From DNA to actionable result. Right now





Exatype: automated SARS-CoV-2 variant typing directly from NGS sequence data

The identification of SARS-CoV-2 variants that exhibit increased transmissibility and may impact vaccine efficacy has accelerated the need for global genomic surveillance of SARS-CoV-2.

While DNA sequencing provides the most comprehensive approach for characterising variants, the analysis and interpretation of sequence data at scale remains a challenge.

Hyrax Biosciences' Exatype platform provides fast, automated SARS-CoV-2 lineage/clade typing.

### **Exatype features:**

- Reports detected variants using both Pangolin lineages and Nextstrain clades nomenclature
- Analyses data from a range of sequencing platforms and can process data generated from any amplification protocol/assay
- Requires no expert training and no software needs to be installed
- Easily scaled with **parallel** processing in the cloud or through direct **integration** with sequencers and/or LIMS
- Is secure and complies with all necessary HIPAA requirements
- Uses algorithms with proven superior accuracy in read mapping and variant calling in viral data



# Once you've set up an account, analysing your data is simple:

#### 1. Name your job:

	PE				Simon Travers 🗸	*	EXATYP SARS-Cov-2	E			
Job Details	Sequencing Platform	BED file Select Data	Sample Details Review &	Submit		>>	Job Details	Sequencing Platform	BED file	Select Data	Sample I
Cre	eate a new job			NEXT >	CANCEL JOB	*	Crea	ate a new job			
i Job	Details					i i	Seque	encing Platform			
Job Na	oData		~			ہ د	0 to 0 til	n Torrent umina			
	talls (Optional- for your refer Details (Optional- for your ref (T )	ence)					Are you s	submitting paired-end or si aired-end EV NEXT >	ngle-ended d	ta? Paired-end	data requit
A Fo	r Research Use Only. Not for	use in diagnostic procedure	15.				A For F	Research Use Only. Not for	use in diagn	ostic procedure:	s.
© Hyra	ax Biosciences Terms of use	Privacy policy FAQs in	fo@hyraxbio.co.za Logout								

#### 2. Select the sequencing platform:

( PREV NEXT )

#### 3. Select the assay used to generate the data:

#### 4. Drag and drop your fastq files into the browser:

	*	EXATYPE SARS-Cov-2								Simon Traver	ers 🗸			Ξ**						Simon Travers 🗸
	<b>&gt;&gt;</b>	Job Details	Sequencing Platform	BED file	Select Data	Sample Detai	ails Revie	w & Submit				>> Job D	etails	Sequencing Platform	BED file	Select Data	Sample Details	Review & Submit		
The control of th	n R di	Crea	te a new job					< PREV	NEXT >	CANCEL JO	DB	<ul> <li>▲</li> <li>↓</li> <li>↓</li></ul>	Crea	ite a new job	)			<pre>&lt; PREV</pre>	NEXT >	CANCEL JOB
Partice the type too bus due type too bus	i o	If you hav	e generated your data usi	ing an amplico	n-based protoc ently listed, ple	col, click "Yes" be	elow and sele	ect the protoco	ol used from the			i ~	lease sel	lect the data files that yo	u would like t	o analyse. Only	gz fastq and	fq files are accepted.		
Select BCD format       *         Artic V1       Artic V2         Artic V2       Control Contro	L	Please sel	ect the protocol used fro	m the menu be	elow. If you hav	e used a protocol	ol not current	tly listed, pleas	se contact us at			د [		СНО	DOSE YOUR	SEQUENCE I	DATA FILES			
Arris VI Arris		Select Bl	ED format	~						>	*									
ATTE V3		ARTIC	V1 V2											OR DRAG AN	D DROP YC	OUR SEQUENC	CE DATA FILES H	20 Sample001_R1_001.f Sample001_R2_001.f Sample002_R1_001.f Sample002_R2_001.f	astq.gz astq.gz astq.gz astq.gz	
C Research Use Only. Not for use indiagnostic procedures.		ARTIC	V3		-													Sample003_R1_001.f Sample003_R2_001. Sample004_R1_001.f Sample004_R2_001	astq.gz fastq.gz astq.gz fastq.gz	
▲ For Research Use Only, Not for use in diagnostic procedures.													( PRE	V NEXT >				Sample005_R2_001	astq.gz astq	
▲ For Research Use Only, Not for use in diagnostic procedures.																				
▲ For Research Use Only, Not for use in diagnostic procedures.																				
▲ For Research Use Only, Not for use in diagnostic procedures.																				
▲ For Research Use Only, Not for use in diagnostic procedures.																				
▲ For Research Use Only, Not for use in diagnostic procedures.																				
For Research Use Only. Not for use in diagnostic procedures.																				
For Research Use Only. Not for use in diagnostic procedures.																				
		A For Re	esearch Use Only. Not fo	r use in diagno	ostic procedure	25.							For R	lesearch Use Only. Not f	or use in diag	nostic procedur	es.			
C Hyrax Biosciences Terms of use Privacy policy FAQs info@tyraxbio.co.za Lagout C Hyrax Biosciences Terms of use Privacy policy FAQs info@tyraxbio.co.za Lagout		© Hyrax B	iosciences Terms of use	Privacy pol	icy FAQs in	ifo@hyraxbio.co.z	.za Logout					C	Hyrax B	Biosciences Terms of us	e Privacy p	olicy FAQs in	nfo@hyraxbio.co.za	Logout		

