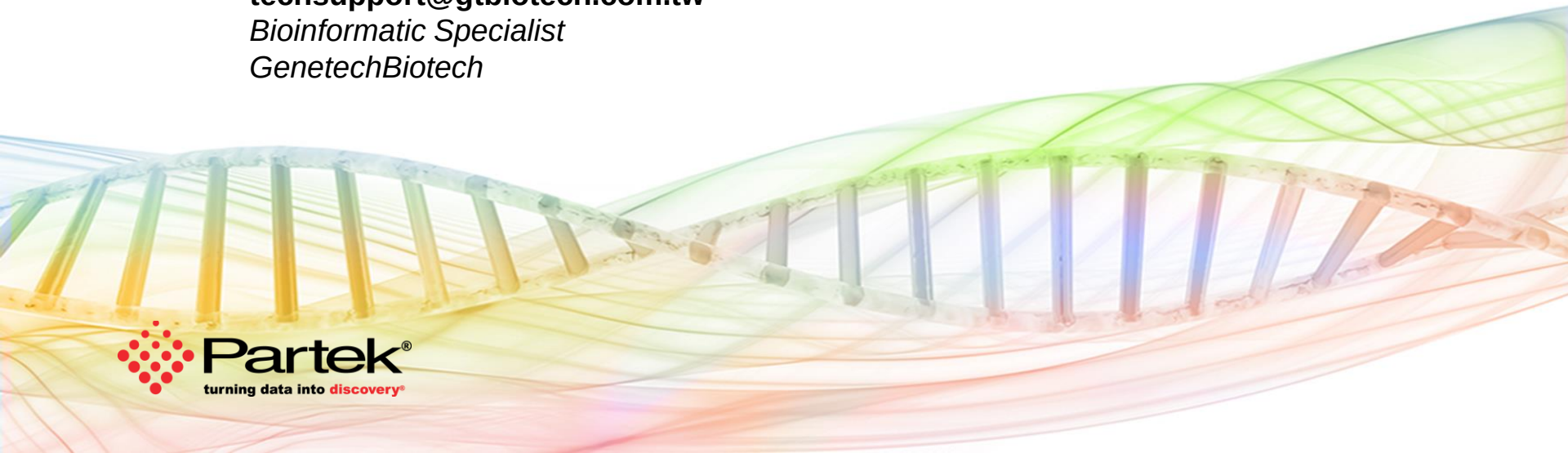
From Raw Data to Pathways: Easy Genomics Analysis with Partek Flow

Tanya Chiu
techsupport@gtbiotech.com.tw
Bioinformatic Specialist
GenetechBiotech



Partek®
turning data into discovery®

Who is Partek

Mission

To empower scientists to make scientific breakthroughs in human genetics, disease relationships, drug discoveries, diagnoses, and disease treatments.

Partek® Flow®

Partek® Pathway

Founded in

1993

for data mining and artificial intelligence

Over

8,500

peer-reviewed citations

More than

40,000

researcher questions answered

Customers in over

40

countries

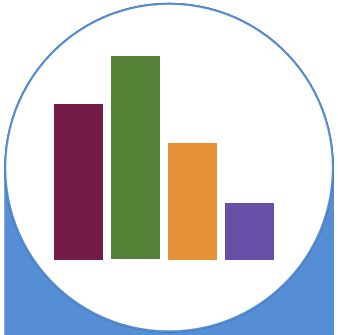
Copyright © 2024 Partek, an Illumina company. All rights reserved.
For Research Use Only.
Not for use in diagnostic procedures.

Partek™ an Illumina company

Partek Flow: Start-to-Finish Bioinformatics Solution



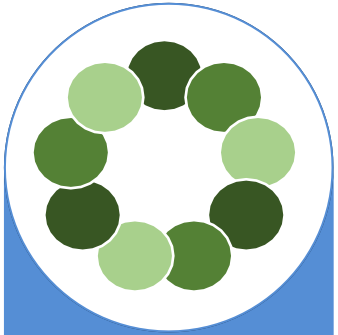
Collaborative Web Environment



Powerful Statistics



Interactive Visualizations



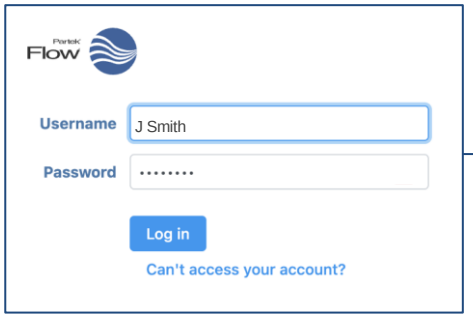
Comprehensive Application Support

Copyright © 2024 Partek, an Illumina company. All rights reserved.
For Research Use Only.
Not for use in diagnostic procedures.

Partek™ an Illumina company

User Friendly Analysis and Visualizations

Access from Your Favorite Browser





Copyright © 2024 Partek, an Illumina company. All rights reserved.
For Research Use Only.
Not for use in diagnostic procedures.

Partek™ an Illumina company

Publicly Available Statistical Algorithms and Tools

Alignment
Bowtie
BWA
Isaac
TopHat
TMAP
Bowtie
GSNAP
STAR
HISAT

Differential analysis
Limma
DESeq2
Poisson
Negative binomial
Non-parametric ANOVA

Metagenomics
Kraken
Alpha and beta diversity
Quantification at taxonomic levels
Differential analysis at taxonomic levels

QA/QC reports
Pre-alignment
Post-alignment
ERCC spike-in
Single cell quality

Clustering
Hierarchical
K-means
Graph-based

Data exploration
PCA
t-SNE
Dot plot
Box plot
Pathway
Bar chart
Bubble map
Heat map
Violin plot
Histograms
Chromosome view
2D & 3D Scatter Plot
Pie chart
UMAP

