

Pillar[®] Biosciences PiVAT[®] User Manual

Software Version 2023.1.0

Released June 2023

Legal Notice

For Research Use Only. Not for use in diagnostic procedures.



Table of Contents

Introduction	4
Software Requirements	
Symbols Used in this User Manual:	
Connect to PiVAT	
Accept Terms and Conditions	6
Using the PiVAT Software	7
Dashboard	
Navigation Bar	7
Setup an Analysis	
Creating Analysis Parameters	
Start Analysis	
Monitoring the Task Status	
View Summary of Task Details	
View Completed Analysis Task	
View Analysis Results	
User password	
Data Management	
SFTP Download	
Troubleshooting	29
Legal Notices	31
Company Information	31
Appendix A: SNV/Indels Caller	32
SNV/Indels Analysis Parameters Setup	
PiVAT Output: CUSTOMER_RESULTS file	
Appendix B: Microsatellite Instability (MSI) Caller	40
MSI Sample Setup	
MSI Analysis Setup	
PiVAT Output: MSI_RESULTS file	
Appendix C: Somatic CNV Caller	44
CNV/ Sample Setur	ЛЛ
CNV Analysis Setup	
CNV Analysis Setup	
cfDNA_CNV_Caller_OC_parameters:	
PiVAT Output: CNV RESULTS file	
	5/
CNV Eiltering	
CNV Plot	
CNV Panel Specific Notes	
Appendix D: Thalassemia CNV Caller	60



Thalassemia CNV Sample Setup Thalassemia CNV Analysis Setup Thalassemia CNV QC parameters PiVAT Output: THAL_RESULTS file Thalassemia Type Calls Table Thalassemia Plot	60 60 60 61 62 63
Appendix E: SMA Caller	64
SMA Sample Setup SMA Analysis Setup SMA QC parameters PiVAT Output: SMA_RESULTS file. SMA Calls Table SMA Plot	64 64 64 65 65 66 67
Appendix F: Fusion Caller	69
Fusion Sample Setup Fusion Analysis Parameters Setup PiVAT Output: CUSTOMER_RESULTS file	69 69 70
Appendix G: FASTQ File Name Format	72



Introduction

This user manual describes how to use Pillar Bioscience's (Pillar) PiVAT software. PiVAT is Pillar's genome sequence data software.

Software Requirements

The PiVAT software requires a supported browser to access the user interface. The current versions of the following browsers are supported:

- Google Chrome
- Mozilla Firefox

Symbols Used in this User Manual:



This symbol is a reminder to pay attention to details that could affect either proper installation or performance.



This symbol is an indication of a critical detail not to be overlooked.



Connect to PiVAT

1. Login to PiVAT using the username and password provided by Pillar Biosciences.

P	LLAR AT (Pillar Variant Analysis T	∿≊ ^{English}
Next (Generation Sequencing for Research Use Or	hly
	Username	
	Password	
	Login	
	About Privacy Terms & Conditions Contact Software	List
Disclaimer The information provided by Pilla make no represe reliability, suitab graphics contair strictly at your o	contained in this application is for general information purposes only. The r Biosciences Inc. and while we endeavour to keep the information up to intations or warranties of any kind, express or implied, about the complete lity or availability with respect to the website or the information, products ed in the application for any purpose. Any reliance you place on such infor wn risk.	e information is date and correct, we eness, accuracy, s, services, or related rmation is therefore

Figure 1 PiVAT login page



Accept Terms and Conditions

- 1. The following screen is displayed upon user's first login.
- 2. Read the terms and conditions and click the Agree button to continue.

	DSCIENCES	
PiVAT	(Pillar Variant Analysis To	olkit)
Next Gen	eration Sequencing for Research Use Only	
	usermanu12	
	PILLAR BIOSCIENCES, INC. ("Pillar") grants vou a non-	
	exclusive license to use the software and any associated	
	documentation ("Software") as indicated below.	
A	IMPORTANT NOTICE. Read this License Agreement carefully	t
	before downloading or using the Software. BY	
	YOU ACKNOWLEDGE THAT YOU HAVE READ, UNDERSTAND	
Disclaimer	AND AGREE TO THE TERMS OF THIS LICENSE AGREEMENT. IF	formation is
provided by Pillar Bio	YOU DO NOT AGREE TO THESE TERMS, DO NOT DOWNLOAD	e and correct, we
make no representat	AND PROMPTLY DELETE OR DESTROY ANY COPIES OF THE	ss, accuracy,
reliability, suitability	SOFTWARE IN YOUR POSSESSION. YOU ARE SOLELY	rvices, or related
graphics contained i	RESPONSIBLE FOR THE PROTECTION OF ANY PERSONALLY	ation is therefore
strictly at your own i	By clicking 'Agree' you agree to the Terms & Conditions	
	Cancel Agree	

Figure 2 PiVAT License Agreement



Using the PiVAT Software

Dashboard

1. User may navigate through the application using the Dashboard screen.

PILLAR Dashboard An	alysis Tool	s				(usermanual 🔻	🕱 English
System Tasks In Queue	0	Your Tasks Running / In Queue / Total	0/0/0	Average Task Analysis Time	0 H	Data Storage disk usage / quo	ta	0.0 G /100.0G
Task Number Monthly Trend (the la	ast 6 months) Q No c	lata yet!		Task Number Per Panel	Q No	data yet!		
Recent Task Activity								More Tasks >
ID	Task		Panel		Date	SA Status	Stage 🚺	Operation
			No analysis	task found				
DUAT EVID-surgice 2023 1.0 selected on 2023 0.4	28						0 Biller	Biscience Inc 2015 2023
FINAL INFO VERSION 2023-1.0 Teledised on 2023-04-2	20						© Pillar	prosulences iniC 2010-2025

Figure 3 PiVAT Dashboard page

Navigation Bar

- 1. The following Navigation Bar is displayed on the top of PiVAT Dashboard page. Additional items are displayed for admin users and detailed in the PiVAT Admin Guide document.
- 2. The page being displayed will be highlighted on the Navigation Bar.
- 3. Logged in user is indicated next to $\stackrel{\bullet}{=}$ icon.
- 4. PiVAT software is enabled for English and Chinese languages.



Figure 4 Navigation Bar on PiVAT Dashboard



Dashboard: Returns the user to the Dashboard screen

PILLAR Dashboard Analysis Too	ols					usermanual 🔹	🏂 English 🕜 Help
System Tasks In Querue	Your Tasks Running / In Queue / Total	0/0/0	Average Task Analysis Time	0н	Data Storage disk usage / quot	ta	0.0 G /100.0G
Task Number Monthly Trend (the last 6 months)	data yet!		Task Number Per Panel	Q NO	data yet!		
Recent Task Activity							More Tasks >
ID Task		Panel		Date	SA Status	Stage 🕕	Operation
		No analysis f	task found				
8047 810 series 2021 1 0 relevant on 2023 Ab 28						C Piler	Biosciences Inc 2015-2028
Priviti NUU Version 2023-1.0 released on 2023-04-28						C Mar	biosciences inc 2015-2023

Figure 5 PiVAT Dashboard page

Analysis: Displays the screen for creating, starting, and viewing analysis tasks on the left menu bar include:

- **Parameters**: Used to create a custom set of secondary analysis parameters for running analysis (optional section for advanced users).
- Start Analysis: Most users will start here after uploading data in the Data Management section.
- Analysis Results: Detailed view of queued and completed analysis tasks.

PILLAR Dashboard	Analysis Tools					🖸 usermanual 👻 🖥	🕻 English 🕜 Help
✿ Parameters	🌴 / Analysis						
🗱 Start Analysis	Analysis Guide						
₩ Analysis Results	1 Create Custom Paramete	rs It's an optional step				Creat	e New Parameters
	ID	Parameters File Name	Pane	el		Date	Operation
			No parameters file found				
	2 Start An Analysis						
	Go To Start Analysis						
	3 View Analysis Results						More Tasks >
	ID	Task	Panel	Date	SA Status	Stage 🚺	Operation
			No analysis task found				
<							

Figure 6 PiVAT Analysis page



Tools: Displays the screens for Integrative Genomics Viewer (IGV), Data Management and Download on the left menu bar.

- IGV: Choose *bam* and *bai* files from task server, and enter the locus you want to focus on.
- **Data Management**: Sample data files can be uploaded to this section for use as inputs to the analysis tasks. The Project Name corresponds to a folder with the same name in the uploads directory. Sample files used for a task can hence be uploaded to a Project and be used as a bundled input for the Analysis Task.
- **Download**: Select and download panel and analysis parameters in bed format.

PILLAR Dashboard	Analysis Tools	🧿 usermanual 👻 🧏 English 🕜 Help
ez IGV	🐐 / Tools / IGV	
💷 Data Management	IGV Choose barn and bai files from task server, and enter the locus you want to focus	
Download	Locus: ① all v Location Start: Location End: @scalor formet: 11555079	
	Home /	Filter
	Definition Name Size	Modified Date
	🍘 input	2023-02-23 11:41:51
	🐚 user	2023-06-01 14:11:30
	Reset	
(Your Selected Items	: 0
	PIVAT RUO version 2023.1.0 released on 2023-04-28	© Pillar Biosciences Inc 2015-2023

Figure 7 PiVAT Tools page



Help: Displays product information, instructions for SFTP, and a downloadable User Manual in PDF format .

倄 / Help

Help Retrieve Product Documentation

Versions

Pivat 2023.1.0 Analysis Pipeline 2023.1.0

PiVAT User Manual

Download User Manual

*If you want to learn about how to use this software, please download our 'User Manual'.

SFTP Download Instructions

To download files using SFTP (Secure File Transfer Protocol), your installation of PiVAT must have the SFTP feature enabled by the administrator. Please contact your administrator to confirm if this feature is available. To use the SFTP feature, SFTP client software is required such as FileZilla or Cyberduck. Use the following settings.

Address: sftp://gt.pillar-biosciences.net Port: 2225 Username: Your PiVAT username Password: Your PiVAT password

Notes:

* When uploading is interrupted, the file number and total size may be incorrect, please logout & login to the SFTP client again so the system can refresh. Deletion of corrupted files will also require you to logout and login again.

* Some SFTP clients, FileZilla for example, may run into an infinite loop of retrying an upload when a file exceeds the size limitation. Please terminate it manually.

Software List

Package	Version	License	Link
BWA	0.7.16a-r1181	Apache 2.0	https://github.com/lh3/bwa
VEP	106.1	Apache 2.0	https://github.com/Ensembl/ensembl-vep
samtools	1.10	The MIT/Expat License	https://github.com/samtools/samtools
MSIsensor	0.6	MIT License	https://github.com/ding-lab/msisensor
SSW	1.2.5	MIT License	https://github.com/mengyao/Complete-Striped-Smith-Waterman-Libr



Setup an Analysis

Creating Analysis Parameters

- 1. Begin with selecting **Analysis** on the Navigation Bar to display the following Analysis Guide page.
- 2. User may skip parameter setting and select **Go to Start Analysis.**



The Custom parameter setting described in this Creating Analysis Parameters section is not required to start a run.

PILLAR Dashboard	Analysis Tools					🖸 usermanual 👻	🗞 English 🕜 Help
Parameters	帝 / Analysis						
🖈 Start Analysis	Analysis Guide						
盟 Analysis Results	1 Create Custom Paramet	ters It's an optional step				Cre	ate New Parameters
	ID	Parameters File Name	F	Panel		Date	Operation
			No parameters file found				
	2 Start An Analysis						
	Go To Start Analysis						
	3 View Analysis Results						More Tasks >
	ID	Task	Panel	Date	SA Status	Stage 🕕	Operation
(No analysis task found				

Figure 8 Analysis Guide page

- 3. Select **Create New Parameters** to display the following Analysis Parameters page. This enables user to customize analysis parameter.
 - **Parameters Name:** Enter desired parameters file name. The user defined parameters will be saved with this name for future use.
 - Panel Selection: Select desired panel using drop-down menu.



4. Select ① Create New Parameters and navigate across the tabs to customize settings for CORE_MODULE, Local Realignment, Paired End Assembly, Variant Call Reduce, Annotation/Filtering and CNV Processing.

PLLAR	Dashboard	Analysis	Tools					\rm O use	ermanual 🔻	菟 English	? Help	
Parameters		🅱 / Analysis / P	arameters							Save Para	ameters	
🛠 Start Analysis		Analysis Par	ameters									
🕮 Analysis Results		Parameters	Name 🕕	usermanu	al							
		Panel Se	lection 🕚	ONCO/Re	veal Lung and Col	on Cancer Panel (L	2103)			×	~	
		1 Secondary A	1alysis Paran	neters		2 Post Filter Parameters				3 QC Parameters		
		CORE_MODU	LE BWA	Alignment	BAM Import	Indel Processir	g Local Realignme	nt Combine Indels	Paired End	Assembly		
		Bam Export	Read To P	osition	Precall Filtering	Variant Call	Variant Call Reduce	Annotation/Filtering	Run Sum	mary		
				Param	eter			Value				
		PARALLELISM_	ТҮРЕ				sample				~	
		NUMBER_CPU									20	
		WRITE_DEBUG	_EXCEL_FILES				False				~	
	K	VCF_FORMAT_	FIELD_VER				True				~	

Figure 9 Analysis Parameters: ① Secondary Analysis Parameters page

5. Select **(2)** Post Filter Parameters to customize settings for Filtering, Grouping and Column Layout.

PLLAR	Dashboard	Analysis Too	ls Admin		💽 admin 🔹 落 English ? Help						
Parameters		🕱 / Analysis / Parame	eters		Save Parameters						
🗱 Start Analysis		Analysis Parame	eters								
III. Analusia Desulta											
m Analysis Results	5	Parameters Name	e 🛈 usermanual								
		Panel Selection	ONCO/Reveal Lung and C	slon Cancer Panel (LC103)	x ~						
		1 Secondary Analys	is Parameters	2 Post Filter Parameters	3 QC Parameters						
		Use Default Post Filte	Use Default Post Filter Parameters								
		✓ Filtering			+ Add New Filter						
		How to Add Filters Note: If you do not h	nave an SA Call Match True Posi	ive condition for any Filter, False Positives will pass for that Filter.							
		1. Click:		+ Add Condition							
		2. Choose a variar	nt report column, ex:	SA Call							
		3. Choose an oper	rator, ex:	Match 0							
		4. Enter a value, e	x:								
		Repeat the proc conditions.	cess, until you have fully specified	the Filter							
		(Note: all condition	ons in a Filter must pass for a variant to	pass that Filter)							
		6. TO add a new P	inter, click:	+ Add New Filter							
		Rules 1. All Conditions 2. Different Filter 3. If any Filter doe	within a Filter are AND'd, which rs are OR'd, meaning that they are as not have an SA Call Match Tru	neans that they must all pass or no results will pass for that Filter. independent, and additive. Positive condition, the default PiVAT QC will not apply, only the conditions defined for that Filter.							
		4. Text values are	case-sensitive.								
		Filter Name: 1	Default-filter	You can add some description here.	+ Add Condition 🛅 Remove						
		SA Call	▼ Match	True Positive True Positive							

Figure 100 Analysis Parameters: 2 Post Filter Parameters page



6. In ② **Post Filter Parameters,** use the CNV Threshold Grid to customize thresholds for filtering CNV calls. The CNV Threshold Grid will only appear if the selected panel is a CNV panel.

