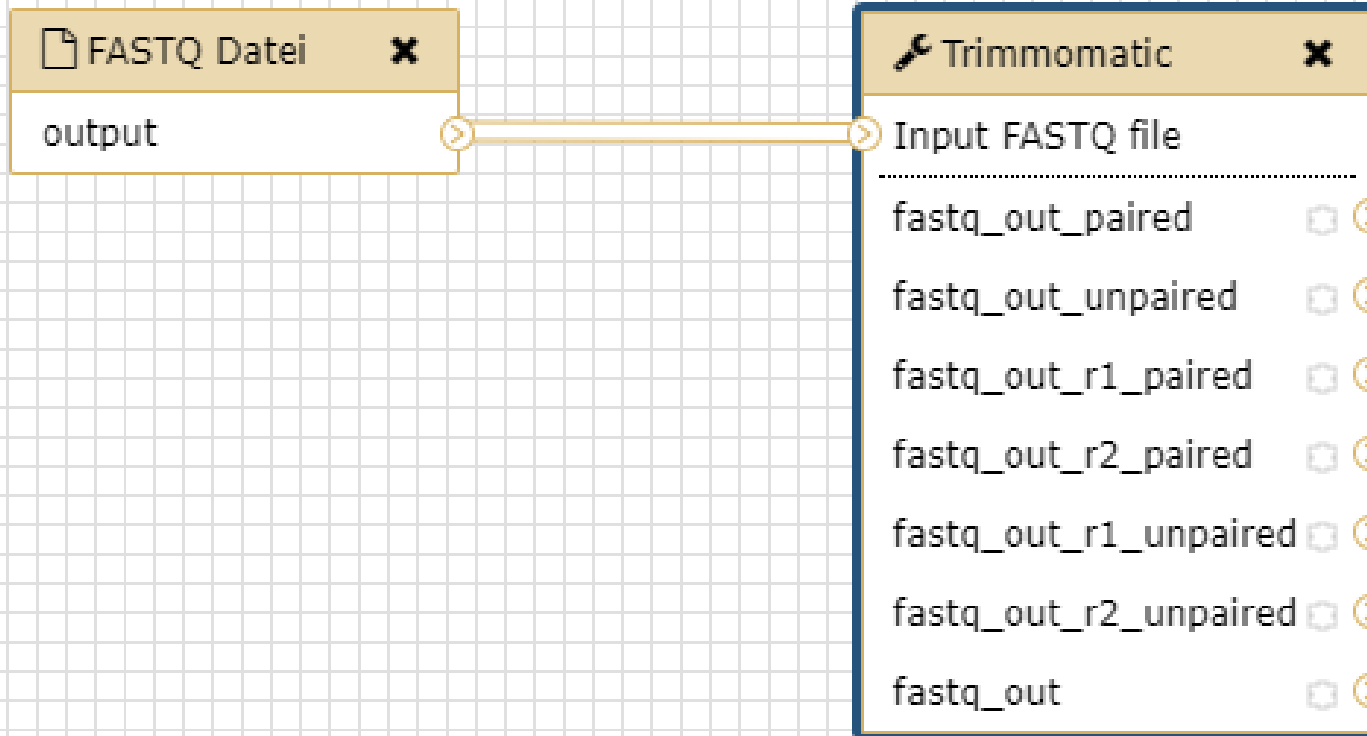
- Nachteil: Viele verschiedene Kopien pro Zelle
 - ▶ Gefundene Variationen oft nicht sehr eindeutig