Variant calling
Samtools
LoFreq
CNVkit
FreeBayes
Strelka
GATK

Variant annotation
SnpEff
dbSNP
VEP
Custom databases

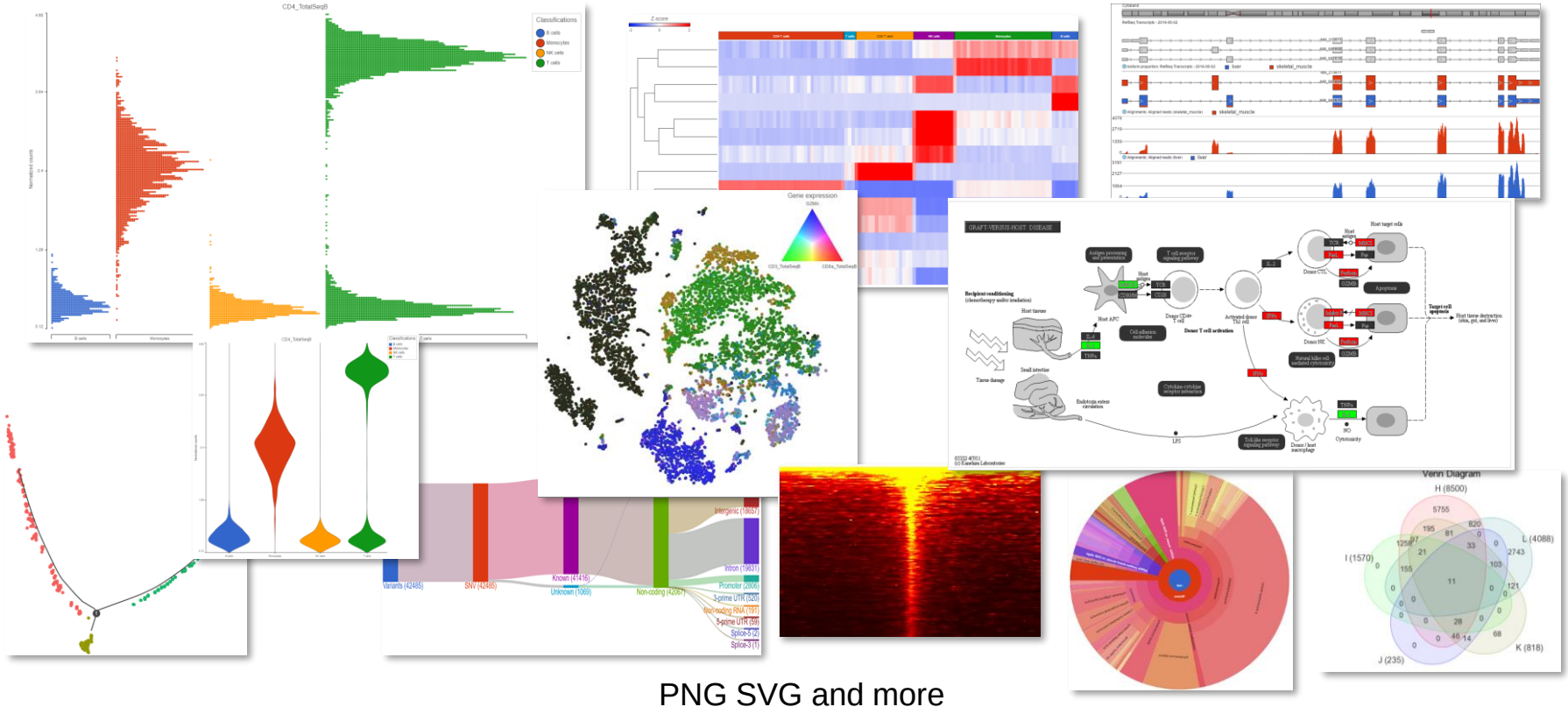
Peak calling
MACS2
TSS plot
Motif detection

Quantification
Partek E/M
HTSeq
Cufflinks

Copyright © 2024 Partek, an Illumina company. All rights reserved.
For Research Use Only.
Not for use in diagnostic procedures.

Partek™ an Illumina company

Compelling and Publishable Visualizations



PNG SVG and more

Copyright © 2019 Partek Incorporated. All rights reserved.