PILLAR	Dashboard	Analysis Tools				0	usermanual 🔹 🕉 English 🕜 Help
Parameters		倄 / Analysis / Parameters	5				Save Parameters
🖈 Start Analysis		Analysis Paramete	ers				
🕮 Analysis Results		Parameters Name (
		Panel Selection		XV			
		1 Secondary Analysis Pa	rameters	2	Post Filter Parameters		G QC Parameters
		✓ CNV Threshold Grid					
		Gene Symbol	Copy Gain Threshold 🚺	Enable Copy Gain	Copy Loss Threshold	Enable Copy Loss	Amplicon Count Threshold
		CCNE1	1.4		0.8		5
		EGFR	1.4		0.8		5
		ERBB2	1.4		0.8		5
		FGFR1	1.4		0.8		5
		FGFR2	1.4		0.8		5
		FGFR3	1.4		0.8		5
		FLT3	1.4		0.8		5
		KDR	1.4		0.8		5
		KIT	1.4		0.8		5
	۲	KRAS	1.4		0.8		5

Figure 111 Analysis Parameters: 2 Post Filter Parameters page CNV Threshold Grid

Note: When enabling Copy Loss for panels using the cfDNA CNV Caller (see Appendix C, currently only P-LBX-01): the **CNV_MAX_COPY_NUMBER** parameter will need to be adjusted to ensure that calls below the Copy Loss Threshold are reported. See Appendix C.

7. Select **3 QC Parameters** to customize QC parameters for **Sample** and controls (**NTC**, **PosCtrl**, **NegCtrl**).

PILLAR Dashboard	Analysis Tools				🔂 usermanual 👻 🖏 English 🕥 Help
Parameters	會 / Analysis / Parameters				Save Parameters
🛠 Start Analysis	Analysis Parameters				
Analysis Results	Parameters Name 🜒 user				
	Panel Selection () ONC	O/Reveal Lung and Colon Cancer Pan	x ~		
	3 QC Parameters				
	QC Parameters only control SN Use Default QC Parameters				
	✓ NTC				
					+ Add Condition
	Absolute Amplicon Coverage Ma	ax 🕶 Less(<) 💌		100 📋	
	Relative Amplicon Coverage Max	x 💌 Less(<) 💌		10 💼	
	No Mutations Detected	▼ Is True ▼	Ē		
	✓ PosCtrl				
					+ Add Condition
	Q30 Reads 💌 🛛	Greater or equal(>=) ▼	25	亩	
	Effective On-Target Rate 💌 🛛	Sreater or equal(>=) 👻	0	亩	
	Amplicon Coverage Min 👻 🛛	Sreater or equal(>=) 👻	0	亩	
	No Mutations Detected 💌	s False 👻 🛅			

Figure 12 Analysis Parameters: ③ QC Parameters page



QC Parameters only control SNV/Indel QC calling. Other Variant QC calls are driven by Secondary Analysis Parameters.



Start Analysis

Please refer to the following appendices for caller-specific instructions.

- Appendix A: SNV/Indels Caller
- Appendix B: Microsatellite Instability
- Appendix C: Somatic CNV Caller
- Appendix D: Thalassemia CNV Caller
- Appendix E: SMA Thalassemia Caller

Appendix F: Fusion CallerAppendix E: SMA Caller

The SMA analysis is based on the double normalization method. The normalization baseline is calculated from negative reference samples. The SMA Caller calculates the copy number ratios of Exon-07 and Exon-08 amplicons on the SMN1 and SMN2 genes.

SMA Sample Setup

For each SMA analysis run, the user should provide 3-5 (minimum 2) in-run normal (negative) reference samples with similar sample condition and preparation process as the positive samples. If less than 2 negative reference samples are provided, the run will fail.



See <u>CNV Sample Setup</u> section in <u>Appendix C: Somatic CNV Caller</u> for instructions to define normal samples to be used as negative references.

SMA Analysis Setup

- 1. See <u>CNV Analysis Setup</u> section in <u>Appendix C: Somatic CNV Caller</u> for instructions to setup a SMA Thalassemia analysis.
- 2. The SMA_RESULTS excel file is output upon completion of the analysis.
- 3. Detailed SMA call information is in the "SMA Call Report" sheet of the SMA_RESULTS excel file.
- 4. Each sample's detailed SMA Report can be downloaded as a PDF file in the "Report" section of the results page.
- 5. The fully normalized copy number ratios can be found in "Fully Normalized" sheet in SMA_RESULTS excel file. Note that the copy number ratio in PiVAT is defined as the copy number ratio of a potentially positive sample to that of the negative reference samples (diploid with 2 copy number).

SMA QC parameters

See <u>CNV QC parameters</u> section in <u>Appendix C: Somatic CNV Caller</u> for a list of user adjustable QC parameters.



PiVAT Output: SMA_RESULTS file

Definition and/or description of result columns reported in SMA_RESULTS file sheets are provided below.

Sheet Tab	Column Name	Definition/Description			
	Sample ID	Unique Sample ID for each sample.			
	Gene-Exon	Name of the gene/exon pair, or if a control target, name of the control amplicon.			
SIVIA Call Report	Location	Genomic coordinates of exon (hg19).			
	Copy Number	Copy number of the gene-exon or of the control amplicon			
	Copy Number Ratio	The copy number ratio of the gene-exon or control-amplicon to the copy number (2) of a negative normal sample.			
	QC Criteria	Indicate which run QC criteria that each row is reporting, including "Negative_Reference", "Sample" and "Run_Status"			
SMA Run QC	QC Status	Whether the QC is passed for each run QC criteria. If run QC is passed, it is labeled as "Pass". It run QC is failed, it is labeled as "Fail".			
	Sample ID	Unique Sample ID for each sample			
Filtered SMA	Filter Reason	The reason that the sample is filtered out from the SMA Thalassemia analysis.			
Samples	Sample Type	Indicate whether the reported sample is "Sample" or "Negative Reference". This is pre-defined by the user before starting the SA analysis.			
CNV Segment Coverages	This sheet reports the initially added to the the the row indices are stated to the the row indices are stated to the row indi	ne raw segment coverages of each amplicon for all the samples e SMA analysis. The column headers are Amplicon Names and Sample IDs.			
Normalized CoveragesThis sheet reports the normalized copy number ratio of each amplicon for samples that pass the sample QC. Note that the filtered samples are not re this sheet. The column headers are Amplicon Names and the row indices a Sample IDs.					



SMA Calls Table

SMA Calls table contains the copy number ratios for the sample(s) selected in Select Sample.

	Analysis Tools	s Admin				0 admin 🝷 🕉	🛦 English (? Help
Parameters	裔 / Analysis / Analysis	s Results / Analysis T	īask				
🛠 Start Analysis	Analysis Task SM	A_CNV_Test-20230602	2-155203				
Analysis Results	📀 Download Result Zip	Files 🗋 Logfile	Task Information				C Rerun
	QC Summary QC	Stat Variant	Report				
	SNV / Indel	CNV SMA					
	Select Sample:	RDvGMpTHAL172d20120	D2iN1-PB09 ×				x ~
	Select Reference:	RDvGMpTHAL172d20120	02iN1-PB05 ×				× ~
	SMA Calls						
	Sample ID)	Gene-Exon	Location	Copy Number	Copy Nur	mber Ratio
	RDvGMpTHAL172d201	1202iN1-PB SMN1	_Ex07	chr5:70241792-70248150	3.92	1.96	
	RDvGMpTHAL172d201	1202iN1-PB SMN1	_Ex08	chr5:70248390-70248554	3.92	1.96	
	RDvGMpTHAL172d201	1202iN1-PB SMN2	_Ex07	chr5:69366366-69372729	0.00	0.00	
	RDvGMpTHAL172d201	1202iN1-PB SMN2	_Ex08	chr5:69372969-69373133	0.00	0.00	
	RDvGMpTHAL172d201	1202iN1-PB CtrL01	.ZNF648.1Q25.3	chr1	2.04	1.02	
	RDvGMpTHAL172d201	1202iN1-PB CtrL02	.AOX1.2Q33.1	chr2	2.01	1.01	
	RDvGMpTHAL172d201	1202iN1-PB CtrL03	.PLCL2.3P24.3	chr3	2.00	1.00	
	RDvGMpTHAL172d201	1202iN1-PB CtrL04	.PDLIM5.4Q22.3	chr4	2.18	1.09	
	RDvGMpTHAL172d201	1202iN1-PB CtrL05	.TBC1D19.4P15.2	chr4	1.97	0.98	
<	RDvGMpTHAL172d201	1202iN1-PB CtrL06	.PJA2.5Q21.3	chr5	2.31	1.16	

Figure 43 Analysis Results: SMA Calls table

Definition and/or description of result columns reported in the SMA Calls table are provided below.

Column Name	Definition/Description
Sample ID	Unique Sample ID for each sample.
Gene-Exon	Name of the gene/exon pair, or if a control target, name of the control amplicor
Location	Genomic coordinates of exon (hg19).
Copy Number	Copy number of the gene-exon or of the control amplicon
Copy Number Ratio	The ratio of the copy number of called Thalassemia CNV to the copy number (2) negative normal sample.



SMA Plot

A SMA plot for each sample will be located on the Variant > SMA tab below the SMA Calls table and SMA RUN QC table in Analysis Results. Each boxplot represents the copy number ratios of all CNV Normal Samples specified in the run, and the orange data points represent the copy number ratios of the potentially positive sample. The x-axis represents gene-exon labels, and the y-axis represents copy number ratio.



Figure 44 Analysis Results: SMA Sample Box Plot



To save the box plot graph, hover over the graph, and click the "**Download plot as a png**" icon as indicated below.



Figure 45 Analysis Results: SMA Sample Box Plot – Saving as PNG

- Appendix F: Fusion Caller
- 1. Select **Start Analysis** on the left menu bar to display the following Start Analysis: 1. Start From page.
 - Select Panel: Select desired panel from the drop-down menu.
 - **Parameters File**: Parameters file selection is optional. Refer to <u>Creating Analysis Parameters</u> section for instructions to create custom parameters file. Default parameters will be applied if parameters file is NOT selected.
 - Select Samples: Select desired FASTQ OR BAM format files. Refer to <u>Data Management</u> section for instructions to import data.

BAI (BAM Index) files may only be used as PiVAT input if accompanied by their respective BAM files. Using only BAI files without their respective BAM files will result in a run failure.



PILLAR Dashboard	Analysis Tools	🕄 usermanual	🝷 🕱 English 💡 Help
☆ Parameters	备 / Analysis / Start Analysis / Start From		
🛠 Start Analysis	Start Analysis		
III Analysis Results	1. Start From 2. Edit Definition	3. Preview 8	ջ Launch
	Select Panel and Parameters file:		
	Select Panel: Select		
	Parameters File: Select		
	Select Samples:		
	① Choose fastq or bam files, or a samplesheet(.csv) and its related sample files to start an analysis.		
	Home /	Filte	er
	Definition Name	Size	Modified Date
	🔄 🚰 input		2023-02-23 11:41:51
	🔄 🐚 user		2023-06-01 14:11:30
	Reset Next		
	You	ur Selected Items: 0	

Figure 13 Start Analysis: 1. Start From page



- 2. Select **Next** to display the following window: 2. Edit Definition page.
 - Analysis Name: Enter desired name for the analysis to uniquely identify the run.
 - **Parameters File**: Parameters file selection is optional. Refer to <u>Creating Analysis Parameters</u> section for instructions to create custom parameters file. Default parameters will be applied if parameters file is NOT selected.
- 3. Users may edit sample information such as **Sample Name**, **QC Type** as well as edit or skip **Sample Files**.
- 4. Users may also select samples using check boxes to define Tumor-Normal Paired and CNV Normal Samples.



Tumor-Normal Paired definition is required for performing MSI analysis, refer to Appendix B: Microsatellite Instability (MSI) Caller for details.



CNV Normal Samples definition is recommended for CNV analysis, refer to Appendix C: Somatic CNV Caller for details.



P-LBX-01 requires at least 1 CNV Normal to be specified (using Define CNV Normal Samples button as seen in the image below). If sample tumor content is known, it should be specified in the **Tumor Content** % column (as a percentage) for improved CNV calling performance.

	Dashboard	Analysis Tools	🖸 usermanual 🝷 薞 English 🕜 Help
Parameters		🐐 / Analysis / Start Analysis / Edit Definition	
😤 Start Analysis		Start Analysis	
Analysis Results		1. Start From 2. Edit Definition	3. Preview & Launch
		Analysis Name: Please input analysis definition name -20230601-143746 Parameters File: Select	~
		Tumor-Normal Paired 😰 Define CNV Normal Samples 🛃 Edit Files 😾 Skip Samples 🕄	Insert Sample
		Sample Name 🤳 🛛 🔍 QC Type 💦 Sample Files	Tumor Content % Operation
		HotSpotControl-1_S1(2) 🕰	Edit Files 🌌 Skip Sample 🔊
		Back Next	
		PiVAT RUO version 2023.1.0 released on 2023-04-28	© Pillar Biosciences Inc 2015-2023

Figure 14 Start Analysis: 2. Edit Definition page



- 5. Select Next to display the following window: 3. Preview & Launch page. Once user confirms the analysis setup, select Launch Analysis.
- 6. A pop-up message will appear. User may select:
 - Start Another Analysis to redirect to Start Analysis page, or
 - Go to Analysis Results to redirect to Analysis Results page.

PILLAR Dashboard	Analysis Tools			🕒 usermanual 👻 🕱 English 🛛 Ə Help
Parameters	倄 / Analysis / Start Analysis /			
🗱 Start Analysis	Start Analysis			
🖽 Analysis Results			2. Edit Definition	9 3. Preview & Launch
	Email Address: ① Please input E			
	Summary			
	Analysis Name: usermanual -2 Panel: ONCO/Reveal Lung and Parameters file: File extensions: .fastq.gz	Start Another Analysis	d successfully. sis Go To Analysis Results	
	Sample Name			Tumor Content %
	HotSpotControl-1	Sample	HotSpotControl-1_S11_L001_R1_0 HotSpotControl-1_S11_L001_R2_0	01.fastq.gz 01.fastq.gz
			Back Launch Analysis	
	PiVAT RUO version 2023.1.0 released or	1 2023-04-28		© Pillar Biosciences Inc 2015-2023

Figure 15 Start Analysis: 3. Preview & Launch page



Monitoring the Task Status

- 1. The PiVAT Dashboard will display summary of analysis tasks, data storage and **Recent Task Activity** of the user currently logged in.
- 2. To view the complete list of order task, select More Tasks >.

PILLAR	Dashboard	Analysis	Tools					🕄 usermanual 👻	🕱 English 🛛	Help
System Tasks In Queue		0	Your Tasks Running / In Queue / Total	0/0/1	Average Task Analysis Time	0.83 н	Data Sto disk usage	rage / quota	0.0 G /100).0G
Task Numbe	Monthly Trend	l (the last 6 mont	ths)		Task Number Per	Panel				
CO/Reveal Lung and Colon Cancer Panel (LC103):1 2023-2 2023-3 2023-4 2023-5 2023-6										
Recent Task Ac	tivity								More Ta	asks
ID		Task		Panel		Date	SA Status	Stage 🚺	Operation	1
8 userman	ual-20230601-143	3950_8_2023.1.0_UT	C20230601184004 (ONCO/Reveal Lung	; and Colon Cancer	Panel (LC103)	2023-06-01 14:40:04	completed	RP : completed	Abort Re	run
PIVAT RUO version 2	023.1.0 released on 2	2023-04-28						© Pillar	Biosciences Inc 201	5-2023

Figure 16 Dashboard page



- 3. The following information is provided for each task row:
 - **ID:** Unique task ID number for current database.
 - **Task:** Analysis name entered by user during analysis setup.
 - Panel: Panel selected for the analysis.
 - **Date:** Date when the analysis task was created.
 - **Status**: Analysis task status
 - ✓ **queued**: There are other task(s) running. This task is queued until preceding tasks are finished.
 - ✓ **running**: This analysis is currently running.
 - ✓ **completed**: The analysis has completed successfully.
 - ✓ **failed**: An error occurred, and the analysis did not complete.
 - ✓ **aborted**: This task was aborted by a user.
 - Stage: Where in the Analysis the task currently is. Only displayed if in the **running** or **failed** state.
 - ✓ SA: Secondary Analysis Pipeline
 - ✓ FA: Functional Annotation
 - ✓ PF: Post Filter
 - ✓ QC: Quality Control
 - ✓ **RP:** Report Generation
 - **Operation**: User may select **Abort** to terminate an analysis task in running state or **Rerun** to repeat an analysis as-is or with edits to analysis setup.