<https://de.wikipedia.org/wiki/Mitochondrium>

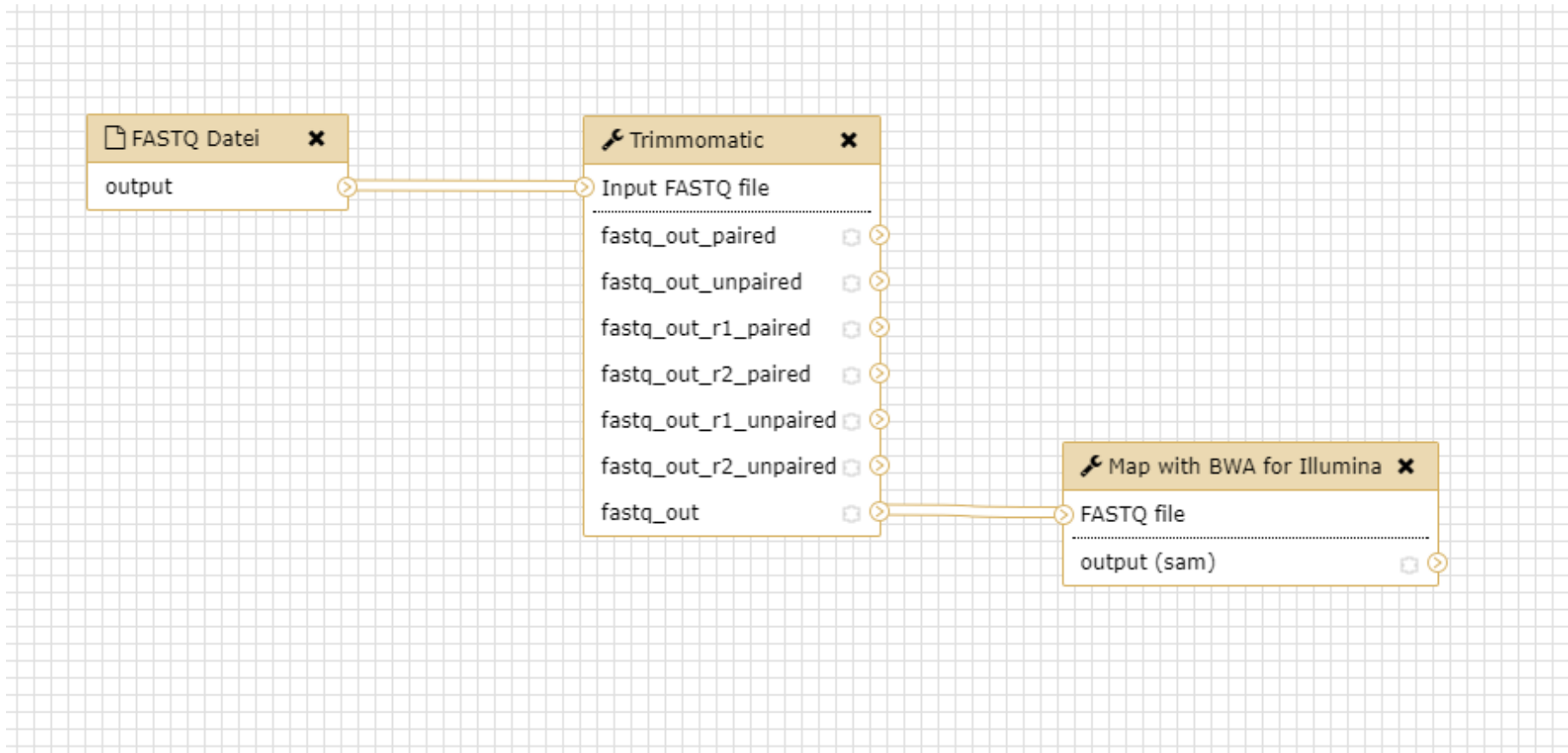
- Webbasiert
- Open-Source
- Leichte Erstellung von Workflows
- Sehr große Sammlung von Bioinformatiktools
- Vor allem zum Thema NGS