5. Exatype automatically groups files for paired-end data and extracts sample names (you can edit these manually if you wish):

lob De	tails Sequencing Pla	atform	BED file	Select Data	Sample Details	Review & Submit		
Cr	eate a new j	ob				< PREV NE	хт )	CANCEL JOB
_								
San	nple Details							
San	nple details   Illi	umina						
	• •							
Sam	ples in this analys	is					REMO	OVE ALL SAN
	Sample ID		File na	me			Size	File status
-	5		Sample	001_R1_001.fa	stq.gz	0%	41.42 MB	Selected
	SampleUU1	A.B.	Sample	001_R2_001.fa	stq.gz	0%	42.18 MB	Selected
-	Sample002		Sample	002_R1_001.fa	stq.gz	0%	50.92 MB	Selected
	Sampleooz	74	Sample	2002_R2_001.fa	stq.gz	0%	51.11 MB	Selected
+	Sample003		Sample	003_R1_001.fa	stq.gz	0%	22.87 MB	Selected
	Jampieooo	<i>w</i>	Sample	:003_R2_001.fa	stq.gz	0%	23.48 MB	Selected
÷.	Sample004	1	Sample	004_R1_001.fa	stq.gz	0%	10.42 MB	Selected
		P	Sample	:004_R2_001.fa	stq.gz	0%	11.14 MB	Selected
<b>金</b>	Sample005	ample002 // ample003 // ample004 // ample005 // ample005	Sample	005_R1_001.fa	stq.gz	0%	35.05 MB	Selected
-			Sample	:005_R2_001.fa	stq.gz	0%	39.62 MB	Selected
<b>金</b>	les in this analysis Sample001 // Sample002 // Sample003 // Sample004 // Sample005 // Sample006 //	Sample	2006_R1_001.fa	stq	0%	121.81 MB	Selected	
			Sample	2006_R2_001.fa	stq	0%	121.74 MB	Selected
會	Sample007		Sample	2007_R1_001.fa	stq	0%	127.38 MB	Selected
			Sample	007_R2_001.fa	stq	0%	127.39 MB	Selected
童	Sample008	1	Sample	000 P2 001 6	stq.gz	0%	10.04 MB	Selected
			Sample	000_R2_001.1a	stq.gz	0%	20.07 MB	Selected
盲	Sample009	. Martin	Sample	009 R2 001.fa	sta.ez	0%	31.65 MB	Selected
-			Sample	010 R1 001.fa	sta.gz	0%	54.26 MB	Selected
盲	Sample010		Sample	010 R2 001.fa	sta.gz	0%	58.62 MB	Selected
-		<b>1</b>						
< 1	PREV NEXT >							
_								

