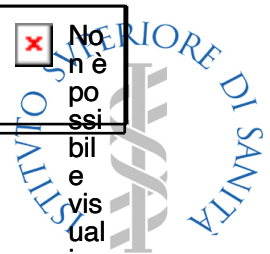




No  
n è  
po  
ssi  
bil  
e  
vis  
ual  
izz  
are  
im  
mag  
in  
gin  
e.



Società Italiana Diagnostica di Laboratorio Veterinaria  
Comitato tecnico per la genomica in sanità pubblica Veterinaria



Centro di Referenza Nazionale  
per **Sequenze Genomiche di  
microrganismi patogeni:**  
banca dati e analisi di Bioinformatica

# Strategie di sequenziamento del genoma di SARS-CoV-2 a confronto

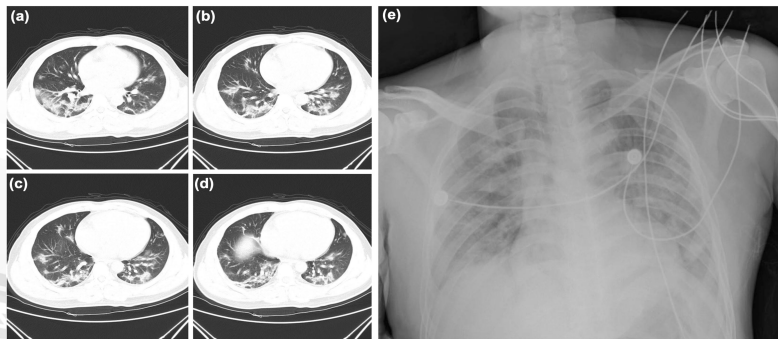
26 Novembre 2020

I laboratori di sanità pubblica veterinaria e la ricerca nella  
genomica di SARS Cov2: Esperienze a confronto



Maurilia Marcacci  
Valentina Curini

Dicembre 2019





## Article

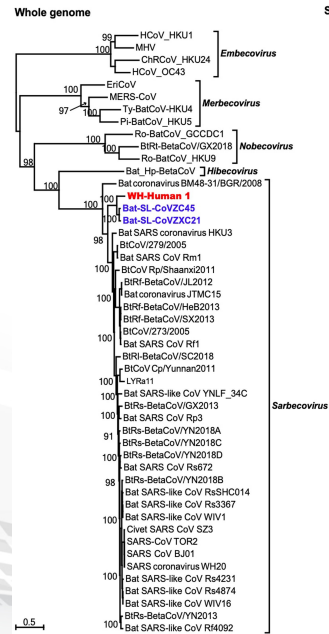
# A new coronavirus associated with human respiratory disease in China

<https://doi.org/10.1038/s41586-020-2008-3>

Received: 7 January 2020

Accepted: 28 January 2020

Fan Wu<sup>1,7</sup>, Su Zhao<sup>2,7</sup>, Bin Yu<sup>3,7</sup>, Yan-Mei Chen<sup>1,7</sup>, Wen Wang<sup>4,7</sup>, Zhi-Gang Song<sup>1,7</sup>, Yi Hu<sup>2,7</sup>, Zhao-Wu Tao<sup>2</sup>, Jun-Hua Tian<sup>2</sup>, Yuan-Yuan Pei<sup>1</sup>, Ming-Li Yuan<sup>2</sup>, Yu-Ling Zhang<sup>1</sup>, Fa-Hui Dai<sup>1</sup>, Yi Liu<sup>1</sup>, Qi-Min Wang<sup>1</sup>, Jiao-Jiao Zheng<sup>1</sup>, Lin Xu<sup>1</sup>, Edward C. Holmes<sup>1,5</sup> & Yong-Zhen Zhang<sup>1,4,6,8,9,10</sup>



# NGS





IZSAM G. CAPOREALE  
TERAMO

Centro di Referenza Nazionale  
per Sequenze Genomiche di  
microorganismi patogeni:  
banca dati e analisi di Bioinformatica





# 11 Marzo 2020 WHO dichiara SARS-CoV-2 pandemia





16 Marzo 2020





# Workflow diagnostico

TaqPath COVID-19 CE-IVD RT-PCR Kit 1,000 rxn	A48067	<a href="#">Request a quote</a>
--	--------	---------------------------------

## Workflow and products



Intended use of the products mentioned in this workflow graphic vary. For specific intended use statements please refer to the instructions for use.