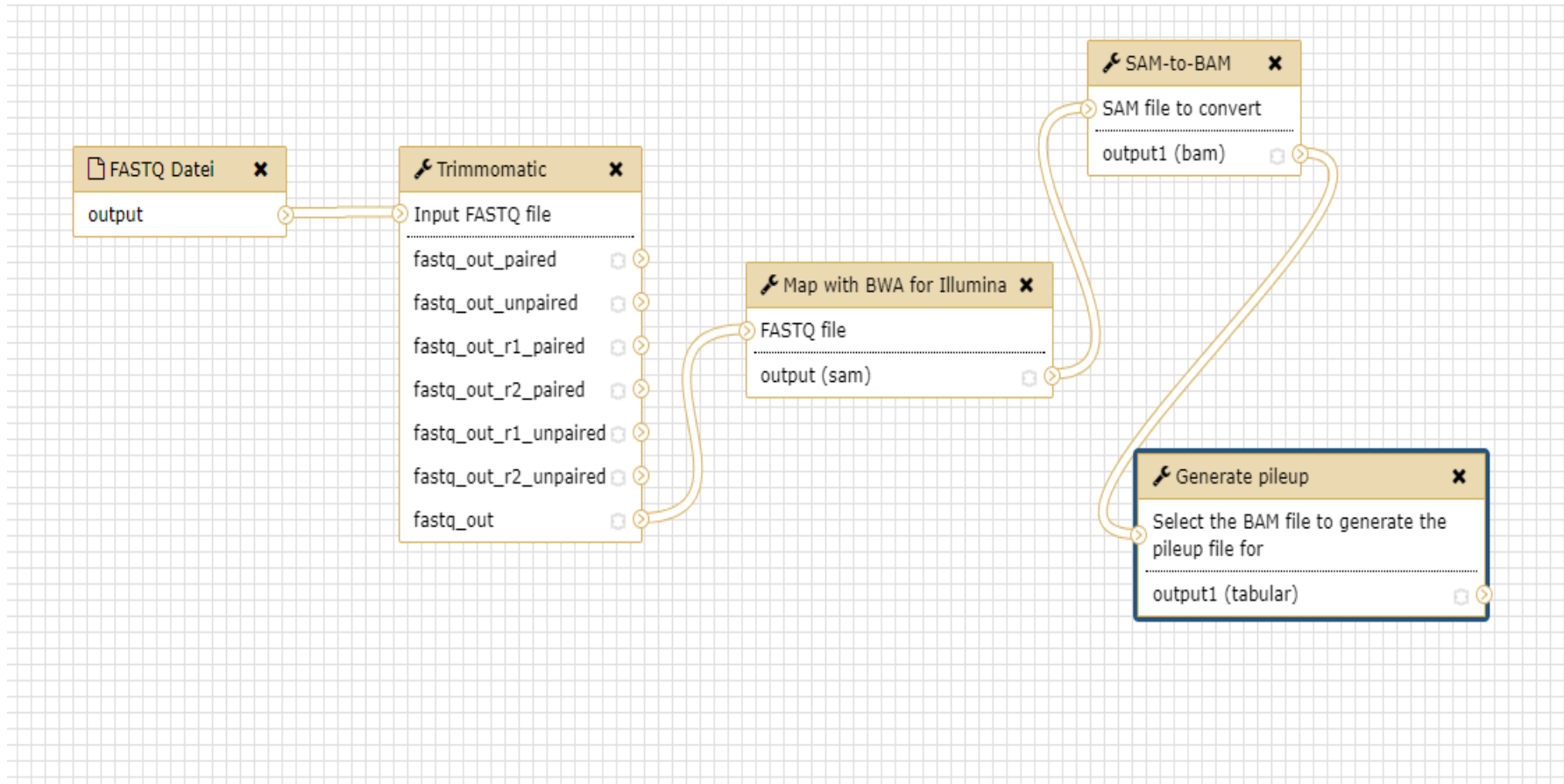
Galaxy Workflow: Read Quality Control

