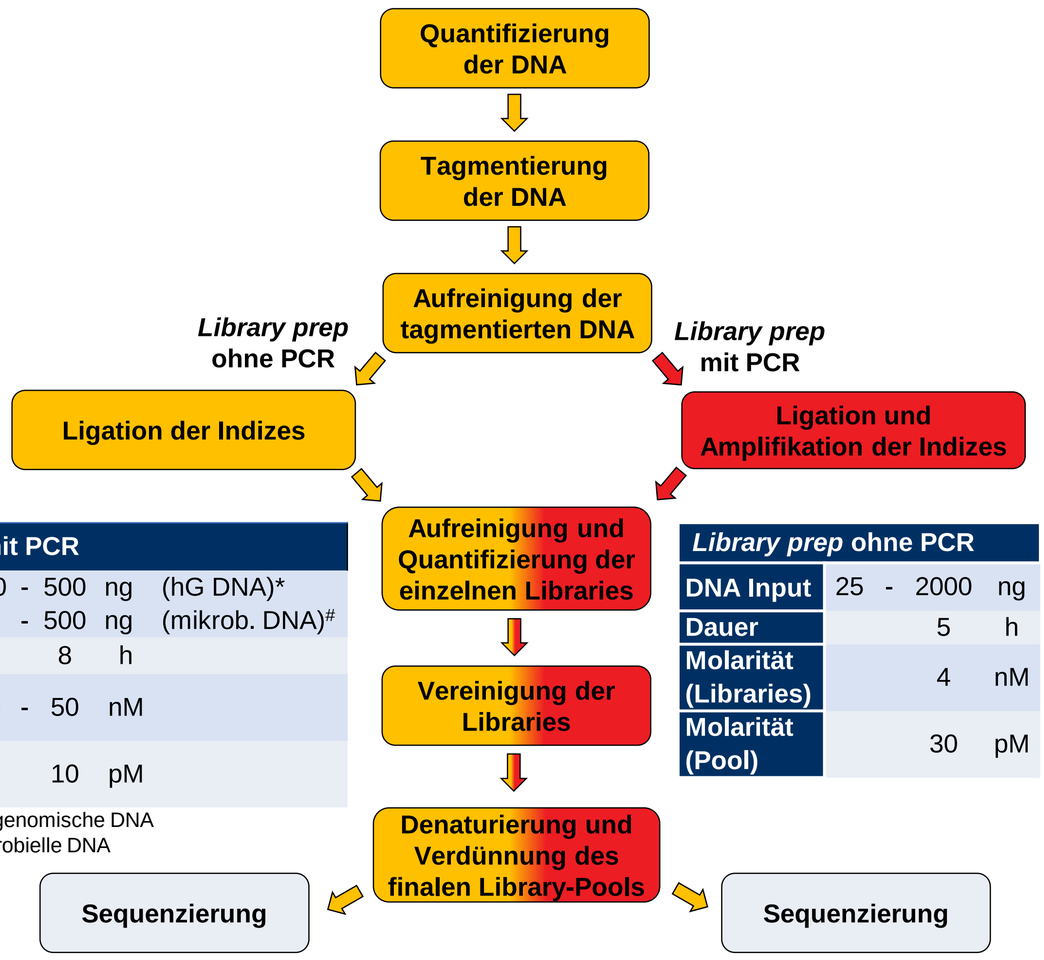
TYPISIERUNG BAKTERIELLER ERREGER IN DER LEBENSMITTEL- UND HUMANANALYTIK

Im täglichen Routinebetrieb unterstützt das NGS-Labor am Institut für Hygiene und Umwelt (NGS@HU) auch die Lebensmittelüberwachung und Aufklärung von Ausbruchsgeschehen. Dabei ist eine ubiquitär eingesetzte Methodik die Typisierung von bakteriellen und viralen Krankheitserregern. Im Gegensatz zur klassischen Serotypisierung wird im NGS-Labor das entsprechende Isolat, genauer gesagt, dessen Gesamtgenom mittels *massive parallel sequencing* (MPS) sequenziert. Im Anschluss einer solchen Ganzgenomsequenzierung (*whole genome sequencing*, WGS) wird mit rechnergestützten Analysen das Isolatgenom in seiner Gesamtheit rekonstruiert (*assembly*) und jedes identifizierte Gen einem funktionellen Kontext zugewiesen (*annotation*). Ganzgenomsequenzierungen erfolgen je nach Vorhandensein mit oder ohne Abgleich zu einem Referenzgenom (*reference based assembly* vs. *de novo assembly*) und ermöglichen in Kombination mit einer Kerngenom Multi-Locus-Sequenztypisierung (cgMLST, *core genome multi locus sequence typing*) oder durch Bestimmungen des individuellen Einzelnukleotid-Polymorphismus (SNP, *single nucleotide polymorphism*) eine Beschreibung des vorherrschenden Typs oder Serovars.