PILLAR Dashboard	I Ani	alysis Tools Ad	min				🕒 admin 👻	🕱 English 🕜 Help
Parameters	谷 / A	nalysis / Analysis Results						
🛠 Start Analysis	Anal	ysis Results View Task	History and Result Status					
III Analysis Results		ID: Please input id numbe	r Task Name:	Please input task name				
	Crea	ated: Please input start time	~ Please input	end time				
	P	anel: All	 Status: 	All	•			
		Search Clear						
	ID	Task		Panel	Date	SA Status	Stage 🌗	Operation
	10	MY766_rebuild_Bulk_Pool_	Testing-201 ONCO/Reveal	Myeloid Panel (MY766)	2023-06-01 16:00:07	completed	RP : completed	Abort Rerun
	9	RD_Pool_Testing_MY766_Re	build-2023(ONCO/Reveal	Myeloid Panel (MY766)	2023-06-01 15:59:09	completed	RP : completed	Abort Rerun
	7	QC_Myeloid_231_like_test	-20230531-€ ONCO/Reveal	Myeloid Panel (MY766)	2023-06-01 10:20:14	completed	RP : completed	Abort Rerun
	6	QC_Myeloid_231_like_test	-20230531-€ ONCO/Reveal	Myeloid Panel (MY766)	2023-05-31 09:37:47	failed		Abort Rerun
	5	QC_Myeloid_231_like_test	-20230530-1 ONCO/Reveal	Myeloid Panel (MY766)	2023-05-30 14:40:42	aborted		Abort Rerun
	4	QC_Myeloid_231_like_test	-20230530-1 ONCO/Reveal	Myeloid Panel (MY766)	2023-05-30 14:06:03	aborted		Abort Rerun
	3	230418_NDX550762_RU0_008	5_AHH5NVAF> P-LBX-01-DN	A2 (P-LBX-01-DNA2)	2023-05-26 17:08:19	failed		Abort Rerun
	2	230418_NDX550762_RU0_008	5_AHH5NVAF> P-LBX-01-DN	A (P-LBX-01-DNA)	2023-05-26 16:58:02	failed		Abort Rerun
	1	230418_NDX550762_RUO_808	5_AHH5NVAF) P-LBX-01-DN	A (P-LBX-01-DNA)	2023-05-26 16:25:52	failed		Abort Rerun
	0							
	PIVAT RU	IO version 2023.1.0 released on 20	23-04-28				© Pillar	Biosciences Inc 2015-2023

Figure 17 Analysis Results: Task History page



View Summary of Task Details

1. Hover the mouse over the task name to view a pop-up summary of analysis.



Figure 18

- 2. The summary information includes:
 - Task name
 - Panel
 - Parameters file name
 - Task status
 - Stage
 - Total number of Samples
 - Duration of the run
 - Created and Completed date/time stamps
 - Note that the time between the Created and Completed time stamps may be longer than the Duration if the task was queued before starting
 - The user who ran the task
 - Email address if configured
 - The output folder location
 - Input files
 - Arguments



View Completed Analysis Task

1. Select task hyperlink from list of completed tasks to display the Analysis Task page. Users may download analysis results, view log file and task information or rerun an analysis task.

	Dashboard	Analysis	Tools	Admin				🖸 admin 🔻	え English	? Help
Parameters		🕱 / Analysis / A	nalysis Results	/ Analysis Task						
🛠 Start Analysis		Analysis Tasl	C_Myeloid	d_231_like_test-20230531-09	93612					
Analysis Results		Download Res	ult Zip Files	🗅 Logfile 📄 Task In	formation				C	Rerun
		QC Summary	QC Stat	Variant Report						
Sam			e Name	Sample Type	Status	Note				
QCvSMpMyeloidd230525iN3-2			d230525iN3-2	3PB0122-TV12-40ng-1-B1	Sample	Success				
		QCvSMpMyeloidd	d230525iN3-2	3PB0122-TV2-40ng-2-B1	Sample	Success				
		QCvSMpMyeloid	1230525iN3-2	3PB0122-TV12-40ng-2-B2	Sample	Success				
		QCvSMpMyeloid	d230525iN3-2	3PB0122-NTC-2-E1	NTC	Success				
		QCvSMpMyeloid	1230525iN3-2	3PB0122-NTC-1-B2	NTC	Success				
		QCvSMpMyeloid	d230525iN3-2	3PB0122-NTC-1-E2	NTC	Success				
		QCvSMpMyeloidd	d230525iN3-2	3PB0122-TV12-40ng-2-E1	Sample	Success				
		QCvSMpMyeloidd	d230525iN3-2	3PB0122-NTC-2-B2	NTC	Success				
		QCvSMpMyeloid	d230525iN3-2	3PB0122-TV12-25ng-1-E2	Sample	Success				
		QCvSMpMyeloidd	d230525iN3-2	3PB0122-TV2-25ng-2-E2	Sample	Success				
		QCvSMpMyeloidd	230525iN3-2	3PB0122-TV2-25ng-1-E2	Sample	Success				
	۲	QCvSMpMyeloidd	d230525iN3-2	3PB0122-TV2-40ng-2-E1	Sample	Success				

Figure 19 Analysis Task page

2. Click the **Download Result Files** button to download a zip formatted file with the files produced by the analysis run.

File download is unavailable if the task is in the **queued** or **deleted** status.



Users may delete the zip file from this screen. Please verify that analysis data has been downloaded and stored in accordance with your organization's IT policies <u>before</u> deleting data from the PiVAT system.



	nboard	Analysis	Tools	Admin			\rm 0 admin 👻	菟 English	? Help
Parameters	ŝ	/ Analysis	s / Analysis Res	sults / Analysis Task /	Download Result Files				
😤 Start Analysis	D	Downloa	ad Result Z	Cip Files QC_Myeloi	d_231_like_test-20230531-0	93612			
Analysis Results		🛅 Delete Z	lipfile 🔨 Re	generate Zipfile				🛅 Delete	Output
	C	Download	Filename				Size		Date
		0	common.zip				261.58 MB	2023-06-01	22:59:41
	(0	CUSTOMER_RESULT testing2.xlsx	TS_QC_Myeloid_231_like	e_test-20230531-093612_Pi	VATrelease-2023.1-	20.2 MB	2023-06-01	23:12:36
		0	log.zip				6.65 MB	2023-06-01	23:08:33
		0	QCvSMpMyeloidd	230525iN3-23PB0122-NTC	C-1-B1.zip		2.89 MB	2023-06-01	23:06:23
		0	QCvSMpMyeloidd2	230525iN3-23PB0122-NTC	-1-B2.zip		2.57 MB	2023-06-01	22:39:46
		0	QCvSMpMyeloidd	2305251N3-23PB0122-NTC	C-1-E1.zip		2.95 MB	2023-06-01	23:06:23
		0	QCvSMpMyeloidd	230525iN3-23PB0122-NTC	-1-E2.zip		2.58 MB	2023-06-01	22:39:47
		0	QCvSMpMyeloidd	230525iN3-23PB0122-NTC	C-2-B1.zip		2.73 MB	2023-06-01	23:03:10
		0	QCvSMpMyeloidd	230525iN3-23PB0122-NTC	C-2-B2.zip		2.67 MB	2023-06-01	22:40:46
		0	QCvSMpMyeloidd2	2305251N3-23PB0122-NTC	C-2-E1.zip		2.57 MB	2023-06-01	22:39:46
		0	QCvSMpMyeloidd	230525iN3-23PB0122-NTC	C-2-E2.zip		2.73 MB	2023-06-01	23:06:24
	٠	M Back							

Figure 20 Analysis Task: Download Result Files page



- 3. A deleted zipfile may be recovered by selecting the **Regenerate Zipfile** button. This selection is only available if the files are still present on the system. If the output zipfiles have been deleted, the regular output files may still be available via your network drive or through SFTP. Refer to the Help page to learn more about SFTP.
- 4. Click the **Back** button or Analysis / Analysis Results / **Analysis Task** to return to Analysis Task page.
- 5. Click the **Logfile** button from Task Analysis page to display information for every stage of this task. The information is not available when the task is in the running, queued or deleted status.
- 6. Click the **Back** button or Analysis / Analysis Results / **Analysis Task** to return to Analysis Task page.

PILLAR Dashboard	Analysis Tools	😧 usermanual 👻	薞 English	? Help
Parameters	🐐 / Analysis / Analysis Results / Analysis Task / Logfile			
🛠 Start Analysis	Logfile Activity Log of the Task			
H Analysis Results	2023-06-02 00:22:51,247-1NFO-Faunal with Pipeline version: release-2023.1-testing2, released on 2023-05-30 2023-06-02 00:22:51,252-NFO-Incoding control file for panel L103 with revision Number 2, hash: Ebeac286397e545cfe0 2023-06-02 00:22:55,352-NFO-Faunal ysis started with 1 samples. 2023-06-02 00:22:55,354-1NFO-Faunal ysis started with 1 samples. 2023-06-02 00:22:55,354-1NFO-Sample in serial 2023-06-02 00:22:55,354-1NFO-Sample intervision Number 2, hash: Ebeac286397e545cfe0 2023-06-02 00:22:16,375-1NFO-Sample intervision Number 2, hash: Ebeac286397e545cfe0 2023-06-02 00:22:16,376,371NFO-Sample MoltspotControl-1 2023-06-02 00:22:16,414-1NFO-Sample MoltspotControl-1 2023-06-02 00:22:16,	Raeld7cf655bb .ng, PEA, BanWrite, Rea	dToPosition,	

Figure 21 Analysis Task: Logfile page



- 7. Click the **Task Information** button to display the summary of analysis, this is the same information you get by hovering mouse over the task name on Analysis Result page.
- 8. Click the **Back** button or Analysis / Analysis Results / **Analysis Task** to return to Analysis Task page.

PLLAR	Dashboard	Analysis	Tools	🖸 usermanual 🝷	🕱 English	? Help
 ✿ Parameters ★ Start Analysis 		☆ / Analysis / Task Inform	Analysis Results / Analysis Task / Task Information hation usermanual-20230601-143950			
Analysis Results		Task: usermanu Panel: ONCO/R Parameters File: SA Status: con Stage: RP : cor	ial-20230601-143950 teveal Lung and Colon Cancer Panel (LC103) pipted npieted			
		Total Samples: Duration: 0:49: Created: 2023- Completed: 20 User: usermanu Email:	1 33 06-01 14:40:04 23-06-01 23:12:41 Jal			
		Notes: Output: /user/output/RUC Inputs: HotSpotControl-1 HotSpotControl-1 Arguments:	//LC103/usermanual-20230601-143950_8_2023.1.0_UTC20230601184004 _S11_L001_R1_001.fastq.gz _S11_L001_R2_001.fastq.gz			
	۲	PRIMARY_CONTR	OL_FILE: /control_files/LC103_primary_control.bxt			

Figure 22 Analysis Task: Task Information page

9. Click the **Rerun** button for the option to repeat the analysis as-is or with edits to analysis setup.

 Parameters Start Analysis 	Analysis / Analysis / Analysis / Analysis Task			
🕮 Analysis Results	🔷 Download Result	Zip Files]		C [†] Rerun
	QC Summary Sample Name	DC Stat Variant Report		
	HotSpotControl-1	Are you sure you want to rerun this task? Cancel Edit Definition	Rerun Immediately	
(

Figure 23 Analysis Task: Rerun confirmation message



View Analysis Results

- 1. Users may navigate across the tabs to view QC Summary, QC Stat, Variant and Report.
- 2. Click the **QC Summary** tab to display summary of Sample Type and Status for each sample.

PILLAR Dashboard	Analysis	Tools Admin				😧 admin 👻	🕆 English 🕜 Help	
 ✿ Parameters ☆ Start Analysis 	Analysis / Analysis / A Analysis Tas	nalysis Results / Analysis T K MY766_rebuild_Bulk_Poo	ask I_Testing-202306	01-155935	5			
III Analysis Results	Download Res	ult Zip Files	🖻 Task Info	mation			C Rerun	
	QC Summary	QC Stat Variant	Report					
	S	ample Name	Sample Type	Status	Note			
	RDvSMpMYDR2c	1230519iM1-Spike-1	Sample	Success				
	RDvSMpMYDR2c	l230519iM1-Original-1	Sample	Success				
	RDvSMpMYDR2c	l230519iM1-Original-2	Sample	Success				
	RDvSMpMYDR2c	I230519iM1-Spike-NTC	Sample	Success				
	RDvSMpMYDR2c	1230519iM1-Spike-2	Sample	Success				
	•							

Figure 24 Analysis Task: QC Summary tab

3. Click the **QC Stat** tab to display and navigate sub-tabs to view Summary, Coverage Depth, Coverage Uniformity, and Segment Stats.

PLLAR	Dashboard	Analysis	Tools	Admin			\rm 9 admir	n 🝷 🕱 English 🕜 Help
Parameters		🏫 / Analysis / A	nalysis R	esults / Analysis Task				
🛠 Start Analysis		Analysis Tas	k мү76	i6_rebuild_Bulk_Pool_Tes	ting-20230601-155935			
Analysis Results		Download Res	ult Zip Fi	iles 🚺 Logfile	Task Information			C Rerun
		QC Summary	QC S	tat Variant Rep	port			
		Summary	Cov	erage Depth C	overage Uniformity	Segment Stats		
		BBAM Su	mm	ary				Copy to Clipboard
		Stat		Sample:RDvSMpMYDR 2d230519iM1-Original- 1	Sample:RDvSMpMYDR 2d230519iM1-Original- 2	Sample:RDvSMpMYDR 2d230519iM1-Spike-1	Sample: RDvSMpMYDR 2d230519iM1-Spike-2	Sample:RDvSMpMYDR 2d230519iM1-Spike- NTC
		Overall:Q=30		95.76	95.7	95.43	95.99	87.31
		Overall:Q=20		96.7	96.7	96.5	96.97	89.73
		Properly Paired R	eads	2077824	1598008	1853732	1813560	34876
		Properly Paired R	ead (%)	97.86	97.52	97.33	98.28	83.32
		Mapping Rate (%)	98.43	98.17	98.16	98.95	88.49
		On Target Rate (%	6)	98.34	98.19	98.02	98.13	93.29
		On Target Reads		2055319	1579586	1832435	1791843	34557
		Q=30: 20% of Me	ean	98.15	97.19	97.05	97.04	26.12
		Q=20: 20% of Me	ean	98.14	97.14	97.02	96.93	29.65

Figure 25 Analysis Task: QC Stat tab



4. Click the **Variant** tab to display and navigate sub-tabs to view variant results. Variant results grouping will be displayed according to the Parameters File applied.