6. That's it! Click upload and walk away. You will receive an email once your analysis is complete:

						_
Cr	eate a new job			l	< PREV	CANCEL
Rev	view & Submit					
Job	overview					
Job n	name: DemoData				Numb	per of samples
Sequ	encing platform: Illumina				Data stru	cture: Paired-
Iotai	rdata quantity: 1.02 GB				Preva	ience cutoir: 1
E 10	20.61		105 00 MD/4 00 CD			
5/2	to files		05.08 MB/1.02 GB		_	
_					_	
		Sample001_R1_001.fast	q.gz	100%	41.42 MB	Uploaded
	Sample001	Sample001_R2_001.fast	q.gz	100%	42.18 MB	Uploading
		Sample002_R1_001.fast	q.gz	73%	50.92 MB	Uploading
	Sample002	Sample002_R2_001.fast	q.gz	100%	51.11 MB	Uploaded
	0	Sample003_R1_001.fast	q.gz	100%	22.87 MB	Uploaded
Т	Sample003	Sample003_R2_001.fast	q.gz	89%	23.48 MB	Uploading
		Sample004_R1_001.fast	q.gz	100%	10.42 MB	Uploaded
~	Sample004	Sample004_R2_001.fast	q.gz	100%	11.14 MB	Uploaded
		Sample005_R1_001.fast	q.gz	89%	35.05 MB	Uploading
	Sample005	Sample005_R2_001.fast	q.gz	67%	39.62 MB	Uploading
1					121.81 MB	Uploading
1	a	Sample006_R1_001.fast	9	14%		
1	Sample006	Sample006_R1_001.fast	q q	14%	121.74 MB	Uploading
1	Sample006	Sample006_R1_001.fast Sample006_R2_001.fast Sample007_R1_001.fast	9 9 9	9%	121.74 MB 127.38 MB	Uploading Uploading
1 1 1	Sample006 Sample007	Sample006_R1_001.fast Sample006_R2_001.fast Sample007_R1_001.fast Sample007_R2_001.fast	9 9 9	14% 9% 8% 8%	121.74 MB 127.38 MB 127.39 MB	Uploading Uploading Uploading
1 1 1	Sample006 Sample007	Sample006_R1_001.fast Sample006_R2_001.fast Sample007_R1_001.fast Sample007_R2_001.fast Sample008_R1_001.fast	q q q q q.gz	14% 9% 8% 8% 100%	121.74 MB 127.38 MB 127.39 MB 10.04 MB	Uploading Uploading Uploading Uploading
1 1 1	Sample006 Sample007 Sample008	Sample006_R1_001.fast Sample006_R2_001.fast Sample007_R1_001.fast Sample007_R2_001.fast Sample008_R1_001.fast Sample008_R2_001.fast	q q q q q.gz q.gz	14% 9% 8% 8% 100% 97%	121.74 MB 127.38 MB 127.39 MB 10.04 MB 10.25 MB	Uploading Uploading Uploading Uploading Uploading
1 1 1 1	Sample006 Sample007 Sample008	Sample006_R1_001.fast Sample006_R2_001.fast Sample007_R1_001.fast Sample007_R2_001.fast Sample008_R1_001.fast Sample008_R2_001.fast Sample009_R1_001.fast	q q q q q q.gz q.gz q.gz	14% 9% 8% 100% 97% 36%	121.74 MB 127.38 MB 127.39 MB 10.04 MB 10.25 MB 29.07 MB	Uploading Uploading Uploading Uploading Uploading Uploading
1 1 1 1	Sample006 Sample007 Sample008 Sample009	Sample006_R1_001fast Sample006_R2_001fast Sample007_R1_001fast Sample008_R1_001fast Sample008_R1_001fast Sample009_R2_001fast Sample009_R2_001fast	9 9 9 9 9 9 8 2 8 2 8 2 8 2 8 2 8 2 8 2	14% 9% 8% 100% 297% 36%	121.74 MB 127.38 MB 127.39 MB 10.04 MB 10.25 MB 29.07 MB 31.65 MB	Uploading Uploading Uploading Uploading Uploading Uploading Uploading
1 1 1 1	Sample006 Sample007 Sample008 Sample009	Sample006_R1_001.fash Sample006_R2_001.fash Sample007_R1_001.fash Sample007_R2_001.fash Sample008_R1_001.fash Sample008_R2_001.fash Sample009_R2_001.fash Sample009_R2_001.fash Sample000_R2_001.fash	9 9 9 9 8 8 2 8 2 8 2 8 2 8 2 8 2 8 2 8	14% 9% 8% 100% 97% 36% 33%	121.74 MB 127.38 MB 127.39 MB 10.04 MB 10.25 MB 29.07 MB 31.65 MB 54.26 MB	Uploading Uploading Uploading Uploading Uploading Uploading Uploading Uploading