Partek®

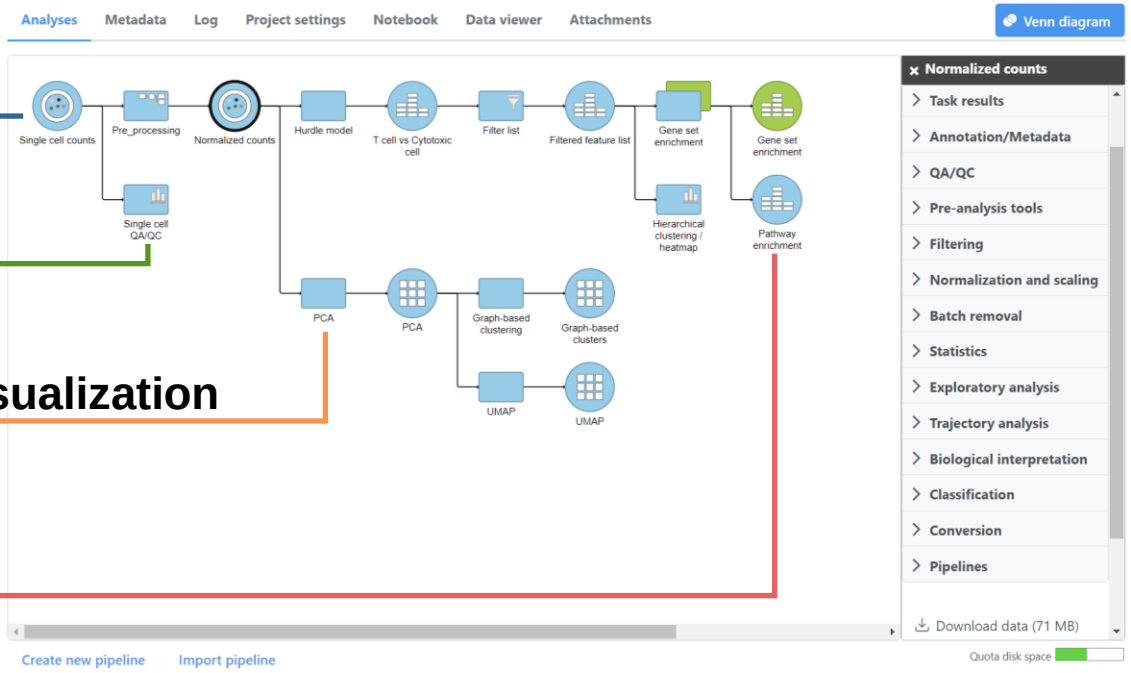
Visual Analysis Process

1 Import Data

2 QA/QC

3 Powerful Statistics & Visualization

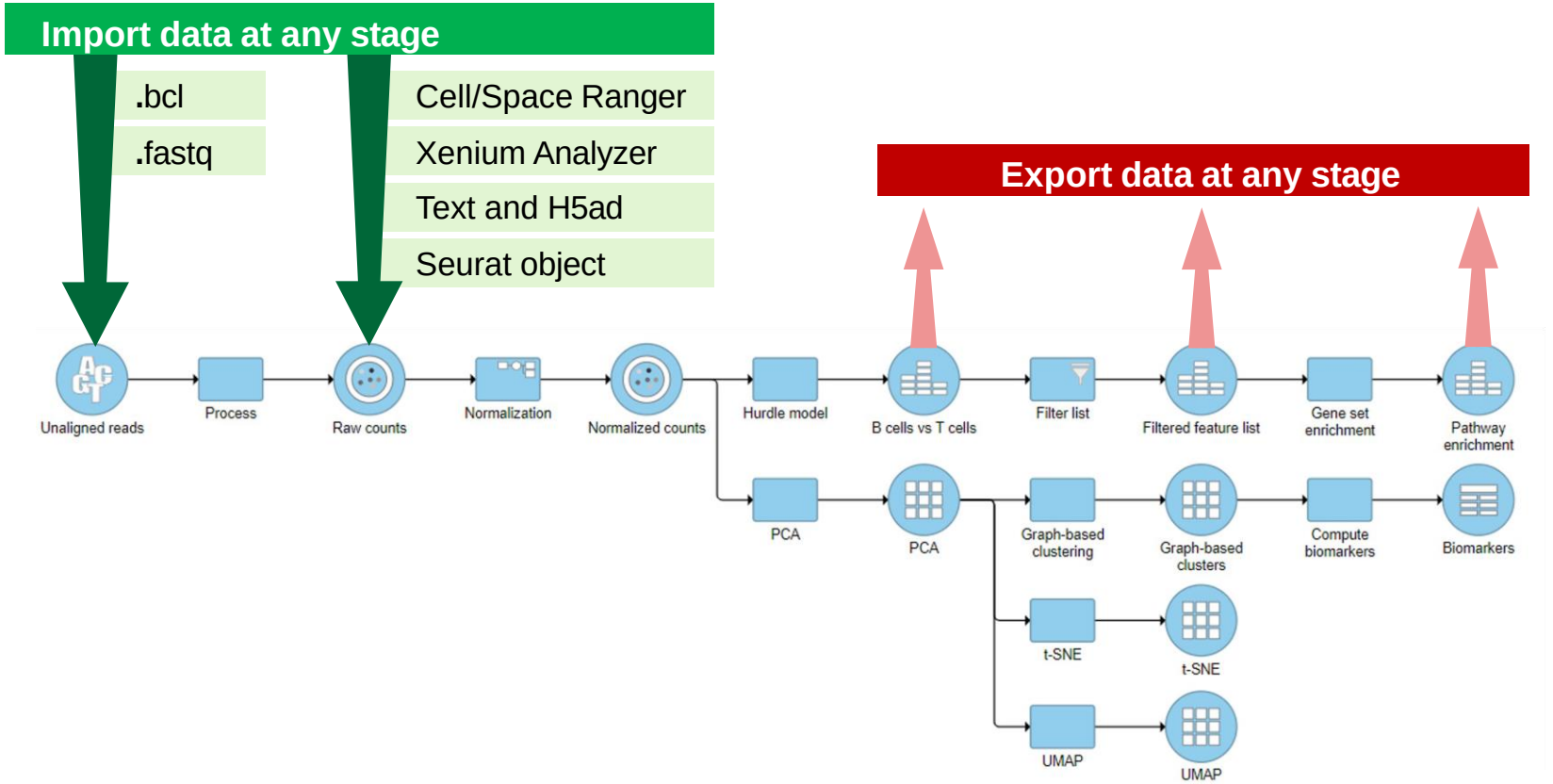
4 Biological Interpretation



Copyright © 2024 Partek, an Illumina company. All rights reserved.
For Research Use Only.
Not for use in diagnostic procedures.

Partek™ an Illumina company

Import and Export Data at Any Stage

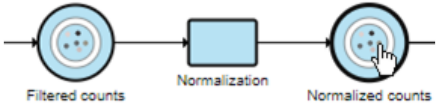


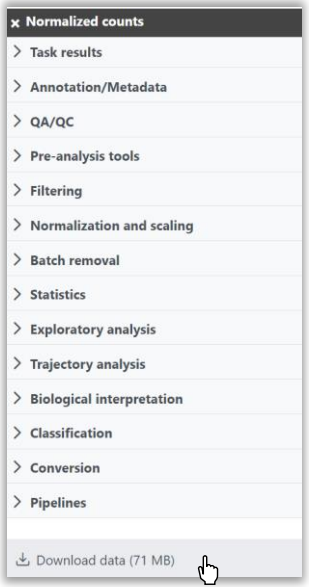
Copyright © 2024 Partek, an Illumina company. All rights reserved.
For Research Use Only.
Not for use in diagnostic procedures.