PILLAR Dash	oard <u>Analysis</u> Tools Admin		🖸 admin 🝷 薞 English 🕜 Help
Parameters	😤 / Analysis / Analysis Results / Analysis Task		
🛠 Start Analysis	Analysis Task MY766_rebuild_Bulk_Pool_Testing-20230601-	155935	
III Analysis Results	💿 Download Result Zip Files	tion	C Rerun
	QC Summary QC Stat Variant Report		
	SNV / Indel		
	Select Variant Filter: System Default 👻 🗹 🖃		
	Search samples V Contains text Q Search	🔷 Download Excel	
	All variants before filters All true positive variants Pot	ential Clin Significant Other variants	
	IGV Sample Type Sample ID	Gene_Symbol	HGVSC 🔒
	IGV Sample RDvSMpMYDR2d230519iM1-Spike-1	DNMT3A DNMT3A	NM_022552.5:c.2173+26C>T NM_15:
	IGV Sample RDvSMpMYDR2d230519iM1-Original-1	DNMT3A DNMT3A	NM_022552.5:c.2173+26C>T NM_15:
	IGV Sample RDvSMpMYDR2d230519iM1-Original-2	DNMT3A DNMT3A	NM_022552.5:c.2173+26C>T NM_15:
	IGV Sample RDvSMpMYDR2d230519iM1-Spike-2	DNMT3A DNMT3A	NM_022552.5:c.2173+26C>T NM_15:
	(
	« <	> >> Show 100 • Total: 20	1302
	PiVAT RUO version 2023.1.0 released on 2023-04-28		© Pillar Biosciences Inc 2015-2023

Figure 26 Analysis Task: Variant tab

5. Click the **Report** tab to view list of samples and option to download all or selected reports.

PILLAR Dashboard	Analysis	Tools	Admin	🖸 admin 🝷	菟 English	Help
Parameters	🕱 / Analysis / A	nalysis Result	s / Analysis Task			
🛠 Start Analysis	Analysis Tas	k MY766_ret	uild_Bulk_Pool_Testing-20230601-155935			
II Analysis Results	Download Res	ult Zip Files	Logfile 🗎 Task Information		C	Rerun
	QC Summary	QC Stat	Variant Report			
	Variant Filter: Syst	em Default		😨 Dow	nload Total F	Report
	* The final report var	iant classification	is fixed, and does not correspond to the custom groups defined in the parameters.			
			Sample Name	c	Operation	
	RDvSMpMYDR2d23	0519iM1-Spike-	1	Downloa	ad Sample Rep	ort
	RDvSMpMYDR2d23	0519iM1-Origin	al-1	Downlow	ad Sample Rep	ort
	RDvSMpMYDR2d23	0519iM1-Origin	al-2	Downloa	ad Sample Rep	ort
	RDvSMpMYDR2d23	0519iM1-Spike-	NTC	Downloa	ad Sample Rep	ort
	RDvSMpMYDR2d23	0519iM1-Spike-	2	Downloa	ad Sample Rep	ort
٠	PiVAT RUO version 202	3.10 released o	n 2023-04-28	© Pillar F	Signation and Signatures Inc.	2015-2023

Figure 27 Analysis Task: Report tab



User password

1. A user can change their own password by selecting the **Profile** option from the drop-down menu for the username. The following page is displayed:

PILLAR	Dashboard Analysis Tools	🖲 usermanual 🝷	🕱 English	? Help
	♠ / Profile Profile Change Your Current Password			
	Change Password 🕦			
	Current Password			
	New Password			
	Repeat Password			
	Update Password			
PiVAT RUO version 2023.	1.0 released on 2023-04-28	© Pillar	Biosciences Inc	2015-2023

Figure 28 Profile: Change Your Current Password page

- 2. Create new password and then click **Update Password** to confirm.
- 3. To exit without changing your password, click $\widehat{}$ or any option from the Navigation Bar.



Record the new password in accordance with your organization's IT policies.



Data Management

- 1. Begin by clicking **Tools** on the Navigation Bar and **Data Management** on the left menu bar to display the following Data Management page.
- 2. The Data Management page is used to upload input data (FASTQ) to the PiVAT system for analysis.
- 3. Storage data is displayed for verification of sufficient free storage before attempting to upload data.

PLLAR	Dashboard	Analysis	Tools				🕒 u:	sermanual 🔻	🕱 English	? Help
⊯ IGV		😤 / Tools / D)ata Management							
🎟 Data Managemen	nt	Data Mana	agement Rep	oository to Upload and Manage	Input Data for Runnin	g Analyses				
Download		Sample data fil directory. Samp Also you can us	Sample data files can be uploaded to this section for use as inputs to the analysis tasks. The Project name corresponds to a folder with th directory. Sample files used for a task can hence be uploaded to a Project and be used as a bundled input for the Analysis Task. Also you can use SFTP to upload sample data files.(DSTP Guide)							
		Storage Usa	ge:0 GB out of	100 GB					+ Create	Project
		ID		Project Name		Number Of Fastqs	Date		Operation	
					No projects	found				
	<	PiVAT RUO version	2023.1.0 released on 2	2023-04-28				© Pillar B	Biosciences Inc.	2015-2023

Figure 29 Data Management page

- 4. Input data can be uploaded to a Project listed on the page. To create a new project, click + Create Project.
- 5. All created projects are listed on this page. Click the Project Name you wish to upload data to and the following page will be displayed:

PILLAR Dashbo	ard	Analysis	Tools				O use	ermanual 👻 🕺 English 🛛	? Help
⊯ IGV		骨 / Tools / Dat	a Management /	Details					
🎫 Data Management		Project Deta	ails Data for pro	oject: usermanual					
Download		🔷 Upload File					🖍 Ren	ame Project 📋 Delete P	roject
				Filename		Size	Date	Operation	
					No files in project				
		Number of fastqs: 0							
		M Back							
	٢	PiVAT RUO version 20	23.1.0 released on 20	023-04-28				© Pillar Biosciences Inc 20	15-2023

Figure 30 Project Details page



6. Click the **Upload File** button to display the Upload Input Files page. From here users may drag & drop files or select the files for upload. Once your files are displayed on the page, select **Start Upload**.

PILLAR Dashboard	Analysis Tools	🖸 usermanual 👻 🕱 Englisi	n ? Help
⊯ IGV	脅 / Tools / Data Management / Upload		
🖽 Data Management	Upload Input Files Data for project: usermanual		
Download	M Backusermanual		
	~		
	Drag and drop files here, or click to select files to upload.		
	No files selected for upload yet, file name can only contain characters: a-zA-ZO-9		
	Start Upload		
•	PIVAT RUO version 2023.1.0 released on 2023-04-28	© Pillar Biosciences I	nc 2015-2023

Figure 31 Upload Input Files page

7. Users may also select Rename or Delete the project from the Project Details page

Take caution when deleting a project, as it will result in all data within the project to be deleted.





SFTP Download

To download files using SFTP (Secure File Transfer Protocol), your installation of PiVAT must have the SFTP feature enabled by the administrator. Please contact your administrator to confirm if this feature is available. To use the SFTP feature, SFTP client software is required such as FileZilla or Cyberduck. Use the following settings.

Address: sftp://[IP address of your PiVAT system] Port: 2225 Username: Your PiVAT username Password: Your PiVAT password



When uploading is interrupted, the file number and total size may be incorrect, please logout & login to the SFTP client again so the system can refresh. Deletion of corrupted files will also require you to logout and login again.



Some SFTP clients, FileZilla for example, may run into an infinite loop of retrying an upload when a file exceeds the size limitation. Please terminate it manually.

Troubleshooting



If the system requires a restart, the user should check that all running and queued analysis tasks have completed before restarting the system.



If a task fails due to insufficient disk space, an administrator user can free up space on the system. There must be at least 100 GB of free space on the system for a task to start.



To avoid performance issues, do not run other applications on the workstation while PiVAT is running.



If the PiVAT application will not load, contact Pillar Biosciences at (800) 514-9307 or <u>support@pillarbiosci.com</u>.



If PiVAT is nonresponsive, it is possible to restart PiVAT using the following command:

- Open a terminal window
- cd /pillar/docker_files
- sudo docker-compose down
 - o Superuser password required
- sudo docker ps
 - Verify that no docker containers are active
- sudo docker-compose up -d
- sudo docker ps



• Verify that the required containers are now running.



Legal Notices

PiVAT[®], Pillar[®] are trademarks of Pillar Biosciences, Inc. Illumina[™] is a registered trademark of Illumina, Inc., Ubuntu is a registered trademark of Canonical Ltd. Chrome is a registered trademark of Google.

Firefox is a trademark of the Mozilla Foundation in the US and other countries.

Company Information

Pillar Biosciences, Inc. 9 Strathmore Road Natick, MA 01760

(800) 514-9307 support@pillarbiosci.com https://pillarbiosci.com


Appendix A: SNV/Indels Caller

The SNV/Indels calling analysis reports all SNV and indel variant calls that have passed PiVAT's filtering criteria based on parameters set. The output will display true positive SNV and indel variants and annotations.

SNV/Indels Analysis Parameters Setup

Please refer to <u>Creating Analysis Parameters</u> section on page 11 for instructions on how to create and save analysis parameters.

1. Navigate to ① Secondary Analysis Parameters page to access the following list of adjustable parameters and default values in [] by tab.

Tab	QC Parameters [default]				
	VCF_FORMAT_FIELD_VER [True]				
CORE_MODULE	This parameter collapses all reports of a mutation across samples to a single entry in VCF files, using the FORMAT field. Default is set to True.				
	SOFT_CLIP_COUNT (count) [10]				
Local Realignment	The number of same soft clips threshold to keep the data. Default is set to 10.				
	PAIRED_READ_ONLY [False]				
Paired End Assembly	This parameter is to set if only paired end reads will be accepted for calling. Default is set to False.				
	VCR_FILTER_QUALITY (q-value)				
Variant Call Reduce	Minimum variant quality acceptance. Default value may change depending on the chosen panel for the analysis.				
	VCR_FEATURE_FILTER [True]				
	Filters variants by the Feature column with NM identifiers. Default is set to True.				
	VCR_FILTER_QUALITY (q-value)				
	Minimum variant quality acceptance. Default value may change depending on the chosen panel for the analysis.				
	VCR_FILTER_FREQUENCY (pct)				
Annotation/Filtering	Minimum variant frequency. Default value may change depending on the chosen panel for the analysis.				
	VCR_FILTER_REPEAT_FREQUENCY (pct)				
	Minimum variant frequency for repeat regions. Default value may change depending on the chosen panel for the analysis.				
	VCR_FILTER_COVERAGE (count) [50]				
	Minimum total coverage. Default is set to 50.				



Tab	QC Parameters [default]
	VCR_FILTER_MIN_RATIO [-2]
	Minimum variant read direction ratio2 implies variant is on the reverse read while reference on the forward read. 0 means no preference for forward or reverse read of variant with respective to the reference. Default is set to -2.
	VCR_FILTER_MAX_RATIO [2]
	Maximum variant read direction ratio. Setting to 2 implies variant is on the reverse read while reference on the forward read. Setting to 0 means no preference for forward or reverse read of variant with respective to the reference. Default is set to 2.
	VCR_FILTER_VARIANT_COVERAGE (count)
	Minimum variant coverage. (Total Coverage * Variant Read Frequency)
	VCR_REPEAT_REGION_QUALITY_FILTER (q-value) [30]
	Minimum variant quality within a repeat region. Default is set to 30.
	SM_OVERLAP_THRESHOLD [5]
	The threshold parameter is used for merging variants from overlapping amplicons. If a single variant has a frequency above this value and all others are below, that variant is taken alone, otherwise all variants are merged together. Default value is set to 5.
	FLANKING_DISTANCE [5000]
	The distance of upstream and downstream between a variant and a transcript for which VEP will assign the upstream_gene_variant or downstream_gene_variant consequences. Default is set to 5000.
	CUSTOM_ANNOTATION_FLAG [False]
	This parameter is to set whether custom VEP annotations should be run. Default is False.



PiVAT Output: CUSTOMER_RESULTS file

Definition and/or description of result columns reported in CUSTOMER-RESULTS file sheet tabs are provided below.

Sheet Tab	Column Name	Definition/Description					
	Sample_ID	User defined library ID, derived from FASTQ/BAM file name					
	SampleType	Type of sample used for analysis. i.e., Sample, NTC, PosCtrl, or NegCtrl					
	Chromosome	Chromosome number in human genome where variant occurs					
Variant Report	Position	Base position on given chromosome where variant occurs (hg19)					
	REF	Sequence found in human reference genome (hg19)					
	ALT	Variant sequence found in sample					
	Location	Genomic coordinates of variant (hg19)					
	Variant_Type	SNV = single nucleotide variant; Deletion = one or more base deletion; Insertion=one or more base insertion; Delins = "deletion- insertion", variant characterized by both a deletion of one or more bases AND an insertion of one of more bases on the same allele. i.e., AGACTA> ATTCTA resulting from deletion of GA and insertion of TT or AGACTA> ATCTA resulting from deletion of GA and insertion of T					
	VARIANT_CLASS	Sequence Ontology (SO) terms to identify variant type					
	Variant_Length	Number of affected bases					
	Variant_Net_Length	The net length between the difference of a variant between the length of the reference sequence and the alternate sequence. Variant_Net_Length = (length of alternate sequence – length of Reference sequence)					
	Amplicon_ID	Amplicon in assay in which variant occurs.					
	Variant_Read_Frequency_ (%)	Occurrence of variant in sequencing data, as a percentage of total sequenced reads in the given segment. (Variant_Coverage / Total_Coverage)*100					
	Variant_Coverage	Number of reads that contain variant in segment.					
	Total_Coverage	Total number of reads in segment.					
	Variant_Quality	Measure of variant call quality. Score is from 0 (low quality) to 40 (highest quality)					



Sheet Tab	Column Name	Definition/Description					
	Variant_Read_Direction_R atio	Measure of strand bias in variant calling. Value of 0 indicates variant was found on both the positive strand and negative strand at equal rates (ideal). Values greater than zero or less than zero indicate percentage of bias towards the negative strand (negative values) or positive strand (positive values)					
	Zygosity	Germline mutation panels only. HETEROZYGOUS = one out of two alleles contain variant; HOMOZYGOUS = both alleles contain variant					
	Consequence	Type of mutation. Examples: frameshift_variant = variant produces an mRNA transcript that is out of the normal reading frame. synonymous_variant = variant does not affect amino acid sequence of final protein due to redundancy of codons. missense_variant = variant leads to change in amino acid sequence. inframe_deletion = deletions of 3, or multiples of 3, which remove whole codon sequences. intron_variant = variant occurs in intron. splice_region_variant = variant in splice site, may lead to splicing error in final mRNA transcript. stop_gained = variant results in premature stop codon in mRNA transcript. 3_prime_UTR_variant = variant occurs in 3" untranslated region (UTR)					
	Impact	Impact of variant on normal gene function					
	Gene_Symbol	Gene name					
	Gene_ID	NCBI gene ID number					
	Feature	NCBI transcript identifier					
	All_Features	List of all NCBI transcript identifiers associated with the variant					
	HGVSC	HGVSC variant name. Describes position and variant change in gene (nucleic acid) sequence in Feature					
	HGVSP	HGVSP variant name. Describes position and variant change in protein (amino acid) sequence in Feature					
	Transcript	The VEP annotated transcript name (e.g., NM_001127500.3)					
	c_dot	VEP HGVSc annotation with transcript name removed (e.g., `c.3912C>T`)					
	p_dot	VEP HGVSp annotation with transcript removed (e.g., `p.Asp1304=`)					
	Exon	Affected exon number in Feature out of total number of exons. Applies to variants within exon regions only. (affected exon # / total # of exons)					



Sheet Tab	Column Name	Definition/Description					
	Intron	Affected intron number in Feature out of total number of introns. Applies to variants within intron regions only. (affected intron # / total # of introns)					
	CDS_Position	cDNA sequence position in Feature					
	Protein_Position	Affected amino acid number in Feature					
	Amino_Acids	Amino acid resulting from variant (reference amino acid / variant amino acid)					
	Codons	Position in codon where variant occurs and resulting codon with variant					
	Co- located_Known_Variation	Variants that occur at the same position in publicly available databases					
	COSMIC_LINK	URL link to the COSMIC database for the called variant					
	Strand	Transcribed strand1 = negative strand, 1 = positive strand					
	Repeat	Number of repeat bases, including repeat sequence (if applicable). i.e. 14G indicates a repeat of 14 Gs					
	HGVS_Offset	Correction factor to synchronize genomic base position for indels in repeat regions in negative strand genes. Read alignment uses left-align paradigm (left most base assumed to be indel; first genomic position) while HGVS uses right-align paradigm (right most base assumed to be indel; first cDNA position)					
	SIFT	Output from SIFT (v5.2.2). PREDICTION (SIFT score). Predicts impact of variant on protein function. Prediction possibilities: tolerated, deleterious. SIFT score is a range from 0 - 1. Scores > 0.05 are tolerated, scores of < 0.05 are deleterious					
	POLYPHEN	Output from PolyPhen (v2.2.2). PREDICTION (PolyPhen score). Predicts impact of variant on protein function. Prediction possibilities: benign, possibly damaging, probably damaging. PolyPhen score is a range from 0 (benign) - 1 (probably damaging) and is the probability of variant leading to a damaging mutation. Benign = variant unlikely to cause damage to protein function. Possibly damaging = variant likely to cause damage to protein function but prediction is with low confidence. Probably damaging = variant likely to cause damage to protein function is with high confidence.					
	AF	Allele frequency of variant found in global population (1000 Genomes version Phase 3)					
	AFR_AF	Allele frequency of variant found in African populations (1000 Genomes version Phase 3)					