## Outputs

1. Explore per-sample results through the front-end:



# 2. Exatype uses both Pangolin lineages (cov-lineages.org) and Nextstrain clades (nextstrain.org) and outputs a table showing calls from each approach. It includes amino acid substitutions and deletions observed in each sample.

Sample	Pangolin (2.1.10) <sup>1</sup>	Nextstrain clade (0.12.0) <sup>2</sup>	Nextstrain AA substitutions (0.12.0) <sup>2</sup>	Nextstrain AA deletions (0.12.0) <sup>2</sup>
Sample003	B.1.338	✓ 20C	ORF1a:       72651, A2129V         ORF1b:       7314L         S:       D614G         ORF3a:       Q57H	None detected
Sample005	B.1	▲ 20C	ORF1a:       72651, 11206K         ORF1b:       7314L, 52430R         S:       D614G         ORF3a:       Q57H	ORF1a: P1207, K1208, E1209, E1210, V1211, K1212-
Sample011	P.1	20J/501YV3     Sample011     Missing data: Good	ORF1a:         \$1188L, K1795Q, \$2947N, A3523V           ORF1b:         \$314L, £1264D           S:         £18F, T20N, P265, D138Y, R190S, K417T, E484K, N501Y, D614G, H655Y, T1027I, V1176F           ORF3a:         \$253P           F8:         E92K	ORF1a: \$3675-, G3676-, F3677-
<sup>1</sup> https://github.c <sup>2</sup> https://clades.n	om/cov-lineages/pangolii extstrain.org/	<ul> <li>Missing data: Good</li> <li>Mixed sites: Good</li> <li>Private mutations: Good</li> <li>SNP clusters: Good</li> </ul>	bd	

#### 3. QC plots provide a valuable oversight of data quality:



4. A wide range of outputs in standardised formatting are available for download for further analysis or for publication of outputs to public repositories e.g. consensus sequences can be uploaded to GISAID:

	HYRAX BIOSCIENCES
	Downloads ^
	All consensus sequences from SARS-CoV-2 samples
	Combined 210120_Demo_consensus_sequences.sars-cov-2.fasta
	One-per-file 10120_Demo_aligned_consensus_sequences.zip
	All variants in SARS-CoV-2 samples present at >= 10% prevalence
	3 210120_Demo_prevalences.zip
	Lineage/Clade files
l	1210120_Demo_nextclade.csv
L	1210120_Demo_nextclade.json
L	210120_Demo_nextclade.tree
l	10120_Demo_pangolin.csv

## How to get in touch

Visit <u>sars-cov-2.exatype.com</u> to sign up for a free account and upload your first sample. Contact us at <u>info@hyraxbio.com</u> to discuss how we can provide companion software for your needs.