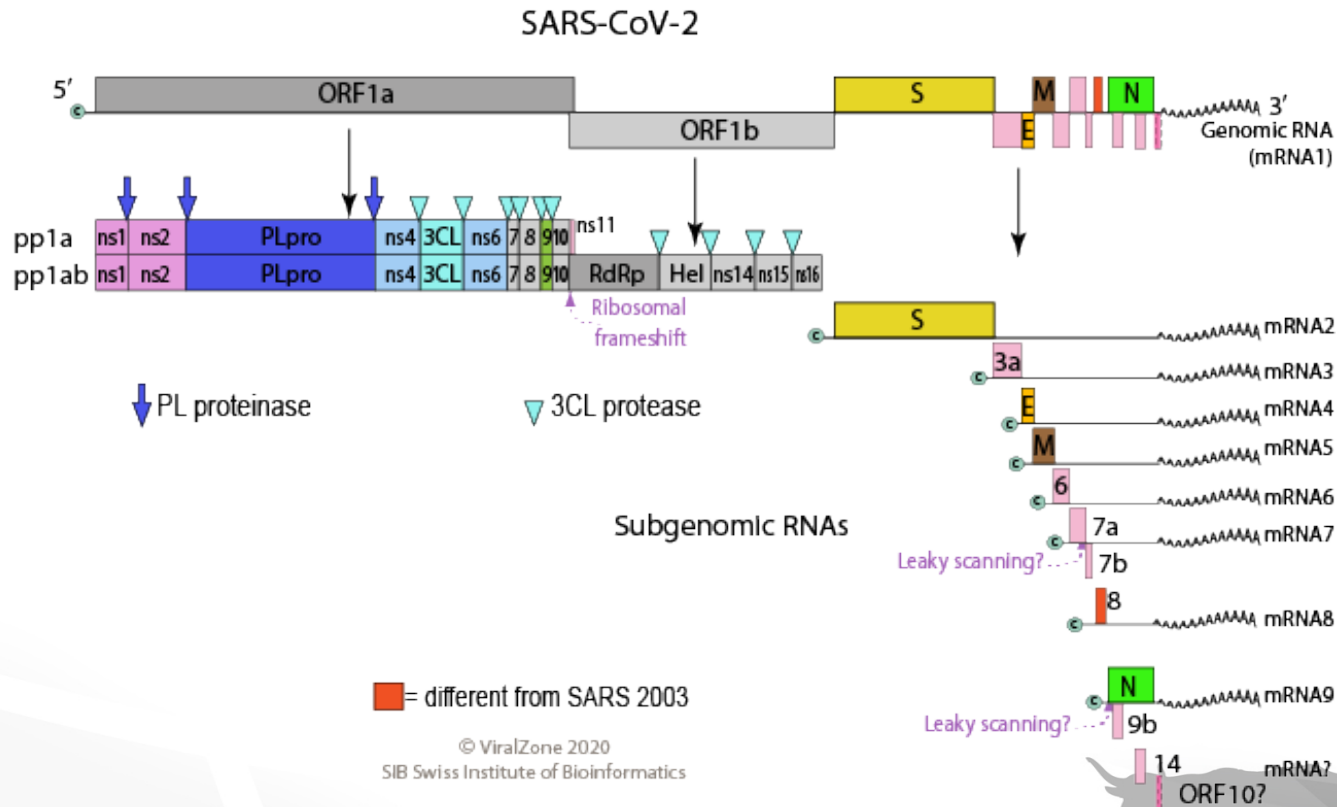


IZZAM G. CAPORALE  
TERAMO

Centro di Referenza Nazionale  
per Sequenze Genomiche di  
microorganismi patogeni:  
banca dati e analisi di Bioinformatica



# Whole genome sequencing Shotgun metagenomic approach







IZSAM G. CAPORALE  
TERAMO

Centro di Referenza Nazionale  
per Sequenze Genomiche di  
microrganismi patogeni:  
banca dati e analisi di Bioinformatica



# «Universal» protocol

## SISPA

Sequence-Independent Single-Primer Amplification





IZSAM G. CAPORALE  
TERAMO

Centro di Referenza Nazionale  
per Sequenze Genomiche di  
microrganismi patogeni:  
banca dati e analisi di Bioinformatica



# RT using special primers

Random FR26RV-N

5' GCC GGA GCT CTG CAG ATA TC **NNN NN** 3'

FR40RV-T

5' GCC GGA GCT CTG CAG ATA TC **TTT TTT TTT TTT TTT TTT TT** 3'

Virus specific primer (es. BTV)

5' GCC GGA GCT CTG CAG ATA TC **CTAA ACN** 3'