I - Erstellung von DNA-Bibliotheken (*library prep*)



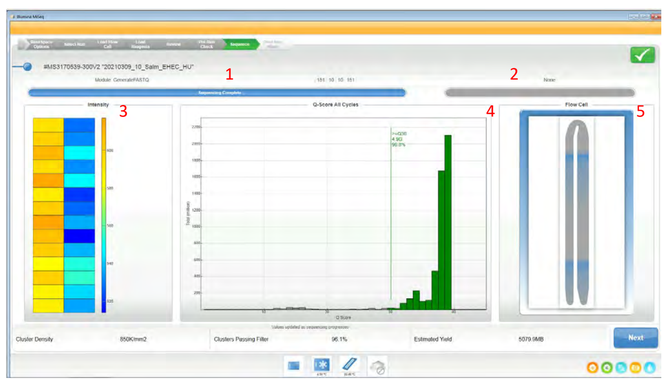
Library prep mit PCR		
DNA Input	100 - 500 ng	(hG DNA)*
	1 - 500 ng	(mikrob. DNA)#
Dauer	8 h	
Molarität (Libraries)	30 - 50 nM	
Molarität (Pool)	10 pM	

*hgDNA = humane genomische DNA
#mikrob. DNA = mikrobielle DNA

Library prep ohne PCR	
DNA Input	25 - 2000 ng
Dauer	5 h
Molarität (Libraries)	4 nM
Molarität (Pool)	30 pM

Bei der DNA-Bibliothekserstellung ist die ubiquitär eingesetzte Methode, die oben dargestellte *Library prep* mit PCR. Hierbei werden die Indizes abweichend zur *Library prep* ohne PCR neben der Ligation auch amplifiziert. Daraus resultiert, dass bei der *Library prep* mit PCR die gesuchten Zielsequenzen in den finalen Libraries höher konzentriert sind. Durch die Amplifikation kann es jedoch zu Sequenzmaskierungen kommen. Zusätzlich unterscheiden sich beide Methoden bei der Qualitäts- und Quantitätskontrolle der finalen Libraries. Mittels Fluorometer wird bei *Library prep* mit PCR die Quantität der doppelsträngigen DNA (dsDNA) gemessen, wohingegen bei der *Library prep* ohne PCR die einzelsträngige (ssDNA) bestimmt wird. Eine Qualitätskontrolle mittels Fragment Analyzer ist nur bei dsDNA möglich.

II - Sequenzierung und Laufparameter (Illumina MiSeq™)



Sequenzierparameter zur Typisierung von bakteriellen Isolaten aus Lebensmitteln und humanen Proben zur Aufklärung von Infektionsketten. Die Übersicht zeigt den Fortschritt[1], die Ausfälle[2], die Clusterdichte[3], den Anteil über Qualitätswert Q30[4] sowie eine Darstellung der Fließzelle[5].

III – Bioinformatik

1. ASSEMBLIERUNG und QUALITÄTS-CHECK via Aquamis *pipeline*[1]

Sample Name	QC Vote	Species	Total Length	# Contigs	Q30 Base Fraction	Coverage Depth	Read Fraction Majority Genus	Contig Fraction Majority Genus	Contam. Status	Contam. # SNVs	MLST Loci w/ Multiple Alleles
						A					
58	FAIL	Salmonella enterica	5,081,243	117	0.869	94.8	0.984	0.996	False	0	0
59	PASS	Salmonella enterica	5,072,636	105	0.895	95.8	0.998	0.998	False	0	0
96	PASS	Salmonella enterica	4,658,133	90	0.805	96.5	0.998	0.998	False	1	0