Sheet Tab	Column Name	Definition/Description					
	AMR_AF	Allele frequency of variant found in American populations (1000 Genomes version Phase 3)					
	EAS_AF	Allele frequency of variant found in East Asian populations (1000 Genomes version Phase 3)					
	EUR_AF	Allele frequency of variant found in European populations (1000 Genomes version Phase 3)					
	SAS_AF	Allele frequency of variant found in South Asian populations (1000 Genomes version Phase 3)					
	gnomAD_AF	Allele frequency of variant found in global population (gnomAD version r2.1)					
	gnomAD_AFR_AF	Allele frequency of variant found in African/African American populations (gnomAD version r2.1)					
	gnomAD_AMR_AF	Allele frequency of variant found in Latino/Admixed American populations (gnomAD version r2.1)					
	gnomAD_ASJ_AF	Allele frequency of variant found in Ashkenazi Jewish populations (gnomAD version r2.1)					
	gnomAD_EAS_AF	Allele frequency of variant found in East Asian populations (gnomAD version r2.1)					
	gnomAD_FIN_AF	Allele frequency of variant found in European (Finnish) populations (gnomAD version r2.1)					
	gnomAD_NFE_AF	Allele frequency of variant found in European (non-Finnish) populations (gnomAD version r2.1)					
	gnomAD_OTH_AF	Allele frequency of variant found in Other populations (gnomAD version r2.1)					
	gnomAD_SAS_AF	Allele frequency of variant found in South Asian populations (gnomAD version r2.1)					
	CLIN_SIG	Clinical significance (ClinVar v201912)					
	PiVAT_ClinSig	PiVAT clinical significance based on CLIN_SIG and IMPACT					
Filtered False Positives	This sheet reports all false p	ositive variants that do not pass PiVAT's filtering parameters.					
Quercl	Total Reads	Total number of sequencing reads for the library					
Overall Stats	Overall:Q=30	Percentage of bases with Q score greater than or equal to 30					
	Overall:Q=20	Percentage of bases with Q score greater than or equal to 20					



Sheet Tab	Column Name	Definition/Description					
	Properly Paired Read	Number of read mates properly paired, from paired end sequencing					
	Properly Paired Read (%)	Percentage of total reads properly paired					
	Mapped Reads	Percentage of total reads that map to human genome (hg19)					
	Mapping Rate (%)	Percentage of total reads that map to human genome (hg19)					
	On Target Reads	Reads that map to target amplicon regions of interest (ROIs)					
	On Target Rate (%)	Percentage of mapped reads that map to target amplicon regions of interest (ROIs)					
	Insert Size Mean	Mean size of the library insert (varies based on panel)					
	Insert Size Median	Median size of the library insert					
	Insert Size Std Dev	Standard deviation of library insert size					
	Coverage_Mean	Mean base coverage of all bases within the defined ROI; with paired end sequencing, merged paired-end reads (forward and reverse) create a coverage of 1x					
	STDEV	Standard deviation of mean base coverage					
	Coverage_Median	Median base coverage of all bases within the defined ROI; with paired end sequencing, merged paired-end reads (forward and reverse) create a coverage of 1x					
	Coverage_Max	Maximum base coverage of all bases within the defined ROI					
	Coverage_Min	Minimum base coverage of all bases within the defined ROI					
	Total_Number_Of_Reads	Total number of reads on target for all amplicons					
	Total_Valid_Reads	Total number of reads contributing to paired end assembly after filtering					
	On_Target_Ratio	The ratio between Total_Valid_Reads and Total_Number_Of_Reads. On_Target_Ratio = (Total_Valid_Reads / Total_Number_Of_Reads)					
	Base_Coverage_Depth_>_ (Nx)	Percent of bases that have a minimum base coverage greater than or equal to Nx (absolute coverage)					
	Base_Coverage_Depth_>_ (Nx)_ Relative_to_Mean_Covera ge	Percent of bases that obtain at least Nx*mean base coverage, usually described as percent of mean base coverage. Used to determine uniformity of base coverage across ROIs in panel. Value of 100 indicates 100% of bases in all ROI are above the given Nx relative to the Coverage_Mean.					



Sheet Tab	Column Name	Definition/Description
Segment Coverage	This sheet reports the Cover samples initially added to th identifier), Region (ROI segn GC_Content_(%) (percentag	rage_Mean, Coverage_Min and STDEV of each amplicon ROI for all ne analysis. Additional columns include Target_Name (amplicon ROI nent position), Segment_Size (base pair size of ROI segment), and ge of GC content within ROI segment).



Appendix B: Microsatellite Instability (MSI) Caller

Microsatellite instability (MSI) is measured using MSIsensor (<u>https://github.com/ding-lab/msisensor</u>) in conjunction with Pillar's custom MSI46 panel. This software compares the distribution of read lengths at 53 microsatellite (MS) sites in the tumor genome with the same sites in a matched normal reference. MSI status (instable—MSI-high, or stable—MSS) is determined by the number of sites which differ.

MSI Sample Setup

The user must provide a matched normal sample for use as a reference in determining each microsatellite site's baseline read length distribution. Ideally this sample should be prepared with the same methodology and sequenced in parallel with its corresponding tumor sample.

MSI Analysis Setup

This section provides specific instructions to start an MSI analysis. Please refer to

Start Analysis section on page 14 for general instructions on how to start a PiVAT analysis.

- 1. Select **Start Analysis** on the left menu bar.
- 2. From the Start Analysis: 1. Start From page select the **MSI46** panel from the drop-down menu and select the desired samples to include in analysis.
- 3. Select **NEXT** to proceed.
- 4. From the Start Analysis: 2. Edit Definition page enter desired **Analysis Name**.
- 5. The samples you have selected will appear in the table as shown below.



Figure 32 Start Analysis: 2. Edit Definition page – MSI Analysis



6. For each tumor/normal pair, (A) check the boxes to the left of the Sample Name in the table and (B) select the **Tumor Normal Paired** button. This will link the two input files and they will appear at the bottom of the table. Ensure that the files are defined correctly. (C) The definitions can be swapped by clicking the button to the left of the files.



Figure 33 Start Analysis: 2. Edit Definition page – define tumor/normal sample pair

- 7. Once all the samples have been paired select **NEXT**.
- 8. Review the analysis input and select Launch Analysis to begin the run. The task status may now be monitored on the PiVAT Dashboard.



PiVAT Output: MSI_RESULTS file

Definition and/or description of result columns reported in MSI_RESULTS file sheet tabs are provided below.

Sheet Tab	Column Name	neDefinition/DescriptionoleUnique Sample ID for each tumor sample in a matched pairupleUnique Sample ID for each normal sample in a matched pairoble_TypeThe sample type associated with the tumor sample from a matched pairople_TypeThe sample type associated with the normal sample from a matched pairople_TypeThe sample type associated with the normal sample from a matched pairople_TypeThe sample type associated with the normal sample from a matched pairople_TypeThe number of MS sites that are different between the tum and the matched normal sampleSomatic_SitesThe number of MS sites that are different between the tum and the matched normal samplese_RatioThe ratio of differing MS sites to total MS sites assessedobleUnique Sample ID for each tumor sample in a matched pairople_TypeThe sample type associated with the tumor sample from a matched pairople_TypeThe sample type associated with the normal sample from a matched pairople_TypeThe sample type associated with the normal sample from a matched pairople_TypeThe sample type associated with the normal sample from a matched pairople_TypeThe sample type associated with the normal sample from a matched pairople_TypeThe sample type associated with the normal sample from a matched pairople_TypeThe sample type associated with the normal sample from a matched pairople_TypeThe sample type associated with the normal sample from a matched pairople_TypeThe sample type associated with the normal				
	Tumor_Sample	Unique Sample ID for each tumor sample in a matched pair				
	Normal_Sample	Unique Sample ID for each normal sample in a matched pair				
	Tumor_Sample_Type	The sample type associated with the tumor sample from a matched pair				
MSI Call	Normal_Sample_Type	The sample type associated with the normal sample from a matched pair				
Somatic Sites	MSI_Call	Whether the tumor sample exhibits microsatellite instability (True = MSI-High, False = MSS)				
	Total_Number_of_Sites	The number of MS sites assessed for a tumor/normal pair				
	Number_of_Somatic_Sites	The number of MS sites that are different between the tumor and the matched normal sample				
	Somatic_Sites_Ratio	The ratio of differing MS sites to total MS sites assessed				
	Tumor_Sample	Unique Sample ID for each tumor sample in a matched pair				
	Normal_Sample	Unique Sample ID for each normal sample in a matched pair				
	Tumor_Sample_Type	The sample type associated with the tumor sample from a matched pair				
	Normal_Sample_Type	The sample type associated with the normal sample from a matched pair				
	Target_Name	The amplicon ID covering the MS site				
Read Count	Chromosome	The chromosome the MS site is located on				
Read Count Distribution	MSI_Start	The location of the MS start site (hg19)				
	Left_Flank	The 5-bp sequence to the left (upstream) of the MS site				
	Repeat_Unit_Bases	The reference number of repeats and the [repeating sequence unit] comprising the MS site				
	Right_Flank	The 5-bp sequence to the right (downstream) of the MS site				
	Number_of_Repeat_Units	The number of possible repeating units (1-100)				
	Tumor_Read_Count	The number of reads in the tumor sample with a repeat length specified by Number_of_Repeat_Units field				
Somatic	Normal_Read_Count	The number of reads in the normal sample with a repeat length specified by Number_of_Repeat_Units field				
Sites	Tumor_Sample	Unique Sample ID for each tumor sample in a matched pair				
	Normal_Sample	Unique Sample ID for each normal sample in a matched pair				



Sheet Tab	Column Name	Definition/Description					
	Tumor_Sample_Type	The sample type associated with the tumor sample from a matched pair					
	Normal_Sample_Type	The sample type associated with the normal sample from a matched pair					
	Target_Name	The amplicon ID covering the MS site					
	Chromosome	The chromosome the MS site is located on					
	MSI_Start	The location of the MS start site (hg19)					
	Left_Flank	The 5-bp sequence to the left (upstream) of the MS site					
	Repeat_Times	The reference number of repeats of the repeating sequence unit					
	Repeat_Unit_Bases	The bases comprising the repeating unit					
	Right_Flank	The 5-bp sequence to the right (downstream) of the MS site					
	Difference	The difference score (0-1) assigned to the MS site between tumor and normal					
	P_Value	The p-value associated with the difference score					
	FDR	The FDR-adjusted p-value associated with the difference score					
	Rank	The rank, by p-value, of the MS site					
	Tumor_Sample	Unique Sample ID for each tumor sample in a matched pair					
	Normal_Sample	Unique Sample ID for each normal sample in a matched pair					
	Tumor_Sample_Type	The sample type associated with the tumor sample from a matched pair					
	Normal_Sample_Type	The sample type associated with the normal sample from a matched pair					
	Target_Name	The amplicon ID covering the MS site					
Germline	Chromosome	The chromosome the MS site is located on					
Sites	MSI_Start	The location of the MS start site (hg19)					
	Left_Flank	The 5-bp sequence to the left (upstream) of the MS site					
	Repeat_Times	The reference number of repeats of the repeating sequence unit					
	Repeat_Unit_Bases	The bases comprising the repeating unit					
	Right_Flank	The 5-bp sequence to the right (downstream) of the MS site					
	Genotype	The repeat lengths of alleles in the Normal sample					



Appendix C: Somatic CNV Caller

There are two copy number variation (CNV) callers implemented within PiVAT, to handle FFPE and cfDNA panels respectively. The FFPE CNV calls are reported at gene- and exon-levels, cfDNA calls are currently restricted to gene-level only. One CNV call is displayed as one row in the report. All amplicons of the panel are used in the CNV analysis.

CNV Sample Setup

FFPE CNV caller: For the best CNV caller performance, **3 to 5 (minimum 2) normal samples** to be used as normalization reference is recommended. These normal samples should have similar sample condition and preparation process as the potentially positive samples. See section below for recommended setting for CNV analysis.

cfDNA CNV caller: The cfDNA CNV caller requires at least one CNV normal sample, with the normal samples having similar sample condition and preparation process as the potentially positive samples. The quality of the calls improves with increasing numbers of normal samples.

CNV Analysis Setup

This section provides specific instructions to start a CNV analysis. Please refer to

Start Analysis section on page 14 for general instructions on how to start a PiVAT analysis.

- 1. Select Start Analysis on the left menu bar.
- 2. From the Start Analysis: 1. Start From page, select a desired CNV panel from the drop-down menu and select the desired samples to include in analysis.
- 3. Select Next to proceed.
- 4. From the Start Analysis: 2. Edit Definition page, enter desired **Analysis Name**.
- 5. The samples you have selected will appear in the table as shown below.
- 6. To define normal samples, (A) check the boxes to the left of the Sample Name in the Edit Definition table and (B) select the **Define CNV Normal Samples** button.
- If using the cfDNA CNV Caller (currently only applies to the P-LBX-01 panel): Tumor Content % should be specified if known. In the Tumor Content % column enter a value from 1e-8 to 100, representing the tumor content as a percentage for the sample on that row. If this is left empty the tumor content will either be inferred or given a default (CNV_DEFAULT_TUMOR_CONTENT_LOWER, CNV_DEFAULT_TUMOR_CONTENT_UPPER may be overridden when defining a Custom Parameter for P-LBX-01).



PILLAR	Dashboard	Analysis	Tools						O usermanual	• 泓 Engl	lish ? Help
Parameters		倄 / Analysis /	Start Analysis / Edit Definition								
🗱 Start Analysis		Start Analy	sis								
Analysis Results			1. Start From		2. Edit	: Definit	tion		3. Preview &	Launch	
		Analysis Name:	Please input analysis definition nan	ne	-20230604-18174	1 Para	meters File:	Select			\ \
		Tumor-Normal F	Pailed B Define CNV Normal S	Samples 🛨	Edit Files 🗫	Skip S	amples 🕤			In	nsert Sample
			Sample Name 🦊		QC Тур	e	S	Sample Files	Tumor Content %	Oper	ation
	_ []	Undetern	nined-d221107iM1		Sample	•	Undeterm	iined-d2211 (2) 🗳		Edit Files 🗭	Skip Sample 🔊
		A 🛛 RDvSMp0	Core2Ad221107iM1-SD10-CNV36-1	x-NTC	Sample	•	RDvSMpC	ore2Ad2211 (2) 🗳		Edit Files 🗭	Skip Sample 🕲
		RDvSMp0	Core2Ad221107iM1-SD10-CNV36-1	x-AM	Sample	•	RDvSMpC	Core2Ad2211 (2) 🗳		Edit Files 🔁	Skip Sample 🔊
		RDvSMp0	Core2Ad221107iM1-SD10-CNV36-1	p5x-AM	Sample	•	RDvSMpC	Core2Ad2211 (2) 🗳		Edit Files 🌮	Skip Sample 🖸
		RDvSMp0	Core2Ad221107iM1-SD10-CNV36-C	lp5x-AM	Sample	•	RDvSMpC	Core2Ad2211 (2) 🗳		Edit Files 🕫	Sample 🖸
		RDvSMp0	Core2Ad221107iM1-SD10-AM		Sample	¥	RDvSMpC	Core2Ad2211 (2) 🕰		Edit Files	Skip Sample
	•				Back	N	lext				

Figure 34 Start Analysis: 2. Edit Definition page – define CNV normal sample(s)

8. The selected samples will be grouped as **Normal Samples**. Samples can be added or removed from the list of Normal Samples using the same **Define CNV Normal Samples** button or the **Undefine Normal Sample** buttons.