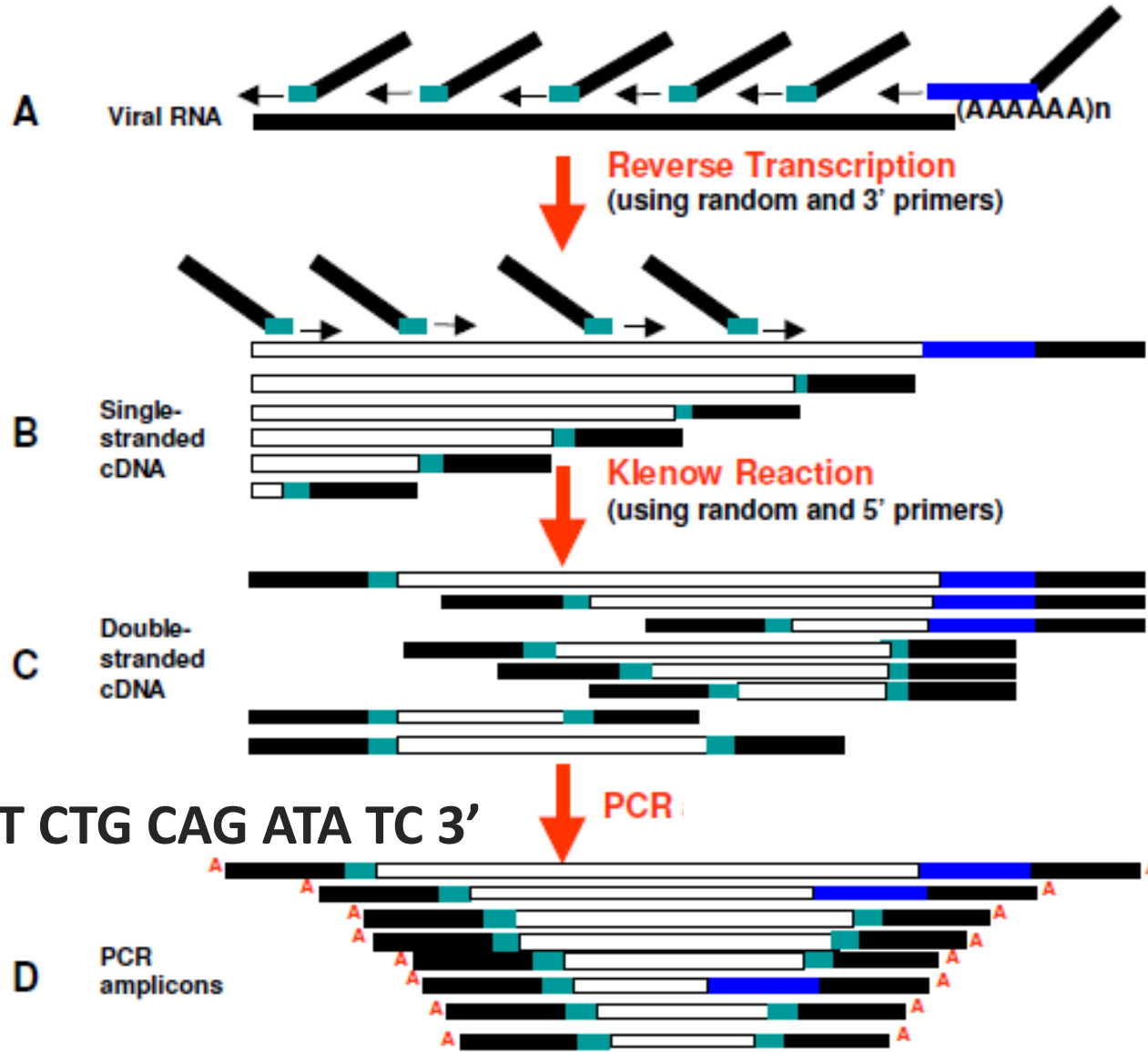


IZZAM G. CAPORALE  
TERAMO

Centro di Referenza Nazionale  
per Sequenze Genomiche di  
microorganismi patogeni:  
banca dati e analisi di Bioinfor



SISPA





Centro di Referenza Nazionale per Sequenze Genomiche di microrganismi patogeni: banca dati e analisi di Bioinformatica