Partek™ an Illumina company

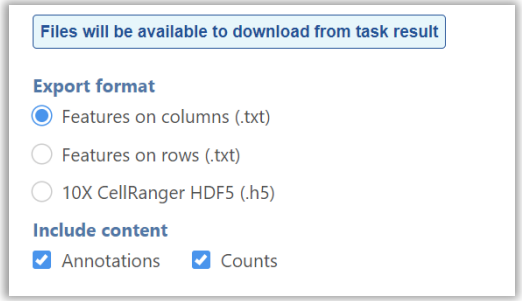
Export Data

Choose Any Data



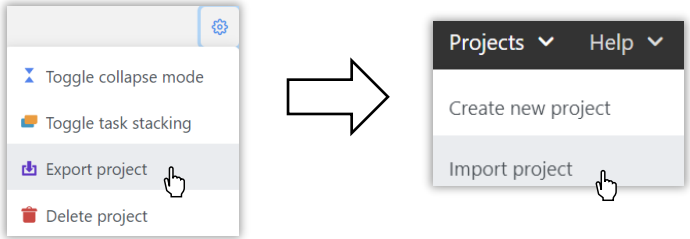


Download in Industry Standard Formats



FASTQ, BAM, TXT, and more

Export and Import Analysis Projects

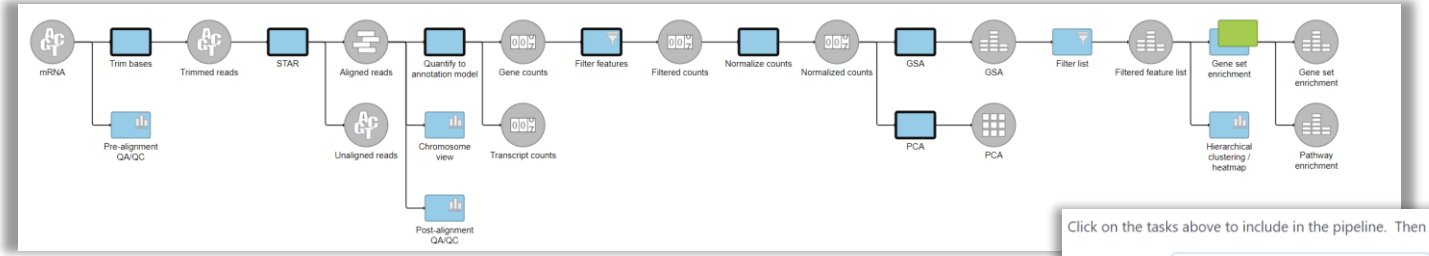


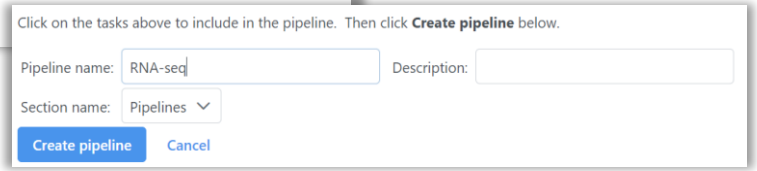
Copyright © 2024 Partek, an Illumina company. All rights reserved.
For Research Use Only.
Not for use in diagnostic procedures.

Partek™ an Illumina company

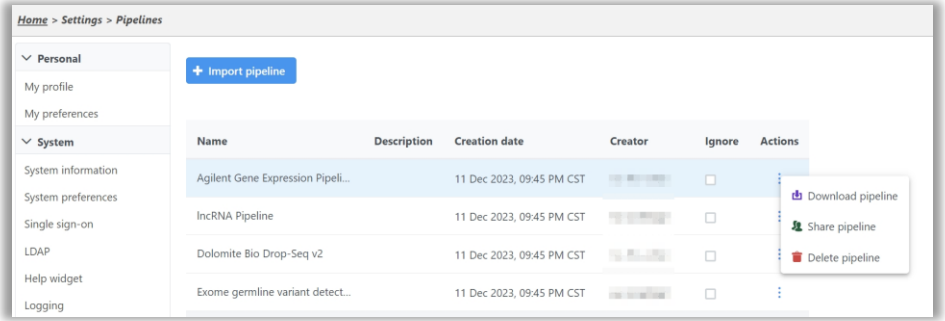
Build, Reuse, and Share Analysis Pipelines

Build Analysis Pipelines





Save, Share, and Manage

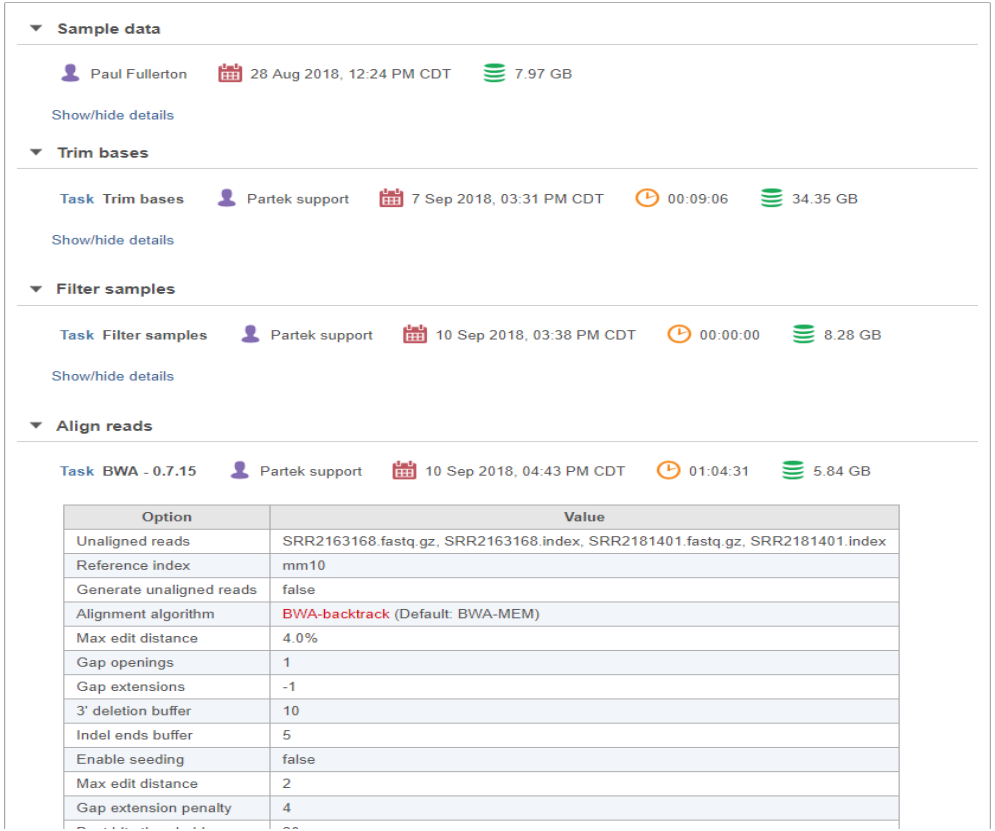


Copyright © 2024 Partek, an Illumina company. All rights reserved.
For Research Use Only.
Not for use in diagnostic procedures.

Partek™ an Illumina company

Summary Report

- Who
- When
- What
- How long
- How much



Sample data

Paul Fullerton | 28 Aug 2018, 12:24 PM CDT | 7.97 GB

Show/hide details

Trim bases

Task **Trim bases** | Partek support | 7 Sep 2018, 03:31 PM CDT | 00:09:06 | 34.35 GB

Show/hide details

Filter samples

Task **Filter samples** | Partek support | 10 Sep 2018, 03:38 PM CDT | 00:00:00 | 8.28 GB

Show/hide details

Align reads

Task **BWA - 0.7.15** | Partek support | 10 Sep 2018, 04:43 PM CDT | 01:04:31 | 5.84 GB

Option	Value
Unaligned reads	SRR2163168.fastq.gz, SRR2163168.index, SRR2181401.fastq.gz, SRR2181401.index
Reference index	mm10
Generate unaligned reads	false
Alignment algorithm	BWA-backtrack (Default: BWA-MEM)
Max edit distance	4.0%
Gap openings	1
Gap extensions	-1
3' deletion buffer	10
Indel ends buffer	5
Enable seeding	false
Max edit distance	2
Gap extension penalty	4

Copyright © 2019 Partek Incorporated. All rights reserved.