PLLAR	Dashboard	Analysis	Tools					O usermanual	• 🕱 English	? Help
Parameters			1. Charle France		0 E.H.	2. Edit Definition			Laura da	
🛠 Start Analysis			1. Start From		2. Edit I				5. Preview & Launch	
Analysis Results		Analysis Name:	Please input analysis definitio	on name	-20230604-181741	Parameter	rs File: Select			\sim
		Tumor-Normal	Paired 🗐 🛛 Define CNV No	ormal Samples 🛨	Edit Files 醚	Skip Sample	es 🔊		Insert S	iample
			Sample Name 🎍		QC Type		Sample Files	Tumor Content %	Operation	
		RDvSMp	Core2Ad221107iM1-SD10-CN	W36-1p5x-AM	Sample	₩ RD	wSMpCore2Ad2211 (2) 🗳		Edit Files 🚧 Skip Samp	_{le}
	RDvs		Core2Ad221107iM1-SD10-CN	W36-0p5x-AM	Sample	₩ RD	IvSMpCore2Ad2211 (2) 🗳		Edit 💤 Skip Files Samp	_{le} D
		RDvSMpCore2Ad221107iM1		1	Sample	₩ RD	IvSMpCore2Ad2211 (2) 🖴		Edit 😿 Skip Files Samp	_{le}
		Norma	l Samples						[]] Undefi	ne All
			Core2Ad221107iM1-SD10-CN	W36-1x-AM	Sample	₩ RD	IvSMpCore2Ad2211 (2) 🗳		iles Windefin Sample	^e (/)
		RDvSMp	Core2Ad221107iM1-SD10-CN	W36-1x-NTC	Sample	₩ RD	IvSMpCore2Ad2211 (2) 😫		iles Cundefin Sample	° E/3
		Undeter	mined-d221107iM1		Sample	₩ Un	determined-d2211 (2) 🗳		iles Undefin Sample	° 6/3
	٢				Back	Next				

Figure 35 Start Analysis: 2. Edit Definition page – Normal Samples group

- 9. For FFPE panels: If no normal negative reference samples are provided, at least 10 potentially positive samples are required to run CNV analysis. In this setting, the CNV analysis will run with center percentile normalization algorithm. The sample setting configuration will be automatically detected by PiVAT. No user parameter adjustment is needed. In addition, the number of samples with the same CNV type cannot exceed 30% of all samples. We recommend pooling samples with different CNVs (CNVs with different exons/genes/lengths) in this setting.
- 10. For PiVAT 2023.1, there is a post-call CNV filter to filter out CNV calls that fail to pass the gene specific thresholds set in the Post-Filtering section of Analysis Parameters (see page 12). Each CNV panel is designed



with a specific set of genes the panel targets, with thresholds for confident CNV calls on those genes. The default values can be seen in the same Post-Filtering section when creating a new Analysis Parameter.

11. The normalized copy number ratio can be found in 'Normalized Coverages' sheet in CNV_RESULTS excel file. Note that the "Copy Number Ratio" in PiVAT is defined as the copy number ratio of a potentially positive sample to that of the negative reference samples (diploid with 2 copy number).

CNV QC parameters: FFPE panels

The following is a list of user adjustable QC parameters and default values in ().

1. CNV_ALLOW_FEWER_THAN_MIN_NUM_SAMPLES (False)

This parameter allows running with fewer than 10 samples and no negative reference samples. The default value is "False". When the parameter is toggled to True, CNV analysis still needs at least 3 samples. Only toggles on for intended low sample number run mode if number of total remaining samples are between 3 and 10 and number of negative samples is less than 2.

2. CNV_QC_ZERO_COVERAGE_PCT_THRESHOLD_NORMAL_SAMPLE (0.5)

Percentage of zero coverage targets needed for a negative reference sample to pass the negative reference QC filter. The default value is 0.5, which means a negative reference sample passes the negative reference QC filter if less than 50% of its targets have zero coverage.

3. CNV_QC_ZERO_COVERAGE_PCT_THRESHOLD_POS_SAMPLE (0.5)

Percentage of zero coverage targets needed for a potentially positive sample to pass the sample QC filter. The default value is 0.5, which means a sample passes the sample QC filter if less than 50% of its targets have zero coverage.

4. CNV_QC_ABSOLUTE_COVERAGE_THRESHOLD_NORMAL_SAMPLE (50)

Mean segment coverage needed for a negative reference sample to pass the negative reference QC filter. The default value is 50, which means a negative reference sample should have the mean segment coverage >= 50 to pass the negative reference QC filter.

5. CNV_QC_ABSOLUTE_COVERAGE_THRESHOLD_POS_SAMPLE (10)

Mean segment coverage needed for a potentially positive sample to pass the sample QC filter. The default value is 10, which means a sample should have the mean segment coverage >= 10 to pass the sample QC filter.

6. CNV_QC_RELATIVE_COVERAGE_PCT_THRESHOLD_NORMAL_SAMPLE (0.1)

Minimum mean segment coverage percentage compared to other samples' mean segment coverage needed for a negative reference sample to pass the negative reference QC filter. The default value is 0.1, which means a negative reference sample should have the mean segment coverage >= 10% of other normal samples' mean segment coverage to pass the negative reference QC filter.



7. CNV_QC_RELATIVE_COVERAGE_PCT_THRESHOLD_POS_SAMPLE (0.1)

Minimum mean segment coverage percentage compared to other samples' mean segment coverage needed for a potentially positive sample to pass the sample QC filter. The default value is 0.1, which means a sample should have the mean segment coverage >= 10% of the CNV normal samples' mean segment coverage to pass the sample QC filter.

8. CNV_QC_LOW_COVERAGE_PCT_THRESHOLD_NORMAL_SAMPLE (0.1)

Percentage of low coverage targets needed for a negative reference sample to pass the negative reference QC filter. The default value is 0.1, which means a negative reference sample should have <= 10% of low coverage targets to pass the negative reference QC filter.

9. CNV_QC_Q30_THRESHOLD (80)

Minimum threshold for Quality 30 for both normal and potentially positive samples to pass the QC filter. The default value is 80, which means a sample (positive or negative) should have overall Q30 >= 80 to pass the sample QC filter.

10. CNV_QC_ON_TARGET_RATE_THRESHOLD (90)

Minimum On Target Rate threshold for both negative and potentially positive samples to pass the QC filter. The default value is 90, which means a sample (positive or negative) should have an On Target Rate >= 90 to pass the sample QC filter.

11. CNV_QC_CORRELATION_COEFFICIENT_THRESHOLD_NORMAL_SAMPLE (0.5)

Minimum correlation coefficient of a negative reference sample with the rest of other negative reference samples for that sample to pass the negative reference sample QC filter. The default value is 0.5, which means a negative reference sample should have a correlation coefficient >= 0.5 to pass the negative reference QC filter.

12. CNV_QC_CORRELATION_COEFFICIENT_THRESHOLD_POS_SAMPLE (0.2)

Minimum correlation coefficient of a potentially positive sample with the negative reference samples for that sample to pass the sample QC filter. The default value is 0.2, which means a sample should have a correlation coefficient >= 0.2 to pass the sample QC filter.

13. CNV_QC_OUTLIER_CENTER_PERCENTILE_POS_SAMPLE (40)

The center percentile is used to compare the correlation and relative coverage for potentially positive samples. If CNV normal samples are provided, the center percentile method is not used. Otherwise, the correlation and relative coverage is calculated against the center samples for each amplicon.

cfDNA CNV Caller QC parameters:

The following is a list of user adjustable QC parameters and default values in ().

1. CNV_MINIMUM_AMPLICON_COVERAGE (5000)

Minimum amplicon coverage below which to exclude amplicons from CNV analysis. Values are integers from 0 to 1e12.



2. CNV_MINIMUM_KEPT_AMPLICON_PROPORTION (0.8)

If fewer than this proportion of amplicons survive QC fail this sample/normal pair with `Amplicon Drop Out`. Values are floats from 0.0 to 1.0.

cfDNA CNV Caller Model Parameters:

We have provided substantial flexibility to control nearly every aspect of the cfDNA CNV Caller. However, these parameters should only be modified with the help of Pillar support, requiring familiarity with Bayesian nonparametric statistics and inference, and experience with the Pillar cfDNA CNV Caller.

3. CNV_CLUSTER_ALPHA (0.01)

The concentration parameter of the base measure, with larger values leading to more high probability clusters. Values are floats between 1e-5 and 1.0.

4. CNV_DEFAULT_TUMOR_CONTENT_LOWER (1.0)

When no user-specified tumor content is provided, this value is used as the lower bound for sample tumor content percentage. This will be used only when the Tumor Content (%) field in Edit Definition is empty. Values are floats, representing percentages, from 1e-6 to 100.0.

5. CNV_DEFAULT_TUMOR_CONTENT_UPPER (30)

When no user-specified Tumor Content (%) is provided, this value is used as the upper bound for sample tumor content percentage. This will be used only when the Tumor Content (%) field in Edit Definition is empty. Values are floats, representing percentages, from 1e-6 to 100.0, and should be set to be larger than CNV_DEFAULT_TUMOR_CONTENT_LOWER.

6. **CNV_AUTO_INFER_SAMPLE_TUMOR_CONTENT (False)**

Whether to automatically infer the expected tumor content from in-sample SNP/Indel calls. **Experimental**.

7. CNV_USE_METADATA_SAMPLE_TUMOR_CONTENT (True)

Whether to use user-specified sample Tumor Content (%) in Edit Definition, when provided. Possible values are True and False, with True indicating that user-specified sample Tumor Content (%) will be used.

8. CNV_METADATA_SAMPLE_TUMOR_CONTENT_LOWER_FACTOR (0.8)

This factor is multiplied by the Tumor Content (%) provided in Edit Definition to create the lower bound estimate for sample tumor content percentage. Values are floats. Only applies for samples that have user-defined Tumor Content (%) specified in Edit Definition.

9. CNV_METADATA_SAMPLE_TUMOR_CONTENT_UPPER_FACTOR (1.2)

This factor is multiplied by the Tumor Content (%) provided in Edit Definition to create the upper bound estimate for sample tumor content percentage. Values are floats. Only applies for samples that have user-defined Tumor Content (%) specified in Edit Definition.

10. CNV_MAXIMUM_NUMBER_OF_COPY_NUMBER_CLUSTERS (10)

The maximum number of CNV clusters to fit. Values are integers >= 3.



11. CNV_CLUSTER_COPY_NUMBER_DISPERSION (1.0)

The spread (e.g., variance) parameter of the cluster copy number distribution. Only has an effect when CNV_FIXED_COPY_NUMBER is False. For the Poisson distribution this is the mean * CNV_CLUSTER_COPY_NUMBER_DISPERSION, set by an affine transformation. For the GammaPoisson distribution this is the rate parameter, set directly. Values are floats.

12. CNV_CLUSTER_COPY_NUMBER_CONCENTRATION (2.0)

The concentration parameter (e.g., mean) of the cluster copy number distribution. Only has an effect when CNV_FIXED_COPY_NUMBER is false. For the Poisson distribution this is the mean. For the GammaPoisson distribution this is the shape parameter, set directly. Values are floats.

13. CNV_SORT_INITIAL_VALUES (False)

Whether or not to sort the initial values for the copy number clusters. Values are True or False, with sorting enabled when True. Only applies when CNV_INIT_TO_2 and CNV_INIT_TO_UNIFORM are False.

14. CNV_AMPLICON_NOISE_DISPERSION (1.0)

Control the spread of background amplicon centered log copy number ratio noise estimates. Smaller values mean less variation. Values are non-negative floats. Only has an effect when CNV_FIX_NOISE is False.

15. CNV_CLUSTER_COPY_NUMBER_DISTRIBUTION ('GammaPoisson')

Base measure of copy number. Possible values are 'Poisson' and 'GammaPoisson' corresponding to a Poisson distribution and the compound Gamma Poisson distribution respectively. This is the distribution from which each cluster tumor copy number is drawn. The 'GammaPoisson' option allows for higher variance compared to 'Poisson', making large copy numbers more likely possibilities.

16. CNV_FIX_TC (True)

Whether to treat estimated tumor content as fixed or infer tumor content. Setting CNV_FIX_TC to False allows for more Bayesian characterization of uncertainty in the call, resulting in wider credible intervals as opposed to a (fixed) point estimate. Possible values are True and False.

17. CNV_FIX_NOISE (True)

Whether to treat observational noise as fixed. It is strongly recommended that this parameter be set to True. Setting to False can result in excessive shrinkage by the Dirichlet process resulting in a single cluster (no calls). Possible values are True and False.

18. CNV_FIX_CLUSTER_COPY_NUMBER (False)

Whether to treat tumor DNA copy number as fixed. When True, tumor DNA copy number will be taken directly from initial values, and should only be used when CNV_INIT_TO_2 is False, as otherwise all tumor copy numbers will be trivially 2. Possible values are True and False.

19. CNV_MAX_COPY_NUMBER (202.0)

The maximum possible tumor DNA copy number that the cfDNA CNV Caller is allowed to infer. The tumor DNA copy number is the idealized copy number in tumor cells (as opposed to the sample tumor copy number, which is the result of a mixture of normal and tumor DNA). In low



tumor content settings, this parameter, alongside the provided tumor content tunes signal to noise ratio. Values are floats and must be >= CNV_MIN_COPY_NUMBER.

20. CNV_MIN_COPY_NUMBER (1.6)

The minimum tumor DNA copy number. Values close to 2 prevent calling copy number deletions , but improve the ability to detect copy number amplifications. Values are floats.

21. CNV_DISCRETE_HMC_GIBBS (True)

Whether to use a Discrete Gibbs HMC MCMC kernel as opposed to a standard kernel. This results in slower sampling but has desirable theoretical/numerical properties. Possible are True and False, with the Discrete Gibbs HMC MCMC kernel enabled when True.

22. CNV_PARAMETER_STRATEGY ('median')

This defines the averaging strategy of the posterior samples. The default, 'median', is robust to outliers. Possible values are 'median', 'mean', and 'mode'.

23. CNV_GENE_STRATEGY ('mean')

Strategy for averaging amplicons within a gene into a gene-level sample copy number ratio estimates. Possible values are 'median', 'mean', and 'mode'.



24. CNV_TUMOR_CONTENT_STRATEGY ('median')

Strategy for averaging tumor content estimates across the MCMC samples. Only important if tumor content is inferred rather than fixed. Possible values are 'median', 'mean', and 'mode'.

25. CNV_AGGREGATION_STRATEGY ('mean')

Strategy for performing within-sample parameter aggregation from the ensemble of individual putative case / CNV Normal fits. Only applies when more than 1 CNV Normal is provided. Possible values are 'median', 'mean', 'min', 'max'.

26. CNV_SEED (2023)

The random number generator seed, to ensure reproducibility. Values are integers.

27. CNV_GENE_SMOOTHING_WEIGHT (0.7)

The amount of count averaging within a gene. 0 correspond to no averaging (raw counts for each amplicon) while 1.0 corresponds to every amplicon within the gene having identical counts (average counts). Values are floats from 0.0 to 1.0.

28. CNV_CONSIDER_TOP_N_GENE_CALLS (20)

How many of the top (sorted in descending order) amplicon copy number values to consider when averaging amplicons within a gene to generate the gene copy number ratio call. Values are integers from 0 to 100,000.

29. CNV_SHOW_NEUTRAL_CALLS (True)

Whether or not to report genes that are below the Gain or Loss copy number ratio threshold defined for those genes. See page 13 (CNV Threshold Grid) for details on how to edit these thresholds through the PiVAT webapp. When **CNV_SHOW_NEUTRAL_CALLS** is True, genes with inferred copy number ratios below the copy number ratio threshold will appear in the CNV Calls sheets of CNV_RESULTS and CUSTOMER_REPORT, with the label "Neutral" in the Relative Call Gain(Loss) column. Possible values are True and False.

30. CNV_INIT_TO_UNIFORM (False)

Instead of initializing the Cluster Copy Number to random values between 2 and CNV_MAX_COPY_NUMBER, initialize them to a uniformly distributed random point in the support of the Cluster Copy Number distribution. Possible values are True and False. Cannot be set at the same time as CNV_INIT_TO_2 or CNV_INIT_TO_VALUES.