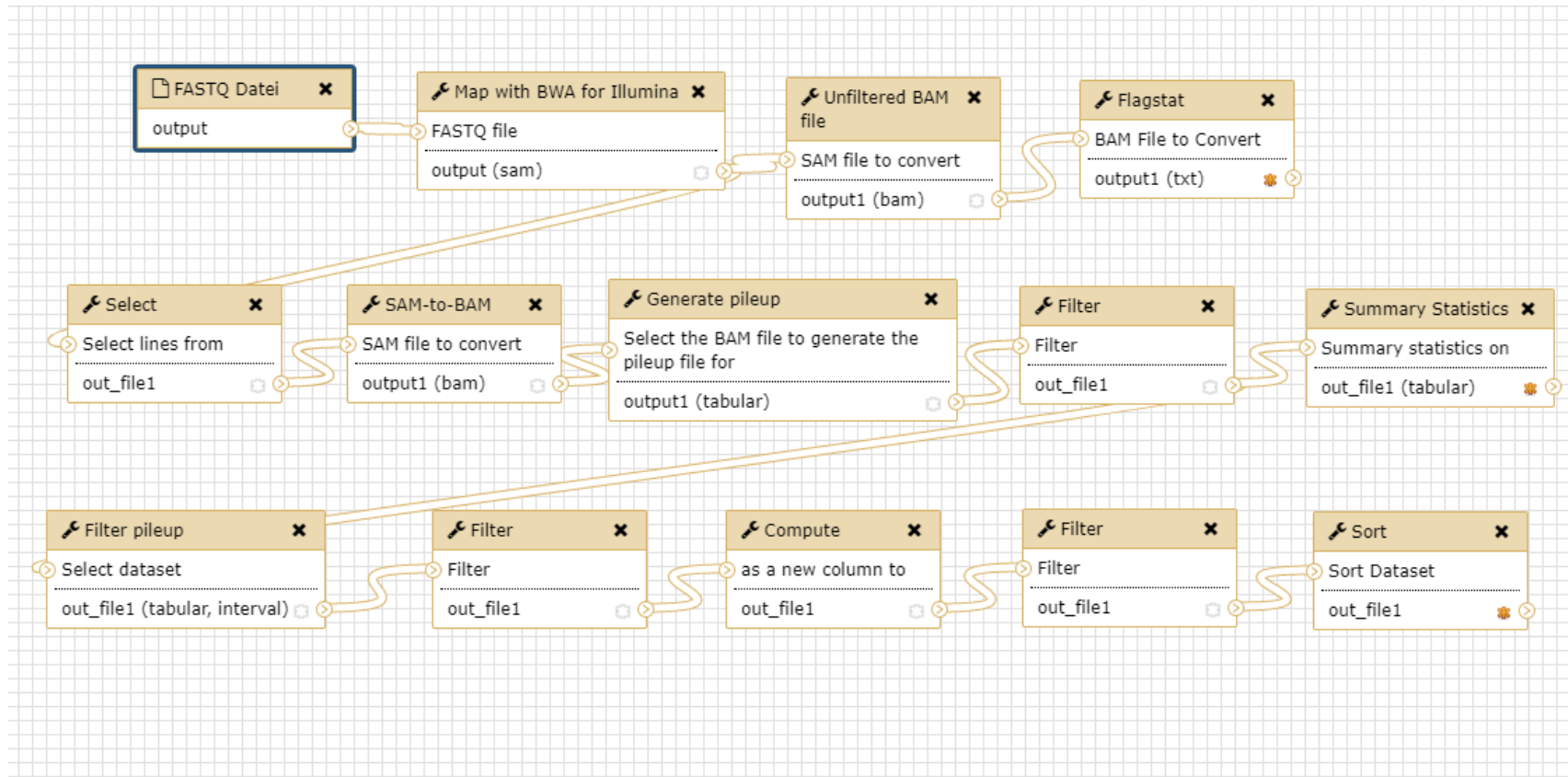


Galaxy Workflow: Mapping



Galaxy Workflow: PileUp

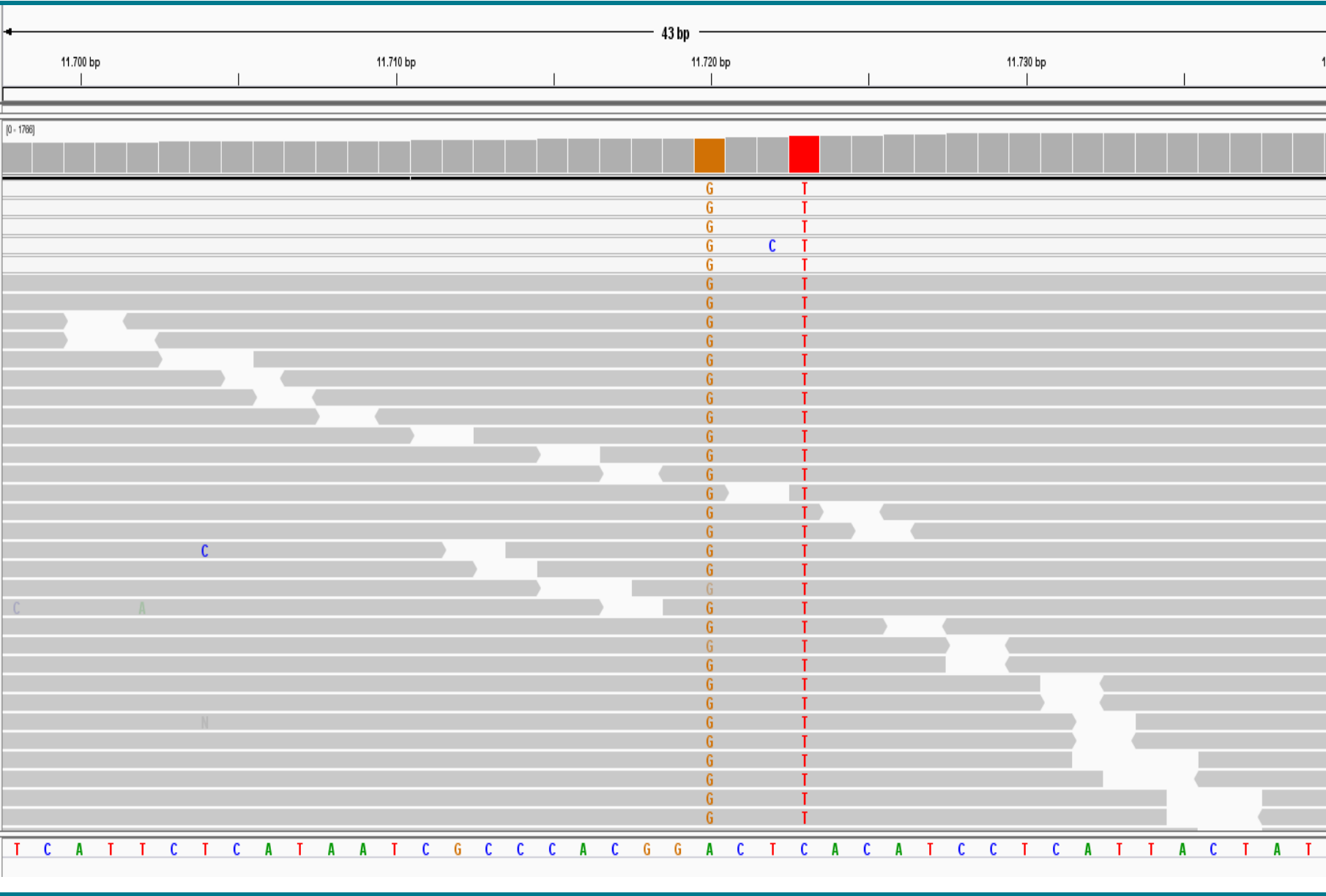