Compatible with All Major Genomics Formats and Assays



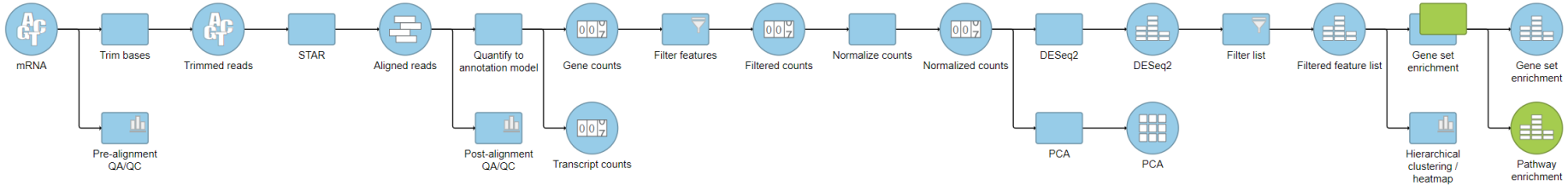
Available Toolkits

- RNA-Seq
- DNA-Seq
- Metagenomics
- Microarray
- ChIP-Seq
- Single Cell

Copyright © 2024 Partek, an Illumina company. All rights reserved.
For Research Use Only.
Not for use in diagnostic procedures.

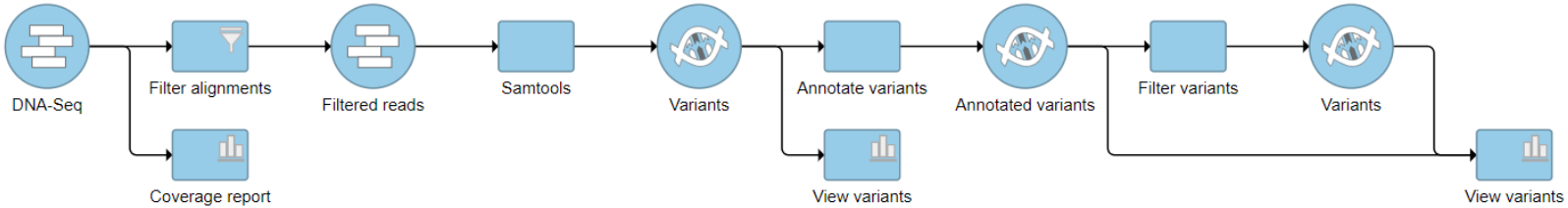
Partek™
an Illumina company

RNA-Seq Demo Pipeline

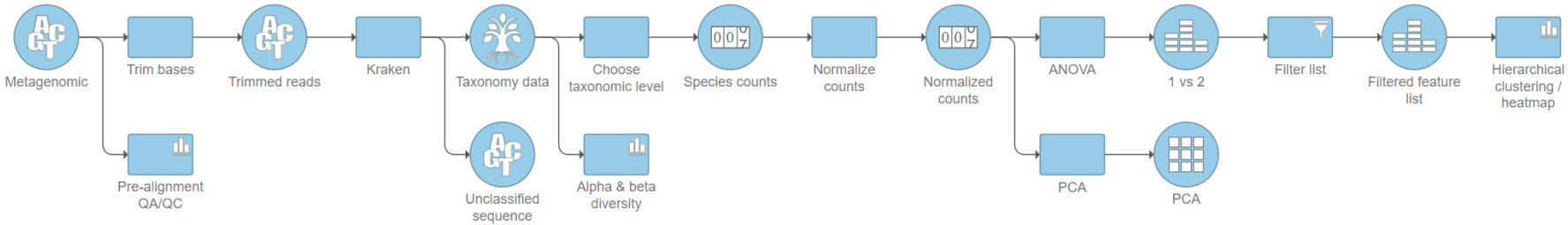


均泰生物科技股份有限公司
Genetech Biotech Co., Ltd.