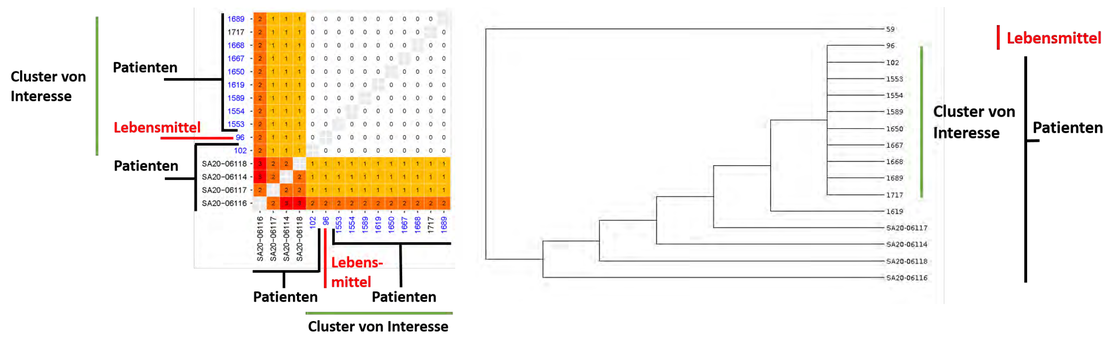
[1] https://gitlab.com/bfr_bioinformatics/AQUAMIS

2. CHARAKTERISIERUNG DES ISOLATS via BakCharak *pipeline*[2]

Sample name	cgmlst_ST	cgmlst distance	cgmlst genome match	cgmlst matching alleles	cgmlst subspecies	h1	h2	mash distance	mash genome	mash match	mash serovar	Mash subspecies	o antigen	serogroup	serovar	Serovar antigen	Serovar cgmlst
58	0,042424	330	SRR5045375	316	enterica	d	12	0,002658	SRR980352	897	Muenchen	enterica	68	C2-C3	Muenchen	Muenchen| Virgina| Manhatten| Yovokome	Muenchen
59	0,030303	330	SRR5045375	320	enterica	d	12	0,002575	SRR980352	900	Muenchen	enterica	68	C2-C3	Muenchen	Muenchen| Virgina| Manhatten| Yovokome	Muenchen
96	0,024242	330	SRR1969406	322	enterica	d	12	0,000264	80577	989	Muenchen	enterica	68	C2-C3	Muenchen	Muenchen| Virgina| Manhatten| Yovokome	Muenchen

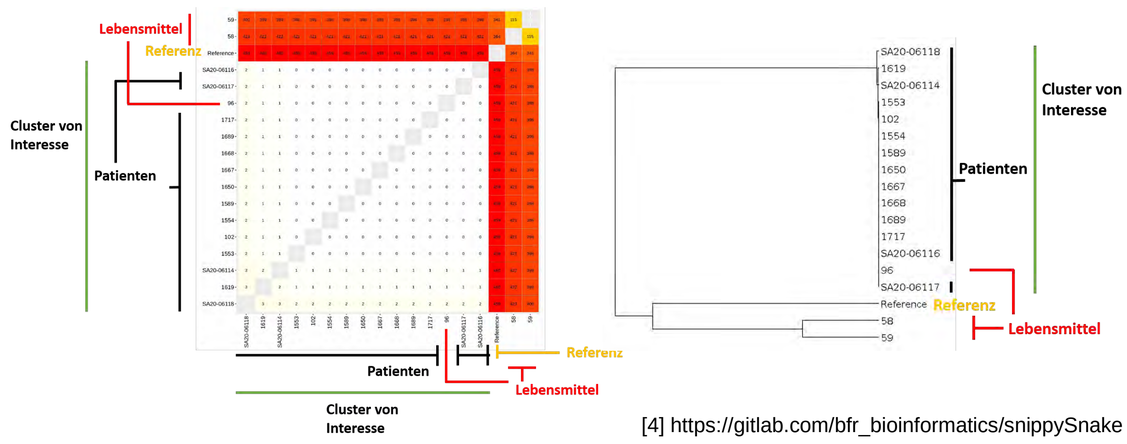
[2] https://gitlab.com/bfr_bioinformatics/bakcharak

3. cgMLST ALLELE CALLING UND HASHING via ChewieSnake *pipeline*[3]



[3] https://gitlab.com/bfr_bioinformatics/chewieSnake

4. SNP-CLUSTER-ANALYSE via snippySnake pipeline[4]



[4] https://gitlab.com/bfr_bioinformatics/snippySnake

NGS-LABOR (NGS@HU)

TYPISIERUNG VON SALMONELLA

Institut für Hygiene und Umwelt
Hamburger Landesinstitut für Lebensmittelsicherheit, Gesundheitsschutz und Umweltuntersuchungen

Hamburg