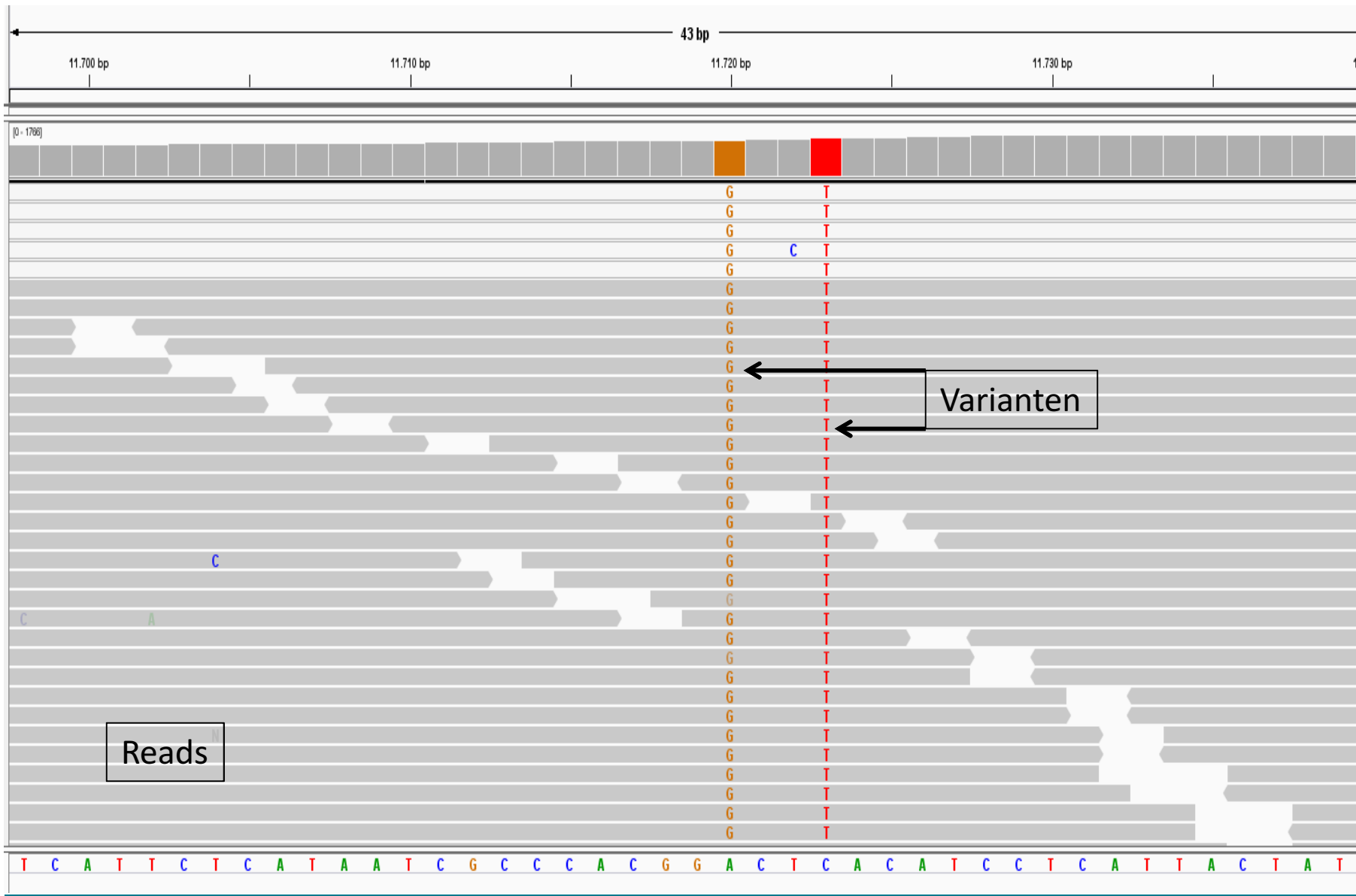
Ergebnis (Die ersten vier)

Chromosom	Position	Referenz-Nukleotid	Variante
chrM	11719	A	G
chrM	11722	C	T
chrM	12705	T	C
chrM	14580	G	A