# SISPA Sequencing Workflow

3 working days

70 euros per sample

6-7 h

5 h

14 h

RNA extraction

Host DNA digestion

SISPA protocol

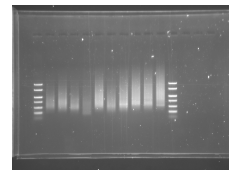
Quality & Quantity Check

Library preparation

Sequencing



RT  
Klenow  
PCR



Reads 150 bp  
Paired-end



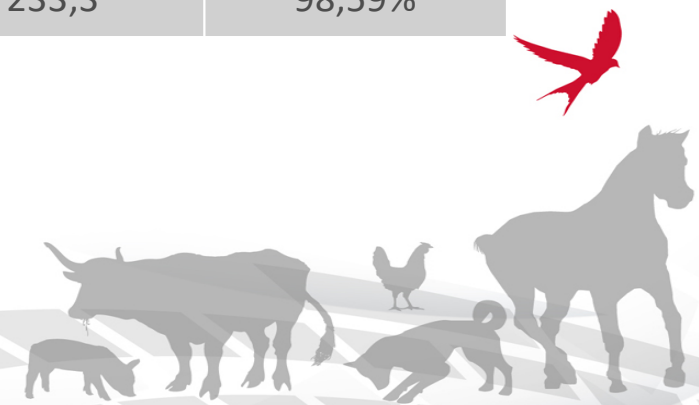


Centro di Referenza Nazionale  
per  
p  
n  
b

# Sispa results

Campione	Ct SARS-CoV-2	TOTAL_RAW READS	Q30_RAW READS	V COV	H COV
1	23	1.154.510	96,11	1,6	26,51%
2	23	922.546	95,71	13,5	91,99%
3	21	1.131.222	96,37	5,3	73,98%
4	19	993.018	94,78	555,9	99,95%
5	17	1.061.604	96,68	570,2	99,97%
6	16	1.483.578	97,07	2.875,1	99,87%
7	16	754.600	92,95	233,3	98,59%

**Quantità di virus nel campione**  
**Integrità dell'RNA virale**



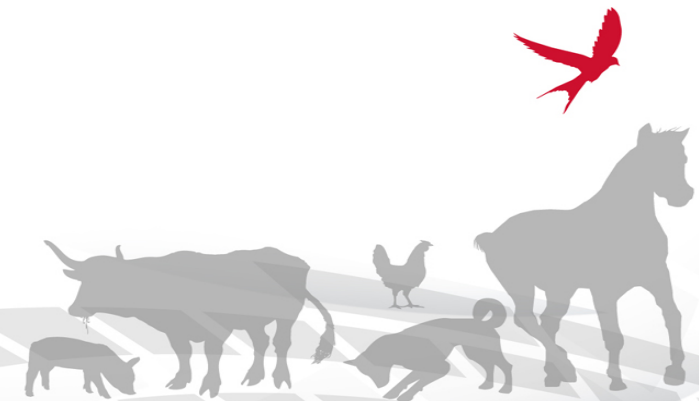


IZSAM G. CAPORALE  
TERAMO

Centro di Referenza Nazionale  
per Sequenze Genomiche di  
microorganismi patogeni:  
banca dati e analisi di Bioinformatica



- Arricchimento
- Preparativa RNA con primers specifici disegnati *ad hoc*
- Ampliseq





IZSAM G. CAPORALE  
TERAMO

Centro di Referenza Nazionale  
per Sequenze Genomiche di  
microorganismi patogeni:  
banca dati e analisi di Bioinformatica



# Confronto fra strategie di sequenziamento

12 campioni  
Ct 16-25

12 campioni  
Ct 19-34

SISPA

ARTIC

ARTIC


Swift Amplicon<sup>®</sup>  
SARS-CoV-2  
Panel kit 

7 campioni  
my Baits  
SARS-CoV-2



myBaits Expert — Predesigned Panels  
myBaits Expert Virus – SARS-CoV-2

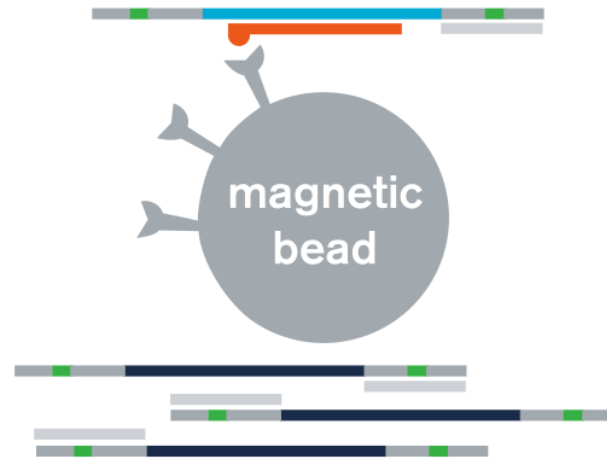
[Home](#) » [Genomics](#) » [Targeted Sequencing](#) » [myBaits – Hyb Capture Kits](#) » [myBaits Expert — Predesigned Panels](#) » myBaits Expert Virus – SARS-CoV-2

 Kit for NGS target enrichment of the whole genome of 2019 novel coronavirus (SARS-CoV-2) responsible for COVID-19 disease.