31. CNV_INIT_TO_2 (True)

Initialize all Cluster Copy Number values to 2. Possible values are either True or False. Cannot be set at the same time as CNV_INIT_TO_UNIFORM or CNV_INIT_TO_VALUES.

32. CNV_INIT_TO_VALUES (None)

Initialize the Cluster Copy Number to values specified in this comma-separated list of numbers. Example: For 2 CNV_MAXIMUM_NUMBER_OF_COPY_NUMBER_CLUSTERS, `2,2` would initialize both clusters to 2. Values are strings of comma-separated integers or floating-point numbers, with as many values as specified CNV_MAXIMUM_NUMBER_OF_COPY_NUMBER_CLUSTERS. Cannot be set at same time as CNV_INIT_TO_UNIFORM or CNV_INIT_TO_2.

33. CNV_AVERAGE_ALL_NORMALS (False)



When multiple normal samples are selected, whether to average the amplicon coverages of those normals and fit the putative case samples against that averaged "normal", instead of fitting the putative case samples against individual normals and averaging the resulting parameter estimates. This results in cheaper computation time (single model fit rather than multiple for unaveraged normal) but can more easily overpolish. Possible values are True and False.



PiVAT Output: CNV_RESULTS file

Definition and/or description of result columns reported in CNV_RESULTS file sheets are provided below.

Sheet Tab	Column Name	Definition/Description			
	Sample ID	Unique Sample ID for each sample.			
	Gene ID	Simplified gene name for the called CNV segment.			
	Copy Number Ratio	The ratio of the copy number of called CNV segment to the copy number (2) of a negative normal sample.			
	Relative Gain/Loss	Relative Gain or Loss label of the called CNV segment. If the copy number ratio is more than 1, it is labeled as "Gain". Otherwise, it is labeled as "Loss".			
CNV Calls	Std Dev	Standard deviation calculated from the copy number ratios of the amplicons in the called CNV segment.			
	P-Value Amplicon Count	P-value of the called CNV segment.			
		Count of amplicons in the called CNV segment.			
	CNV Start	The estimated start location of the called CNV segment.			
	CNV End	The estimated end location of the called CNV segment.			
	Exon List	The list of the exons in the called CNV segment. Only reported for FFPE caller.			
	QC Criteria	Indicate which QC criteria that each row is reporting, including "Negative_Reference", "Sample" and "Run_Status".			
CNV Run QC	QC Status	Whether the QC is passed for each QC criteria. If QC is passed, it is labeled as "Pass". If QC is failed, it is labeled as "Fail". The "Run_Status" is labeled with the actual setting that CNV analysis is run with unless the CNV run is completely failed, where it is labeled as "Fail".			
	Sample ID	Unique Sample ID for each sample			
Filtered CNV	Filter Reason	The reason that the sample is filtered out from the CNV analysis.			
Samples	Sample Type	Indicate whether the reported sample is "Sample" or "Negative Reference". This is pre-defined by the user before starting the SA analysis.			
CNV Segment Coverages	This sheet reports the r added to the CNV analy Sample IDs.	raw segment coverages of each amplicon for all the samples initially ysis. The column headers are Amplicon Names and the row indices are			



Sheet Tab	Column Name	Definition/Description
Normalized Coverages	This sheet reports the r are included in the CNV The column headers are IDs; the row indices are	normalized copy number ratio of each amplicon for the samples that analysis. Note that the filtered samples are not reported in this sheet. the covered region of each amplicon ("Region") plus all the Sample Amplicon Names.

CNV Calls Table

The CNV Calls table contains the CNV calls for the sample(s) selected in Select Sample.

	Dashboard	Analysis Too	ols Admin					🖸 admin 🔹 🖄 English	? Help
Parameters		🅱 / Analysis / Analy	sis Results / Analysis Ta	ask					
🛠 Start Analysis		Analysis Task	R283_CNV_Test-2023060	02-153459					
Analysis Results		Download Result 2	ip Files 🚺 Logfile	Task Information					🖰 Rerun
		QC Summary C	C Stat Variant	Report					
		SNV / Indel	CNV						
		Select Sample:	Select Sample: QCvSMpBRCAFFFEgCNVd200813IM3-BRCNV1-20ng-Beginning X						
		Select Reference:	QCvSMpBRCAFFPEgCNVd.	200813iM3-TV12-20ng-End	×			×	
		Filtered Sel	ect Variant Filter: Syste	em Default 🛛 👻					
		CNV Calls							
		Sample ID	Gene ID	Copy Number Ratio	Relative Gain(Loss)	Std Dev	P-Value	Amplicon Coun	nt
		QCvSMpBRCAFFPEg.	BRCA2	1.26	Gain	0.04	1.16e-06	3	i i
		QCvSMpBRCAFFPEg.	BRCA2	1.36	Gain	0.04	3.32e-07	3	i i
		QCvSMpBRCAFFPEg.	BRCA1	1.67	Gain	0.15	0.e+00	4	į.
									2
	4	CNV RUN QC:	QC Criteria	QC Status					
			Negative_Reference	Pass					

Figure 36 Analysis Results: CNV Calls Table

Column NameDefinition/DescriptionSample IDUnique Sample ID for each sample.Gene IDSimplified gene name for the called CNV segment.Copy Number RatioThe ratio of the copy number of called CNV segment to the copy number (2) of a
negative normal sample.Relative Gain/LossRelative Gain or Loss label of the called CNV segment. If the copy number ratio is
more than 1, it is labeled as "Gain". Otherwise, it is labeled as "Loss".Std DevStandard deviation calculated from the copy number ratios of the amplicons in the
called CNV segment.

Definition and/or description of result columns reported in the CNV Calls table are provided below.



Column Name	Definition/Description
	For FFPE caller: P-value of the called CNV segment.
P-Value	For cfDNA caller : 0 if cfDNA CNV Caller considers the copy number to be non-neutral, or 1 otherwise.
Amplicon Count	Count of amplicons in the called CNV segment.
CNV Start	The estimated start location of the called CNV segment.
CNV End	The estimated end location of the called CNV segment.
Exon List	The list of the exons in the called CNV segment.

CNV Filtering

CNV Calls in the CNV Calls Table can be filtered using the **Filtered** toggle displayed below. If no custom parameter is assigned to this task, **System Default** is selected. If the task was run with a custom Parameter, that custom filter will be selected.

Filtered Select	t Variant Filter: Systen	n Default 🛛 🔻				
CNV Calls						
Sample ID	Gene ID	Copy Number Ratio	Relative Gain(Loss)	Std Dev	P-Value	Amplicon Count
QCvSMpBRCAFFPEg	BRCA2	1.26	Gain	0.04	1.16e-06	3
QCvSMpBRCAFFPEg	BRCA2	1.36	Gain	0.04	3.32e-07	3
QCvSMpBRCAFFPEg	BRCA1	1.67	Gain	0.15	0.e+00	4

Figure 37 Analysis Results: CNV Call Filtering

Apply a custom filter preset by selecting it in the **Select Variant Filter** dropdown. The CNV Calls table will update to show only the CNV Calls that pass the selected filter. Clicking **Save** will assign the chosen filter to this task and **regenerate the PDF reports for this task.** Clicking **Discard** will switch to the selection to the filter that is assigned to the task.

Filtered	Select Variant Filter:	CNV_FILTER_TEST	▼ Save	Discard		
CNV Calls						
Sample ID	Gene ID	Copy Number Ratic	Relative Gain(Loss)	Std Dev	P-Value	Amplicon Count
QCvSMpBRCAFF	PEg BRCA2	1.26	Gain	0.04	1.16e-06	3
QCvSMpBRCAFF	PEg BRCA2	1.36	Gain	0.04	3.32e-07	3
QCvSMpBRCAFF	PEg BRCA1	1.67	Gain	0.15	0.e+00	4

Figure 38 Analysis Results: CNV Call Filtering – Saving the new filter



If the Filtered toggle is off, all CNV calls will appear in the CNV Calls table with no filtering.

CNV Plot

A scatter plot will be displayed for each selected CNV sample. Each point represents the copy number ratio of a target.

	Dashboard	Analysis	Tools	Admin		🖸 admin 🔻	落 English	? Help
Parameters		QC Summary	QC Stat	Variant	Report			
🛠 Start Analysis		SNV / Indel	CNV					
III Analysis Results		Select Sample:	QCvSM	pBRCAFFPEgCN	Vd200813iM3-BRCNV1-20ng-Beginning X		×	~
- ,		Select Reference	CVSM	pBRCAFFPEgCN	Vd200813iM3-TV12-20ng-End X		×	~
			Sam	ple	Pass			
			Run	Status	Pass			
	۲	Sample:QCV Colored Guidel Scale Type: •	SMpBRC ines: Lower Linear	AFFPEgCN Limit 0.8 - Log	Vd200813iM3-BRCNV1-20ng-Beginning Upper Limit 12 Y Axis Range: Min (Default:0): 0 Max (Default:2): 2		nload CNV F	lot

Figure 39 Analysis Results: CNV Plot



Genes can be toggled on/off by clicking the name of the gene in the legend. In the example below, the **BRCA1** gene has been hidden.

		Analysis	Tools	Admin		🖸 admin 🝷	🕆 English	? ⊦
Parameters		QC Summary	QC Stat	Variant	Report			
Start Analysis		SNV / Indel	CNV					
Analysis Result	s	Select Sample:	QCvSMp	BRCAFFPEgCNV	d200813iM3-BRCNV1-20ng-Beginning X		×	~
i indigoio recourt	-	Select Reference	e: QCvSMp	BRCAFFPEgCNV	d200813iM3-TV12-20ng-End 🗙		×	~
		Sample:QC\	/SMpBRC/	۹ ۶ ۶ ۶ ۶ ۶	vd200813iM3-BRCNV1-20ng-Beginning	Dev	unload CNIV	lat
		Colored Guidel	lines: Lower	Limit: 0.8 🗸	Upper Limit: 1.2 • Y Axis Range: Min (Default:0): 0 Max (Default:2): 2			101
		Scale Type: 🔾	l Linear 🔵	Log				
		1.6			• • • •			
		1.4	•	•				
		1.3						
		1.1					٠	
		0.9	• •					
		0.8						
		0.8						
		0.8 0.7 0.6 0.5						
		0.8 0.7 0.6 0.5 0.4						
		0.8 0.7 0.6 0.5 0.4 0.3 0.2						
		0.8 0.7 0.6 0.5 0.4 0.3 0.2 0.1						
		0.8 0.7 0.6 0.5 0.4 0.3 0.2 0.1 0.0				1 1 1 1 1		
		0.8 0.7 0.6 0.5 0.4 0.3 0.2 0.1 0.0 0.7 0.4 0.3 0.2 0.1 0.0 0.5 0.4 0.5 0.4 0.5 0.5 0.4 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5	Lower Limi	t			1 1 1 1	1

Figure 40 Analysis Results: CNV Plot – Toggle Gene



The plot can be saved as a PNG by clicking the blue **Download CNV Plot** button. Any change made to the plot will be reflected in the saved image.



Figure 41 Analysis Results: CNV Sample Plot – Save Plot as PNG

Note that the plot shows the copy number ratios of all targets in the genes of interest for the panel. The data points on the plot do not get filtered out by the selected preset filter.

CNV Panel Specific Notes

Notes in this section pertain to specific panels. Any given note below can be ignored if your PiVAT installation does not include the particular panel.



oncoReveal LBx Core (P-LBX-01) utilizes the **cfDNA caller**. **P-LBX-01 requires at least 1 CNV Normal to be specified** (using Define CNV Normal Samples button in Edit Definition). If sample tumor content is known, it should be specified in the **Tumor Content %** column (as a percentage) for improved CNV calling performance.



For the oncoReveal Multi-Cancer with CNV v4 (HS341), if running cell-line or germline samples, it is recommended to create a custom Parameter to filter copy gain calls at 1.20, copy loss calls at 0.8, and amplicon count at 3 amplicons instead of the default values.



For the oncoReveal BRCA1 & BRCA2 plus CNV Panel (BR283), if running somatic samples, it is recommended to create a custom Parameter to filter copy gain calls at 1.20 and copy loss calls at 0.80 instead of the default values.



 $\mathbf{\Lambda}$

For the oncoReveal BRCA1 & BRCA2 plus CNV Panel (BR283), when viewing the CNV results in the CNV Calls table, there may be multiple CNV calls for the same gene and same exons. This is expected, as the panel was designed with both a gene caller and an exon-level cluster caller, and a copy gain or loss may be detected in multiple clusters within a given exon or gene. The CNV Start and CNV End positions can be used to distinguish the calls apart, and the custom Parameter filter can be used to isolate pertinent CNV calls.



Appendix D: Thalassemia CNV Caller

The Thalassemia analysis is based on the double normalization method, including one per-sample normalization and one per-amplicon normalization. The normalization baseline is calculated from negative reference samples. The copy number variance (CNV) is detected by amplicon clusters and the Thalassemia type of each sample is called by matching the edges of a CNV with the edges of a list of common Thalassemia types.

Thalassemia CNV Sample Setup

For each Thalassemia CNV analysis run, the user should provide 3-5 (minimum 2) in-run normal (negative) reference samples with similar sample condition and preparation process as the positive samples. If less than 2 negative reference samples are provided, the run will fail.

See <u>CNV Sample Setup</u> section in <u>Appendix C: Somatic CNV Caller</u> for instructions to define normal samples to be used as negative references.

Thalassemia CNV Analysis Setup

- 1. See <u>CNV Analysis Setup</u> section in <u>Appendix C: Somatic CNV Caller</u> for instructions to setup a Thalassemia CNV analysis.
- 2. The THAL_RESULTS excel file is output upon completion of the analysis.
- 3. Detailed Thalassemia call information are in the "Thalassemia Calls" sheet of the THAL_RESULTS excel file.
- 4. The fully normalized copy number ratios can be found in "Fully Normalized" sheet in THAL_RESULTS excel file. Note that the copy number ratio in PiVAT is defined as the copy number ratio of a potentially positive sample to that of the negative reference samples (diploid with 2 copy number).

Thalassemia CNV QC parameters

See <u>CNV QC parameters</u> section in <u>Appendix C: Somatic CNV Caller</u> for a list of user adjustable QC parameters.



PiVAT Output: THAL_RESULTS file

Definition and/or description of result columns reported in THAL_RESULTS file sheets are provided below.

Sheet Tab	Column Name	Definition/Description
	Sample ID	Unique Sample ID for each sample.
Thalassemia	Thalassemia Call Status	Whether each sample is called as Thalassemia CNV positive by PiVAT. "TRUE" means a sample is Thalassemia CNV positive while "FALSE" indicates a sample is Thalassemia CNV negative.
Summary	QC Status	Whether each sample passes the sample QC. Sample is labeled as "Pass" if passing sample QC and "Fail" if failing sample QC.
	Sample Type	Indicate whether a sample is defined as "Sample" or "Negative_Reference" by the user.
	Sample ID	Unique Sample ID for each sample.
	Thal Type	The Thalassemia type detected by the "Thal Type Caller" for a sample.
	Copy Number Ratio	The ratio of the copy number of called Thalassemia CNV to the copy number (2) of a negative normal sample.
	P-Value	P-value of the called Thalassemia CNV.
Thalassemia Calls	Relative Gain/Loss	Relative Gain or Loss label of the called CNV segment. If the copy number ratio is more than 1, it is labeled as "Gain". Else (when the copy number ratio is less than 1) it is labeled as "Loss".
	CNV Start Min	The estimated minimum start location of the called Thalassemia CNV.
	CNV Start Max	The estimated maximum start location of the called Thalassemia CNV.
	CNV End Min	The estimated minimum end location of the called Thalassemia CNV.
	CNV End Max	The estimated maximum end location of the called Thalassemia CNV.
Thalassemia	QC Criteria	Indicate which run QC criteria that each row is reporting, including "Negative_Reference", "Sample" and "Run_Status".
Run QC	QC Status	Whether the QC is passed for each run QC criteria. When run QC is passed, it is labeled as "Pass". When run QC is failed, it is labeled as "Fail".
	Sample ID	Unique Sample ID for each sample
THAL Filtered	Filter Reason	The reason that the sample is filtered out from the Thalassemia analysis.
Samples	Sample Type	Indicate whether the reported sample is "Sample" or "Negative_Reference". This is pre-defined by the user before starting the SA analysis.