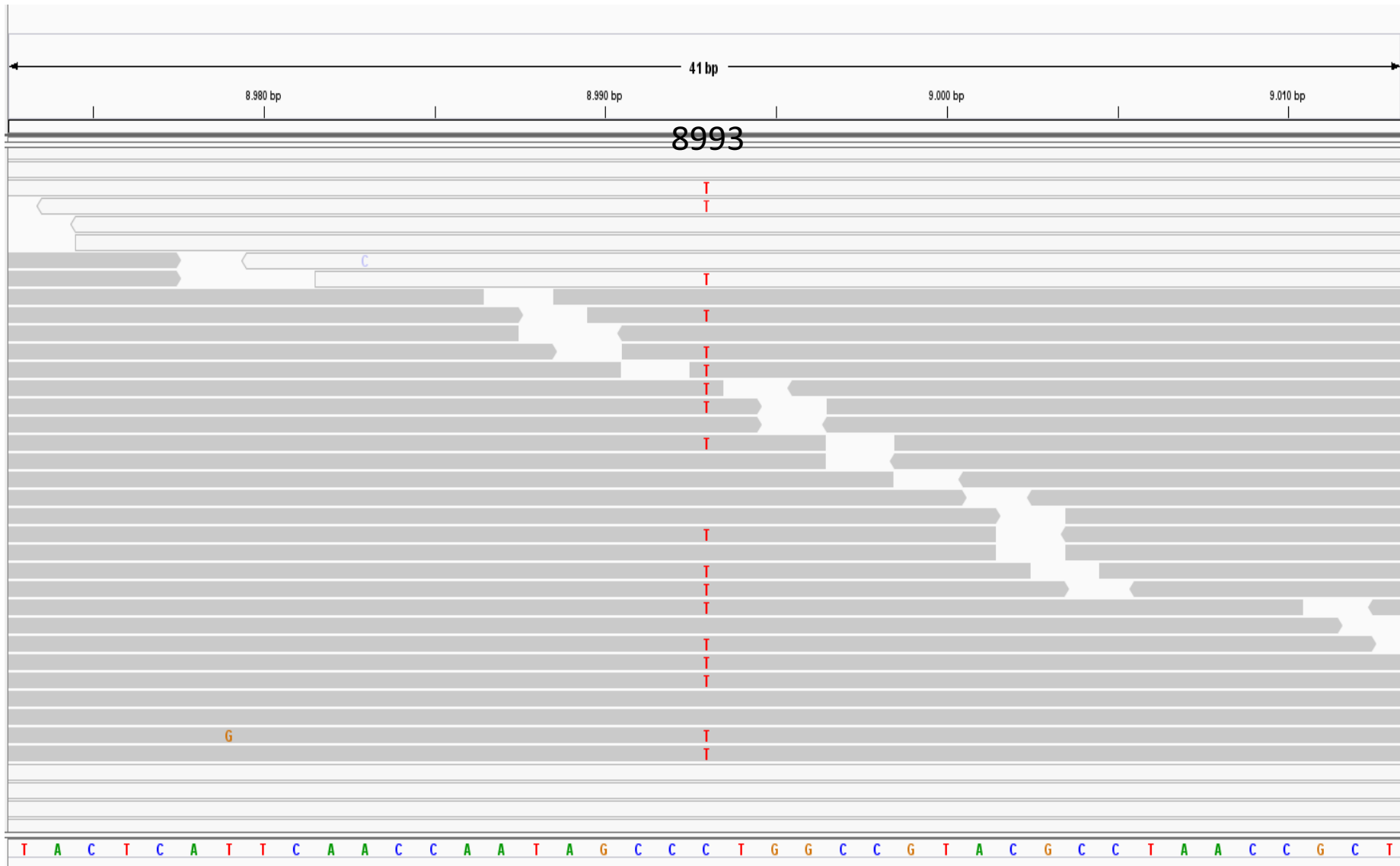
Variant Call Alignierte Reads



Variant Call Alignierte Reads



Leigh-Syndrome



- Differentielle Genexpressionsanalysen:
 - ▶ RNA-Seq
- Finden von DNA-bindenden Motiven in Proteinen:
 - ▶ ChIP-Seq
- Finden von Methylierungsstellen in DNA:
 - ▶ Methyl-Seq
- DeNovo Genomsequenzierung:
 - ▶ Genome-Assembly

<https://usegalaxy.org/>

<http://software.broadinstitute.org/software/igv/>

<http://www.usadellab.org/cms/?page=trimmomatic>

<http://bio-bwa.sourceforge.net/>

<http://samtools.sourceforge.net/pileup.shtml>