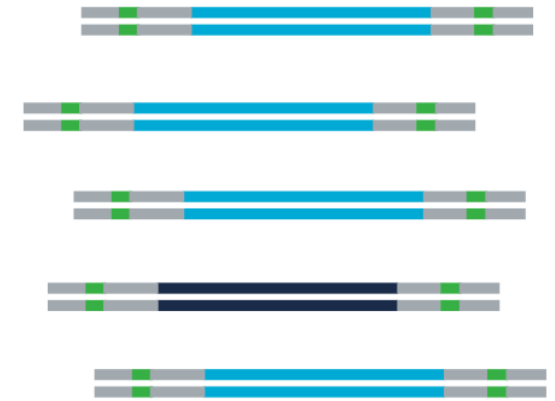
# myBaits Expert Virus-SARS-CoV-2



1 Denature library, bind to blockers and baits



2 Bind to beads, wash away off-target molecules



3 Amplify enriched library, sequence!



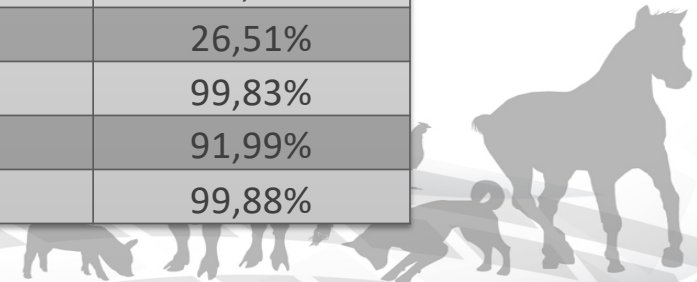


Centro di Referenza Nazionale  
per Sequenze Genomiche di  
microorganismi patogeni:  
banca dati e analisi di Bioinformatica



# SISPA vs myBaits enrichment

NRG	Ct SARS-CoV-2	Metodo	VCOV	HCOV
21400.1.2	25	SISPA	1,6	29,87%
		myBaits	332,6	99,86%
21400.1.3	25	SISPA	1,5	28,16%
		myBaits	522,6	99,87%
21400.1.4	25	SISPA	10,7	88,97%
		myBaits	3740,3	100,00%
21400.1.5	23	SISPA	5,5	69,61%
		myBaits	1998,2	99,87%
21400.1.6	22	SISPA	13,8	89,05%
		myBaits	4036,0	100,00%
21400.1.8	23	SISPA	1,6	26,51%
		myBaits	252,6	99,83%
21400.1.9	23	SISPA	13,5	91,99%
		myBaits	4485,1	99,88%





Centro di Referenza Nazionale per Sequenze Genomiche di microrganismi patogeni: banca dati e analisi di Bioinformatica



# SISPA+myBaits Sequencing Workflow

4 working days

130 euro per sample

6-7 h

5 h

14 h

RNA extraction

Host DNA digestion

SISPA protocol

Quality & Quantity Check

Library preparation

Sequencing



RT  
Klenow  
PCR



Reads 150 bp  
Paired-end

16 h



# ARTIC Protocol



## nCoV-2019 sequencing protocol V.1

Forked from Ebola virus sequencing protocol

Spanish Journal of Legal Medicine

Josh Quick<sup>1</sup>

<sup>1</sup>University of Birmingham

Version 1

Jan 22, 2020

★ Bookmark

Run

Copy / Fork

18 Works for me dx.doi.org/10.17504/protocols.io.bbmuik6w

ARTIC Coronavirus Method Development Community 1 more workspace

Josh Quick

Steps Forks Metadata Metrics

OPEN ACCESS

protocols.io



Apr 08, 2020

## nCoV-2019 sequencing protocol for illumina v.2

Forked from nCoV-2019 sequencing protocol

Kentaro Itokawa<sup>1</sup>, Tsuyoshi Sekizuka<sup>1</sup>, Masanori Hashino<sup>1</sup>, Rina Tanaka<sup>1</sup>, Makoto Kuroda<sup>1</sup>