Sheet Tab	Column Name	Definition/Description		
CNV Segment Coverages	This sheet reports the raw segment coverages of each amplicon for all the samples initially added to the Thalassemia analysis. The column headers are Amplicon Names and the row indices are Sample IDs.			
Normalized Coverages	This sheet reports the pass the sample QC ar filtered samples are n and the row indices ar	normalized copy number ratio of each amplicon for the samples that nd are actually included in the Thalassemia analysis. Note that the ot reported in this sheet. The column headers are Amplicon Names re Sample IDs.		

Thalassemia Type Calls Table

Thalassemia Type Calls table contains all calls for the sample(s) selected in Select Sample.

	Dashboard	Analysis Too	ols Admin					🚺 admin 🔻 🦻	🛦 English 🛛 🥐 Help
Parameters		倄 / Analysis / Analy	sis Results / Analysis Tas	k					
🛠 Start Analysis		Analysis Task s	MA_CNV_Test-20230602-1	155203					
III Analysis Results		Download Result Z	Zip Files 🚺 Logfile	Task Information					C Rerun
		QC Summary C	QC Stat Variant F	Report					
		SNV / Indel	CNV SMA						
		Select Sample:	RDvGMpTHAL172d201202if	N1-PB09 ×					x ~
		Select Reference:	RDvGMpTHAL172d201202if	N1-PB05 ×					x ~
		Thalassemia Type	e Calls						
		Sample ID	Thal Type	Copy Number Ratio	P-Value	Relative Gain(Loss)	Chromosome	CNV Start Min	CNV Start Max
		RDvGMpTHAL172d2.	CNV Call	0.00	0.e+00	Loss	chr5	69366608	69372394
									,
		Thalassemia Run Q	C: QC Criteria	QC Status					
			Negative_Reference	Pass					
			Sample	Pass					
			Run_Status	Pass					
	¢	Sample:RDvGf Colored Guideline: Scale Type: • Li	MpTHAL172d201202 s: Lower Limit: 0.8 ▼ near ○ Log	iN1-PB09 Upper Limit: 1.2 -	Y Axis Range: Min I	(Default:0): 0	Max(Default2.1): 2.1	Downl	oad CNV Plot

Figure 42 Analysis Results: Thalassemia Type Calls Table

Definition and/or description of result columns reported in the Thalassemia Type Calls table are provided below.

Column Name	Definition/Description
Sample ID	Unique Sample ID for each sample.
Thal Type	The Thalassemia type detected by the "Thal Type Caller" for a sample.



Copy Number Ratio	The ratio of the copy number of called Thalassemia CNV to the copy number (2) of a negative normal sample.		
P-Value	P-value of the called Thalassemia CNV.		
Relative Gain/Loss	N/Loss Relative Gain or Loss label of the called CNV segment. If the copy number ratio is more than 1, it is labeled as "Gain". Otherwise, it is labeled as "Loss".		
Chromosome	The chromosome the Thalassemia type call is located on.		
CNV Start Min	The minimum start position of the detected thalassemia.		
CNV Start Max	The maximum start position of the detected thalassemia.		
CNV End Min	The minimum end position of the detected thalassemia.		
CNV End Max	The maximum end position of the detected thalassemia.		

Thalassemia Plot

See <u>CNV Plot</u> section in <u>Appendix C: Somatic CNV Caller</u> for a description of CNV/Thalassemia sample plots, and how to save plots as PNG files.

Note that for the Thalassemia plot, the copy number ratios are grouped by chromosome and not by the target genes. The chromosomes can be toggled to show or hide the data points, similar to the CNV plot.





Appendix E: SMA Caller

The SMA analysis is based on the double normalization method. The normalization baseline is calculated from negative reference samples. The SMA Caller calculates the copy number ratios of Exon-07 and Exon-08 amplicons on the SMN1 and SMN2 genes.

SMA Sample Setup

For each SMA analysis run, the user should provide 3-5 (minimum 2) in-run normal (negative) reference samples with similar sample condition and preparation process as the positive samples. If less than 2 negative reference samples are provided, the run will fail.

See <u>CNV Sample Setup</u> section in <u>Appendix C: Somatic CNV Caller</u> for instructions to define normal samples to be used as negative references.

SMA Analysis Setup

- 5. See <u>CNV Analysis Setup</u> section in <u>Appendix C: Somatic CNV Caller</u> for instructions to setup a SMA Thalassemia analysis.
- 6. The SMA_RESULTS excel file is output upon completion of the analysis.
- 7. Detailed SMA call information is in the "SMA Call Report" sheet of the SMA_RESULTS excel file.
- 8. Each sample's detailed SMA Report can be downloaded as a PDF file in the "Report" section of the results page.
- 9. The fully normalized copy number ratios can be found in "Fully Normalized" sheet in SMA_RESULTS excel file. Note that the copy number ratio in PiVAT is defined as the copy number ratio of a potentially positive sample to that of the negative reference samples (diploid with 2 copy number).

SMA QC parameters

See <u>CNV QC parameters</u> section in <u>Appendix C: Somatic CNV Caller</u> for a list of user adjustable QC parameters.



PiVAT Output: SMA_RESULTS file

Definition and/or description of result columns reported in SMA_RESULTS file sheets are provided below.

Sheet Tab	Column Name	Definition/Description		
SMA Call Report	Sample ID	Unique Sample ID for each sample.		
	Gene-Exon	Name of the gene/exon pair, or if a control target, name of the control amplicon.		
	Location	Genomic coordinates of exon (hg19).		
	Copy Number	Copy number of the gene-exon or of the control amplicon		
	Copy Number Ratio	The copy number ratio of the gene-exon or control-amplicon to the copy number (2) of a negative normal sample.		
SMA Run QC	QC Criteria	Indicate which run QC criteria that each row is reporting, including "Negative_Reference", "Sample" and "Run_Status".		
	QC Status	Whether the QC is passed for each run QC criteria. If run QC is passed, it is labeled as "Pass". It run QC is failed, it is labeled as "Fa		
Filtered SMA Samples	Sample ID	Unique Sample ID for each sample		
	Filter Reason	The reason that the sample is filtered out from the SMA Thalassemia analysis.		
	Sample Type	Indicate whether the reported sample is "Sample" or "Negative Reference". This is pre-defined by the user before starting the SA analysis.		
CNV Segment Coverages	This sheet reports the raw segment coverages of each amplicon for all the samples initially added to the SMA analysis. The column headers are Amplicon Names and the row indices are Sample IDs.			
Normalized Coverages	This sheet reports the normalized copy number ratio of each amplicon for the samples that pass the sample QC. Note that the filtered samples are not reported in this sheet. The column headers are Amplicon Names and the row indices are Sample IDs.			



SMA Calls Table

SMA Calls table contains the copy number ratios for the sample(s) selected in **Select Sample**.

PILLAR PILLAR EDashboard	Analysis Tools	Admin			🖸 admin 🝷 🕉 English 🕜 Help
Parameters	🕱 / Analysis / Analysis R	esults / Analysis Task			
🛠 Start Analysis	Analysis Task SMA_	CNV_Test-20230602-155203			
iii Analysis Results	📀 Download Result Zip Fi	les 🚺 Logfile 🖻 Task Information			C Rerun
	QC Summary QC St	at Variant Report			
	SNV / Indel Cf	IV SMA			
	Select Sample:	vGMpTHAL172d201202iN1-PB09 ×			x ~
	Select Reference: RD	vGMpTHAL172d201202iN1-PB05 ×			x ~
	SMA Calls				
	Sample ID	Gene-Exon	Location	Copy Number	Copy Number Ratio
	RDvGMpTHAL172d20120	2iN1-PB SMN1_Ex07	chr5:70241792-70248150	3.92	1.96
	RDvGMpTHAL172d20120	2iN1-PB SMN1_Ex08	chr5:70248390-70248554	3.92	1.96
	RDvGMpTHAL172d20120	2iN1-PB SMN2_Ex07	chr5:69366366-69372729	0.00	0.00
	RDvGMpTHAL172d20120	2iN1-PB SMN2_Ex08	chr5:69372969-69373133	0.00	0.00
	RDvGMpTHAL172d20120	2iN1-PB CtrL01.ZNF648.1Q25.3	chr1	2.04	1.02
	RDvGMpTHAL172d20120	2iN1-PB CtrL02.AOX1.2Q33.1	chr2	2.01	1.01
	RDvGMpTHAL172d20120	2iN1-PB CtrL03.PLCL2.3P24.3	chr3	2.00	1.00
	RDvGMpTHAL172d20120	2iN1-PB CtrL04.PDLIM5.4Q22.3	chr4	2.18	1.09
	RDvGMpTHAL172d20120	2iN1-PB CtrL05.TBC1D19.4P15.2	chr4	1.97	0.98
<	RDvGMpTHAL172d20120	2iN1-PB CtrL06.PJA2.5Q21.3	chr5	2.31	1.16
			· -		

Figure 43 Analysis Results: SMA Calls table

Definition and/or description of result columns reported in the SMA Calls table are provided below.

Column Name	Definition/Description
Sample ID	Unique Sample ID for each sample.
Gene-Exon	Name of the gene/exon pair, or if a control target, name of the control amplicon.
Location	Genomic coordinates of exon (hg19).
Copy Number	Copy number of the gene-exon or of the control amplicon
Copy Number Ratio	The ratio of the copy number of called Thalassemia CNV to the copy number (2) of a negative normal sample.



SMA Plot

A SMA plot for each sample will be located on the Variant > SMA tab below the SMA Calls table and SMA RUN QC table in Analysis Results. Each boxplot represents the copy number ratios of all CNV Normal Samples specified in the run, and the orange data points represent the copy number ratios of the potentially positive sample. The x-axis represents gene-exon labels, and the y-axis represents copy number ratio.



Figure 44 Analysis Results: SMA Sample Box Plot


To save the box plot graph, hover over the graph, and click the "Download plot as a png" icon as indicated below.



Figure 45 Analysis Results: SMA Sample Box Plot – Saving as PNG



Appendix F: Fusion Caller

The Solid Fusion V2 panel detects common fusion transcripts in a simple, multiplex reaction. The output provides likely fusion transcripts by gene and exon pairs. Only fusions described in the Product Sheet for a given panel will be reported by that panel. Fusions detected, which do not match any described in the Product Sheet, will be filtered from results and recorded in the debug log.

Fusion Sample Setup

Control samples are not required for this panel.

Fusion Analysis Parameters Setup

1. Navigate to ① Secondary Analysis Parameters page to access the following list of adjustable parameters and default values in () by tab.

Tab	QC Parameters [default]		
	Minimum Distance to Breakpoint (12)		
Fusion BWA Alignment	-T parameter input into BWA. Sets a different T value when running bwa which can help getting a shorter alignment as secondary alignment. Decreasing this parameter may increase noise but may help rescue fusion calls for shorter read lengths		
	Fusion Cutoff Threshold (200)		
	Minimum supporting reads for a positive fusion call. Usually far below what would be needed for a positive call due to the normalized fusion count cutoff. May need to be adjust if the normalized cutoff is significantly decreased.		
	Normalized Fusion Cutoff Threshold (250)		
	Minimum normalized reads for a positive fusion call.		
	NormalizedCount = (Fusion count * Constant) / mapped read count		
Fusion Sample	Constant = 100,00		
Summary	Decreasing this value increases sensitivity but may result in more false positive calls. Increasing this value will decrease sensitivity resulting in possible false negative calls.		
	Total Read Count Threshold (100,000)		
	Samples with less total reads than this parameter will fail QC. Decrease if sample input is low. May result in poor quality samples with fusion calls.		
	Internal Control Threshold (1,500)		
	The sum of the counts for TBP HBMS in the IMBALANCE RATIO sheet. Samples with less than 1500 reads for both targets combined will fail QC		



PiVAT Output: CUSTOMER_RESULTS file

Definition and/or description of result columns reported in REPORT_FUSION file sheets are provided below.

Sheet Tab	Column/Row Name	Definition/Description
Fusions	SAMPLE_NAME	Name of sample
	FusionName	Fusion ID = LeftGene(LeftExon):RightGene(RightExon)
	Count	Number of reads traversing fusion breakpoint
	NormalizedCount	Count / MappedReads * Constant
	HGVS Notation	Fusion ID in HGVS Nomenclature
	Breakpoint	Fusion breakpoint LeftPosition, RightPosition
	LeftStrand	Orientation of left gene. 1 positive strand, -1 negative strand
	RightStrand	Orientation of right gene. 1 positive strand, -1 negative strand
IMBALANCE RATIO	SAMPLE_NAME	Name of sample
	Gene	Target gene for balance reads
	Туре	Internal control or 3'/5' comparison
	Count	Number of reads on left Number of reads on right
	Ratio	Imbalance score = (3' balance reads – 5' balance reads)/ reads in control gene
		"Positive" or "Negative" or "No Call" for evidence of fusion
	Detection	Note: This is not factored into the Fusions sheet. If positive calls exist it provides support for a fusion event with this gene. The partner should be in the fusion sheet if it was included in the panel's targets.
	Target_Name	Name of left target Name of right target
RUN STATS	Read Count	The total raw reads in the sample's fastq files
	Assembled	Total count of forward and reverse reads that were paired and merged during paired end assembly. each pair count as 2.
	% Assembled	Percentage of Assembled reads out of Read Count: • (Assembled / Read Count) * 100
	Mapped Reads	Total number of reads that were successfully aligned by bwa.



Sheet Tab	Column/Row Name	Definition/Description
	Manusius Data (9/)	Percentage of Mapped Reads out of Total Reads Count:
	Mapping Rate (%)	 (Mapped Reads / Read Count) * 100
		Reads that have a secondary alignment are considered as fusion reads because they are candidate to a junction point between the primary and secondary alignment.
	Eucion Boads	These reads are counted as following:
	Fusion Reads	 Assembled reads: if there is a secondary alignment: count as 2
		 paired end reads: each read with secondary alignment: count as 1
	FILTERED: Single Side Fusions With Mate Aligned On Different Gene	Unassembled paired end reads where the forward read detected a different junction point than the reverse read.
	FILTERED: Dimers	Reads filtered because fusion was detected to be a dimer.
	FILTERED: Misprimings	Reads filtered because the aligned read was detected to be a mispriming
	Total Valid Fusion Sequences	Total number of reads that cover a junction after filtering.
	Balance Reads	The total number of balance reads.
	FILTERED: Unique Reads Mapped Off Target	All reads that are aligned to untargeted coordinates.
	Total Valid Balance Sequences	Total number of balance reads that are not filtered. Reads are filtered if they are not on targeted regions.
	Total Reads After Filtering	The total reads that were not filtered.
	% Total Unfiltered Reads	Percentage of Total Unfiltered Reads out of Total Reads Count.
		(Total Unfiltered Reads / Total Reads Count) * 100
QC FILTERED SAMPLES	Reason	Reason sample did not pass QC thresholds



Appendix G: FASTQ File Name Format

A full description of the Illumina FASTQ Name Format can be found under the "Naming" section here:

https://support.illumina.com/help/BaseSpace_OLH_009008/Content/Source/Informatics/BS/NamingConvention_n_FASTQ-files-swBS.htm



It is not necessary to name FASTQ files following the Illumina naming convention to analyze them through PiVAT.



The following terms should be avoided as sample names: "*common*" and "*logs*". Sample names should not vary by case only.



Sample names are case insensitive when working in Windows or MacOS environments but, case sensitive when working in Linux. This could result in files being hidden, or overwritten.

Do not use any patient identifiable information in any file names.