DNA-Seq Demo Pipeline

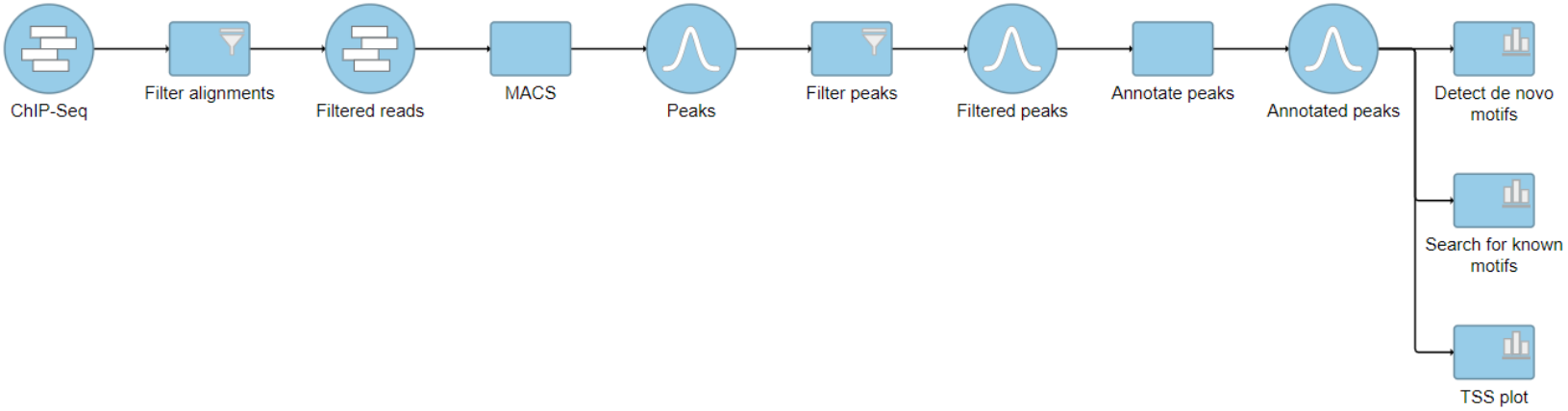


Metagenomics Demo Pipeline

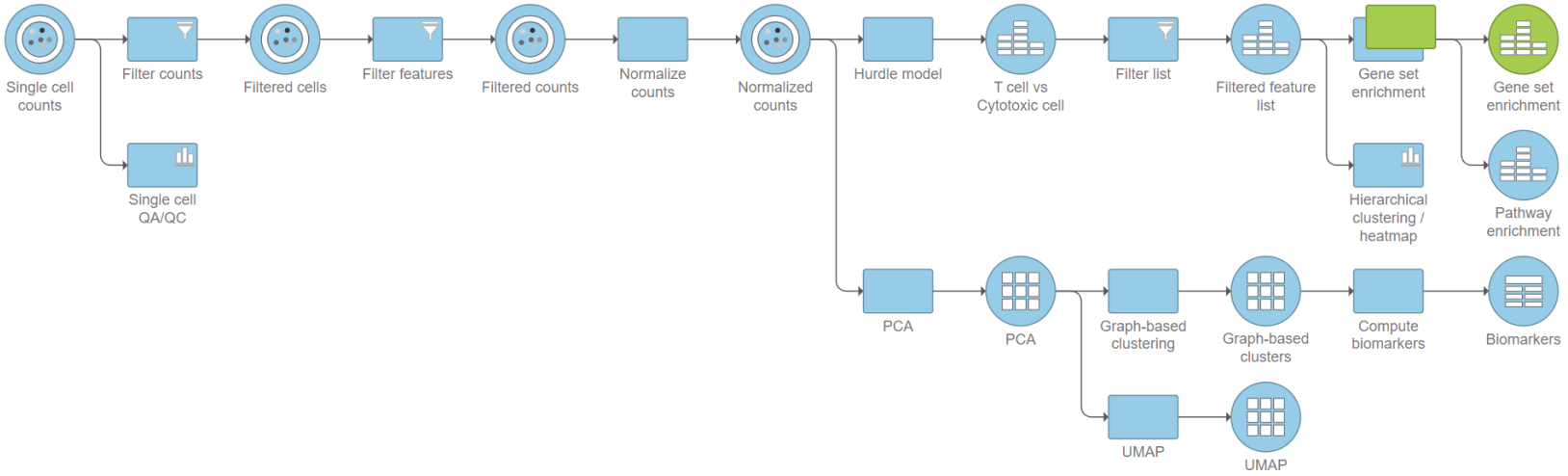


均泰生物科技股份有限公司
Genetech Biotech Co., Ltd.

ChIP-Seq Demo Pipeline



scRNA-Seq Demo Pipeline

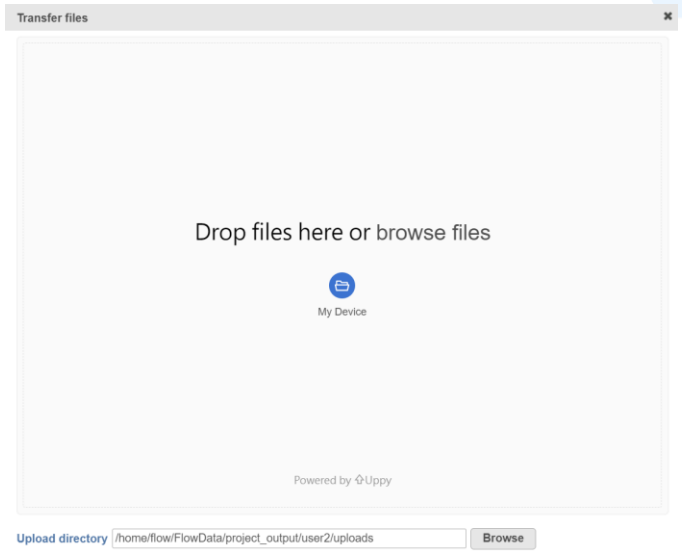


RNA-Seq Analysis Demo

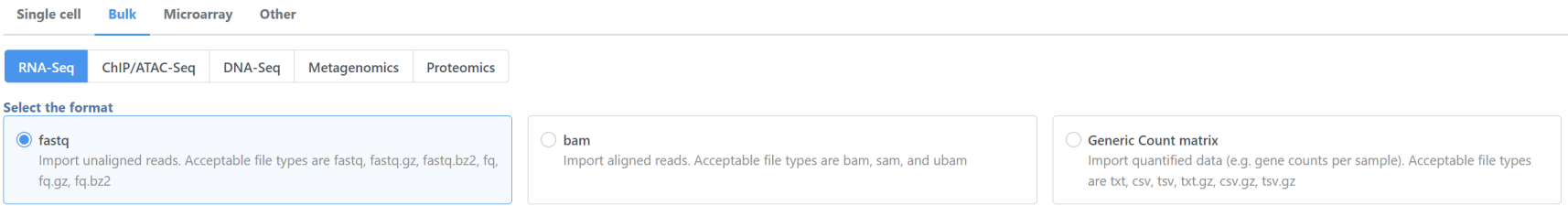
均泰生物科技股份有限公司
Genetech Biotech Co., Ltd.

Transfer Files

- To move files from your local computer to the Partek server, please **Transfer files** first



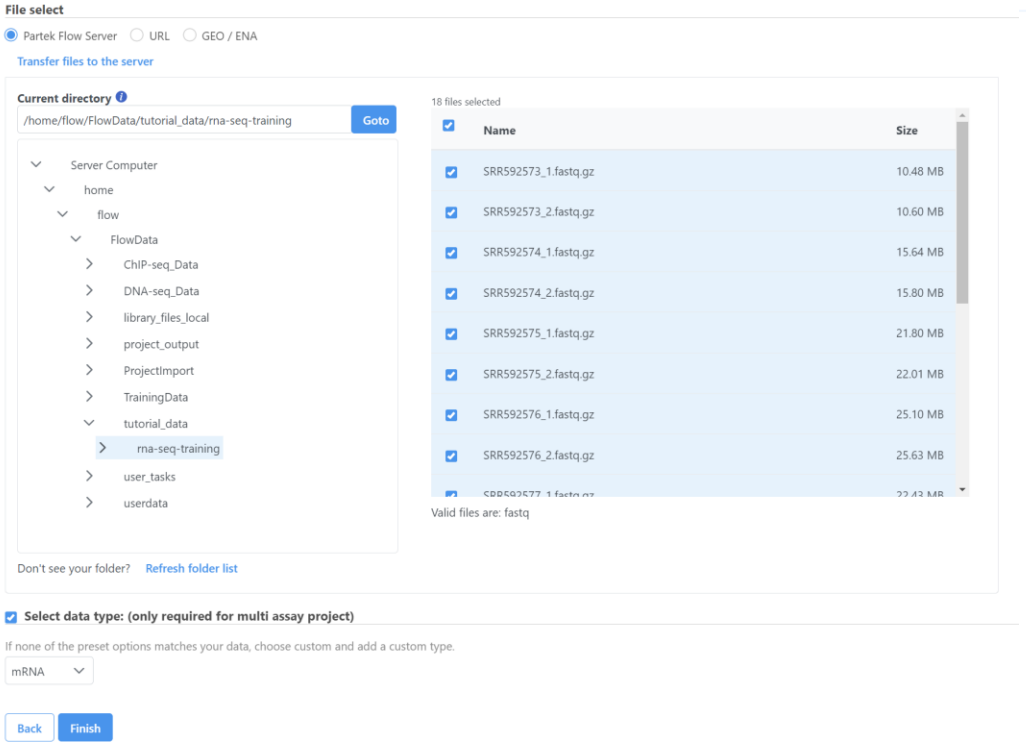
Import Data



- Click **+New project** and enter a project name
- Click **Add data**
- Select the format