<sup>1</sup>National Institute of Infectious Diseases, Japan

1 Works for me dx.doi.org/10.17504/protocols.io.betejeje

Coronavirus Method Development Community

Kentaro Itokawa  
National Institute of Infectious Diseases, Japan

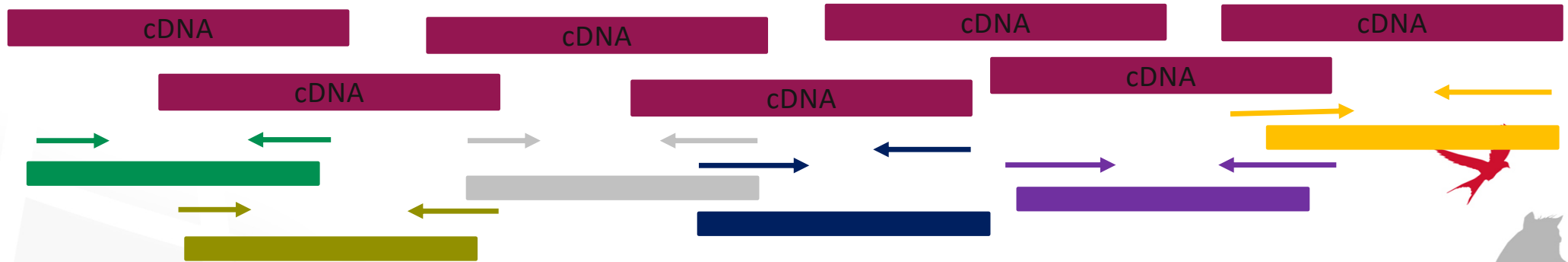
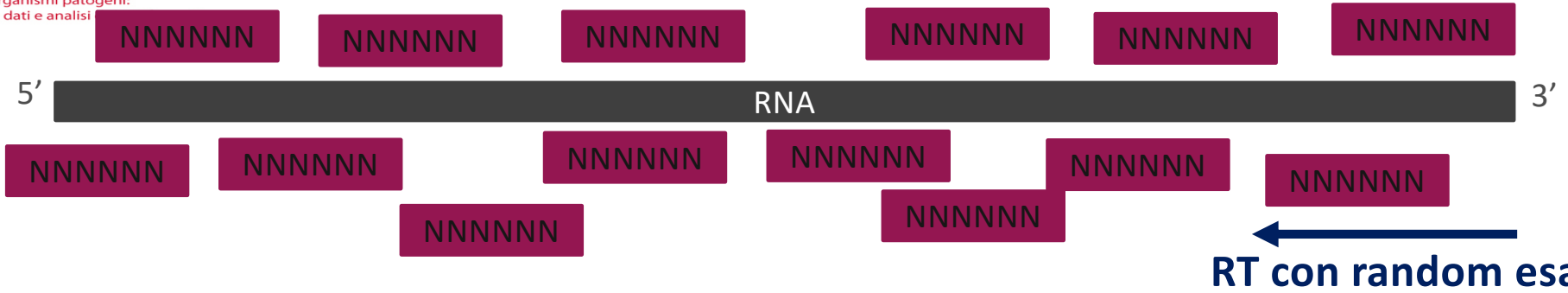
### ABSTRACT

This protocol is forked from "[ARTIC amplicon sequencing protocol for MinION for nCoV-2019](#)" by Josh Quick to adapt it to **illumina sequencers**.

Because the PCR products are fragmented and ligated with adapters, this protocol is not restricted to 250 PE mode of MiSeq.

While the library preparation uses QiaSeq FX by Qiagen and is basically straight forward (as per kit instruction but set to 1/4 scale), some tweaks for much of simplicity and speed were added.

# Targeted protocol



PCR con primers SARS-CoV-2 specifici





Centro di Referenza Nazionale  
per Sequenze Genomiche di  
microorganismi patogeni:  
banca dati e analisi di Bioinformatica



# ARTIC Sequencing Workflow

3 working days

65 euros pr sample

4 h

5 h

14 h

RNA  
extraction

ARTIC protocol

Quantity Check

Library  
preparation

Sequencing



RT random  
PCR SARS-CoV-2  
primers



Reads 150 bp  
Paired-end



NRG	Ct SARS-CoV-2	Metodo	VCOV	HCOV
<b>21400.1.2</b>	25	myBaits	332,6	99,86%
		ARTIC	1546,9	99,18%
<b>21400.1.3</b>	25	myBaits	522,6	99,87%
		ARTIC	526,3	99,32%
<b>21400.1.4</b>	25	myBaits	3740,3	100,00%
		ARTIC	2425,3	99,99%
<b>21400.1.5</b>	23	myBaits	1998,2	99,87%
		ARTIC	1442,9	99,93%
<b>21400.1.6</b>	22	myBaits	4036,0	100,00%
		ARTIC	1470,1	99,97%
<b>21400.1.8</b>	23	myBaits	252,6	99,83%
		ARTIC	686,3	98,67%
<b>21400.1.9</b>	23	myBaits	4485,1	99,88%
		ARTIC	1208,4	99,92%

NRG	Ct SARS-CoV-2	Metodo	VCOV	HCOV
<b>21400.1.7</b>	19	SISPA	555,9	99,95%
		ARTIC	4133,5	99,98%
<b>21400.1.10</b>	16	SISPA	2875,1	99,87%
		ARTIC	4361,4	100,00%
<b>21400.1.11</b>	17	SISPA	570,2	99,97%
		ARTIC	3201,7	100,00%
<b>21400.1.12</b>	16	SISPA	233,3	98,59%
		ARTIC	1439,1	99,91%
<b>21400.1.13</b>	21	SISPA	5,3	73,98%
		ARTIC	2245,5	99,93%

Metodo	241 C	3037 C	14408 C	23404 A	28881
SISPA					AAC
myBaits	T	T	T	G	AAC
ARTIC	T	T	T	G	AAC
SISPA					
myBaits	T	T	T	G	AAC
ARTIC	T	T	T	G	AAC
SISPA	T	T	T	G	AAC
myBaits	T	T	T	G	AAC
ARTIC	T	T	T	G	AAC
SISPA			T	G	AAC
myBaits	T	T	T	G	AAC
ARTIC	T	T	T	G	AAC
SISPA	T		T	G	AAC
myBaits	T	T	T	G	AAC
ARTIC	T	T	T	G	AAC
SISPA			T	A	
myBaits	T	T	T	G	RRS
ARTIC	T	T	T	G	GGG

## Varianti SARS-CoV-2 Europa

Metodo	241 C	3037 C	14408 C	23404 A	28881
SISPA			T	G	AAC
myBaits	T	T	T	G	AAC
ARTIC	T	T	T	G	AAC
SISPA	T	T	T	G	AAC
ARTIC	T	T	T	G	AAC
SISPA	T	T	T	G	GGG
ARTIC	T	T	T	G	GGG
SISPA	T	T	T	G	AAC
ARTIC	T	T	T	G	AAC
SISPA		T	T	G	GGG
ARTIC	T	T	T	G	GGG
SISPA			T	G	AAC
ARTIC	T	T	T	G	AAC

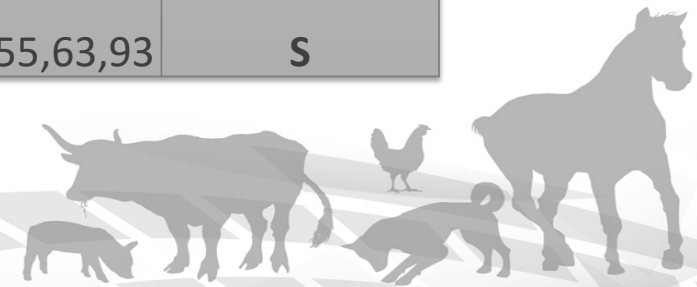


Centro di Referenza Nazionale  
per Sequenze Genomiche di  
microrganismi patogeni:  
banca dati e analisi di Bioinformatica



# myBaits IUPAC pos 28881-2883

NRG	POS	REF	ALT	PERC	n. READS	IUPAC
2020.TE.21400.1.8	28881	G	A	29,40%	123,254,64,93	R
2020.TE.21400.1.8	28882	G	A	29,64%	120,255,64,94	R
2020.TE.21400.1.8	28883	G	C	29,43%	119,255,63,93	S







IZSAM G. CAPORALE  
TERAMO

Centro di Referenza Nazionale  
per Sequenze Genomiche di  
microrganismi patogeni:  
banca dati e analisi di Bioinformatica



# Campioni Ct 19-34

## ARTIC vs Swift



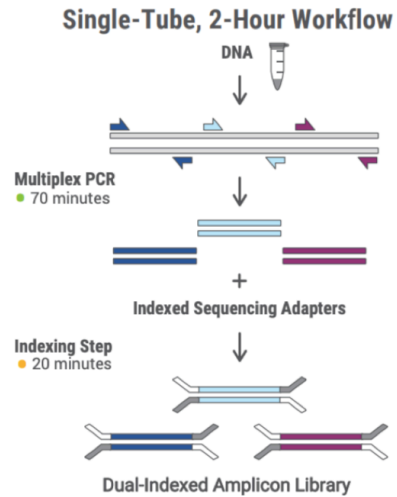


Centro di Referenza Nazionale per Sequenze Genomiche di microrganismi patogeni: banca dati e analisi di Bioinformatica

# Swift Amplicon SARS-CoV-2 Reasearch panel

## 2 working days


## 67 euros per sample



### Reads 150 bp Paired-end






 Centro di Referenza Nazionale  
 per Sequenze Genomiche di  
 microrganismi patogeni:  
 banca dati e analisi di Bioinformatica


# Varianti SARS-CoV-2 Europa

NRG	Ct SARS_CoV2	Metodo	VCOV	HCOV	241	3037	14408	23404	28881
33690.1.3	23	ARTIC	5694,55	99,88%	T	T	T	G	AAC
		SWIFT	4093,51	98,91%	T	T	Y	G	AAC
33690.1.4	33	ARTIC	535,569	65,68%	T			G	AAC
		SWIFT	9,39666	58,44%	Y				
33690.1.5	33	ARTIC	512,911	65,44%					
		SWIFT	8,84501	54,05%					
33690.1.6	30	ARTIC	342,665	41,37%	T	T	T		
		SWIFT	8,03359	52,37%	T	T		G	GGG
33690.1.7	31	ARTIC	4555,97	99,80%	T	T	T	G	GGG
		SWIFT	1741,54	98,85%	T	T	N	G	GGG
33690.1.8	28	ARTIC	6163,35	99,90%	T	T	T	G	RRS
		SWIFT	898,582	97,96%	T	T	Y	G	GGG