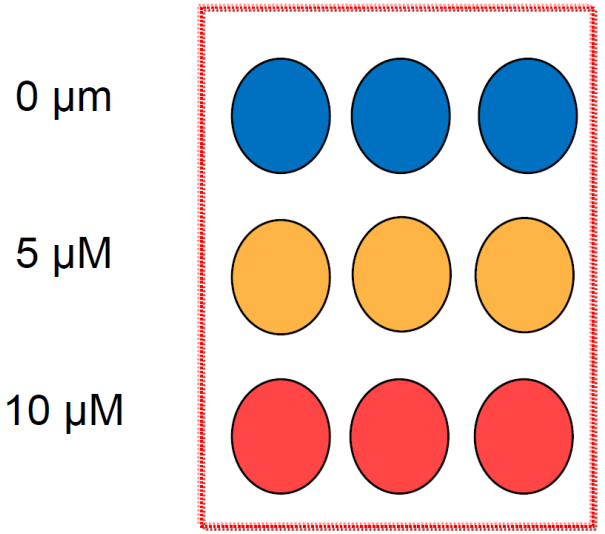
Import Data

- Choose the data needed
  - Partek Flow recognizes paired-end read data if tagged with (_1 or _R1)
- Select data type
  - **mRNA** for RNA-seq data



Experiment Description

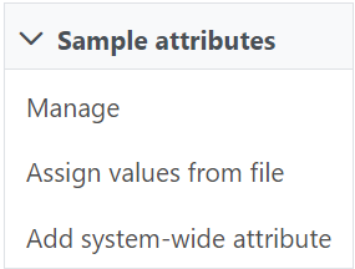
- HT29 colon cancer cells exposed to 5-aza drug with 3 different doses
  - 0 μM (Control)
  - 5 μM
  - 10 μM
- Goal: Identify differentially expressed genes between different groups
- mRNA purified and sequenced using Illumina HiSeq (Paired end reads)
- Xu et al. 2013 BMC Bioinformatics (PMID: 23902433)

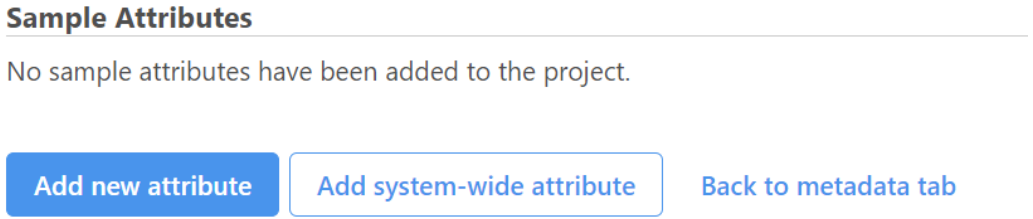


Sample Attribute Assignment

Once the download completes, the sample table will appear in the **Metadata** tab.

- Click **Metadata** tab
- Click **Manage** under Sample attributes
- Click **Add new attribute**

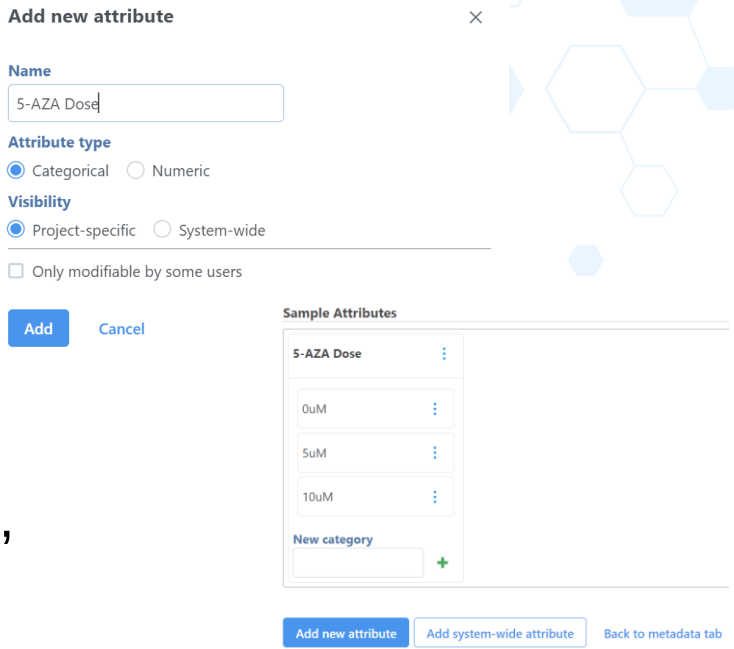




均泰生物科技股份有限公司
Genetech Biotech Co., Ltd.

Sample Attribute Assignment

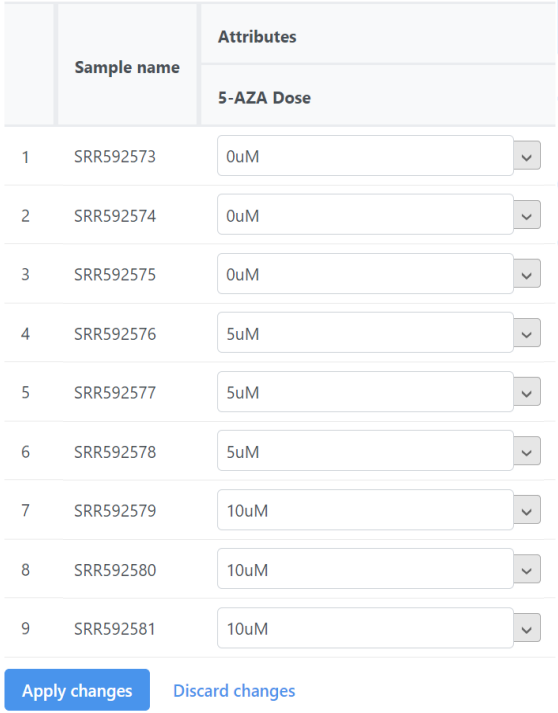
- Name the attribute **5-AZA Dose**
- Click **Categorical** and **Project-specific**
- Click **Add**
- Name the first New category **0uM**
- Click **Add**
- Repeat for two additional categories, **5uM** and **10uM**



均泰生物科技股份有限公司
Genetech Biotech Co., Ltd.