IZSAM G. CAPOREALE  
TERAMO


 Centro di Referenza Nazionale  
 per Sequenze Genomiche di  
 microrganismi patogeni:  
 banca dati e analisi di Bioinformatica

# Varianti SARS-CoV-2 Europa

NRG	Ct SARS_CoV2	Metodo	VCOV	HCOV	241	3037	14408	23404	28881
33690.1.9	34	ARTIC	1808,55	97,55%	T	T	T	G	GGG
		SWIFT	13,9966	66,40%	Y	T			
33690.1.10	33	ARTIC	301,295	56,85%	T			G	AAC
		SWIFT	9,38229	53,82%	T	T			
33690.1.11	27	ARTIC	5063,37	99,87%	T	T	T	G	AAC
		SWIFT	381,791	95,58%	T	T	Y	G	AAC
33690.1.12	19	ARTIC	5501,44	99,90%	T	T	T	G	GGG
		SWIFT	3111,42	99,47%	T	T	Y	G	GGG
33690.1.13	30	ARTIC	5074,36	99,86%	T	T	T	G	AAC
		SWIFT	642,996	96,77%	T	T	Y	G	AAC
33690.1.14	24	ARTIC	5197,05	99,86%	T	T	T	G	GGG
		SWIFT	2664,25	98,82%	Y	T			




 Centro di Referenza Nazionale  
 per Sequenze Genomiche di  
 microrganismi patogeni:  
 banca dati e analisi di Bioinformatica



# Swift IUPAC pos 241, 14408

NRG	POS	REF	ALT	PERC	n. READS	IUPAC
2020.TE.33690.1.4	241	C	T	84,61%	2,0,6,5	Y
2020.TE.33690.1.9	241	C	T	72,22%	4,1,7,6	Y
2020.TE.33690.1.14	241	C	T	66,66%	1,0,1,1	Y
2020.TE.33690.1.2	14408	C	T	32,96%	120,0,21,38	Y
2020.TE.33690.1.3	14408	C	T	22,99%	278,0,31,52	Y
2020.TE.33690.1.8	14408	C	T	17,39%	19,0,2,2	Y
2020.TE.33690.1.11	14408	C	T	14,28%	12,0,1,1	Y
2020.TE.33690.1.12	14408	C	T	32,22%	204,0,26,71	Y
2020.TE.33690.1.13	14408	C	T	36,36%	14,0,4,4	Y



Centro di Referenza Nazionale  
per Sequenze Genomiche di  
microorganismi patogeni:  
banca dati e analisi di Bioinformatica



# ARTIC IUPAC pos 28881-28883

NRG	POS	REF	ALT	PERC	n. READS	IUPAC
2020.TE.33690.1.8	28881	G	A	42,44%	2041,2119,1518,1549	R
2020.TE.33690.1.8	28882	G	A	42,41%	2036,2121,1517,1544	R
2020.TE.33690.1.8	28883	G	C	42,25%	2008,2184,1518,1549	S





IZSAM G. CAPOREALE  
TERAMO

Centro di Referenza Nazionale  
per Sequenze Genomiche di  
microrganismi patogeni:  
banca dati e analisi di Bioinformatica



# SISPA vs myBaits vs ARTIC vs Swift...

## chi vince?





Centro di Referenza Nazionale per Sequenze Genomiche di microrganismi patogeni: banca dati e analisi di Bioinformatica



# ARTIC

3 working days  
65 euros pr sample

4 h

5 h

14 h

RNA extraction

ARTIC protocol

Quantity Check

Library preparation

Sequencing



RT random PCR SARS-CoV-2 primers



Reads 150 bp  
Paired-end







■ ■ Centro di Referenza Nazionale  
per Sequenze Genomiche di  
microorganismi patogeni:  
banca dati e analisi di Bioinformatica



# Cosa manca?

- **Analisi varianti su tutta la sequenza genomica**
- **Diversi tool (Genpat, Galaxy, Lasergene)**





■ ■ Centro di Referenza Nazionale  
per Sequenze Genomiche di  
microrganismi patogeni:  
banca dati e analisi di Bioinformatica



# Ringraziamenti

- Gruppo COVID
- Virologia e colture cellulari
- Biologia molecolare e tecnologie omiche
- Bioinformatica