Sample Attribute Assignment

- Click **Back to metadata tab**
- Select **Assign values** under Sample attributes
- Edit the attribute for each sample with the drop-down menu
  - The first three samples should be 0uM, the next three samples should be 5uM, and the final three samples should be 10uM
- Click **Apply changes**

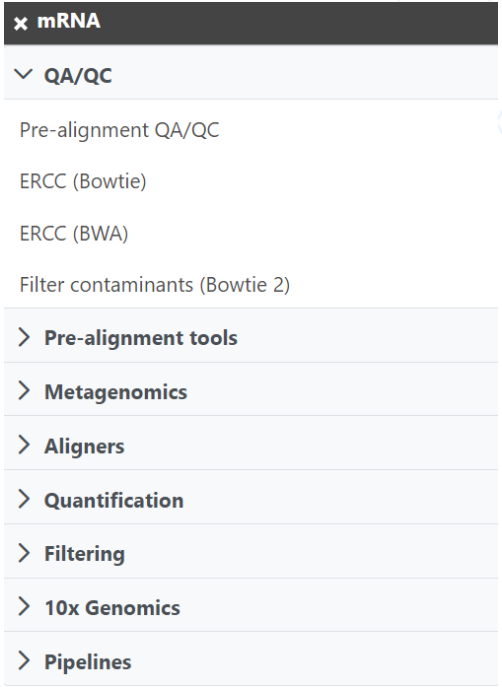


	Sample name	Attributes
		5-AZA Dose
1	SRR592573	0uM
2	SRR592574	0uM
3	SRR592575	0uM
4	SRR592576	5uM
5	SRR592577	5uM
6	SRR592578	5uM
7	SRR592579	10uM
8	SRR592580	10uM
9	SRR592581	10uM

Apply changes Discard changes

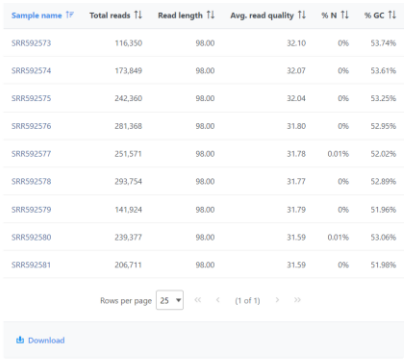
Pre-alignment QA/QC

- Click **Analyses** tab
- Click **mRNA** data node
- Click **Pre-alignment QA/QC** in the QA/QC section of the task menu
- Click **Finish** to run the task with default settings



Pre-alignment QA/QC

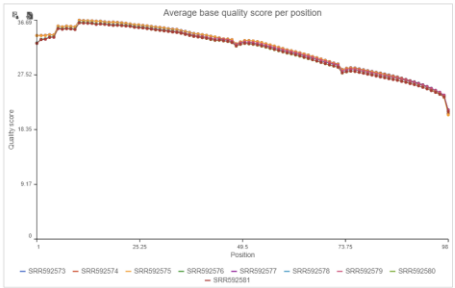
- Double click the **Pre-alignment QA/QC** node



Sample name	Total reads	Read length	Avg. read quality	% N	% GC
SRR592573	116,350	98.00	32.10	0%	53.74%
SRR592574	173,849	98.00	32.07	0%	53.61%
SRR592575	242,360	98.00	32.04	0%	53.25%
SRR592576	281,368	98.00	31.80	0%	52.95%
SRR592577	251,571	98.00	31.78	0.01%	52.02%
SRR592578	293,754	98.00	31.77	0%	52.89%
SRR592579	141,924	98.00	31.79	0%	51.96%
SRR592580	239,377	98.00	31.59	0.01%	53.06%
SRR592581	206,711	98.00	31.59	0%	51.98%

Rows per page 25 (1 of 1)

Download

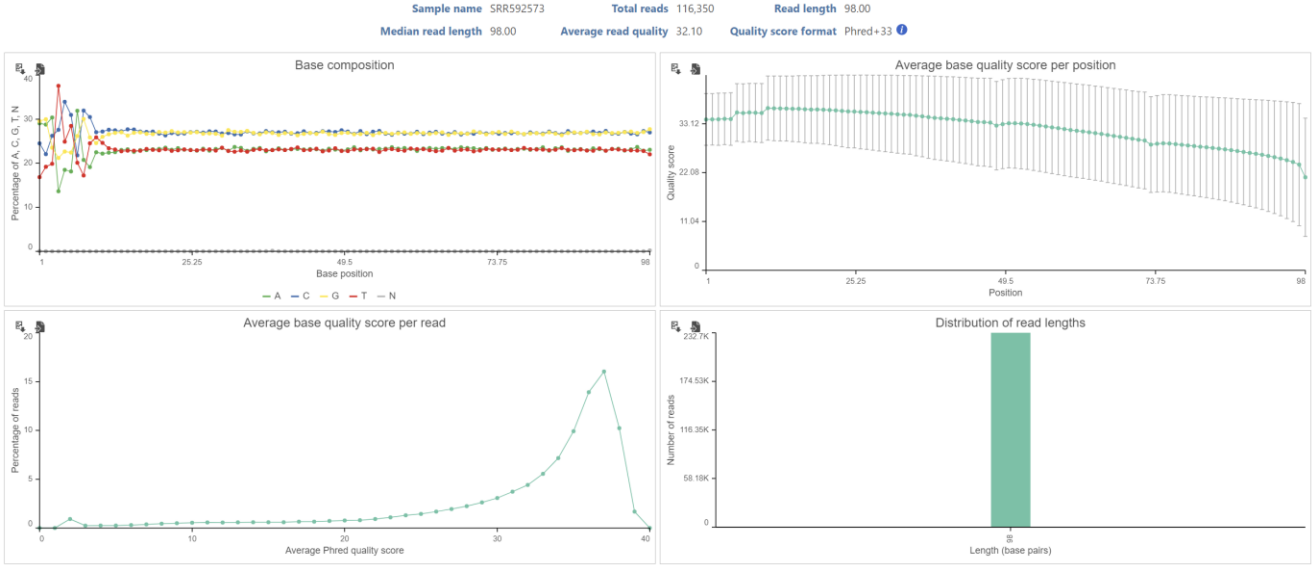




均泰生物科技股份有限公司
Genetech Biotech Co., Ltd.

Pre-alignment QA/QC

- Click a sample name in the table to open a sample-level report

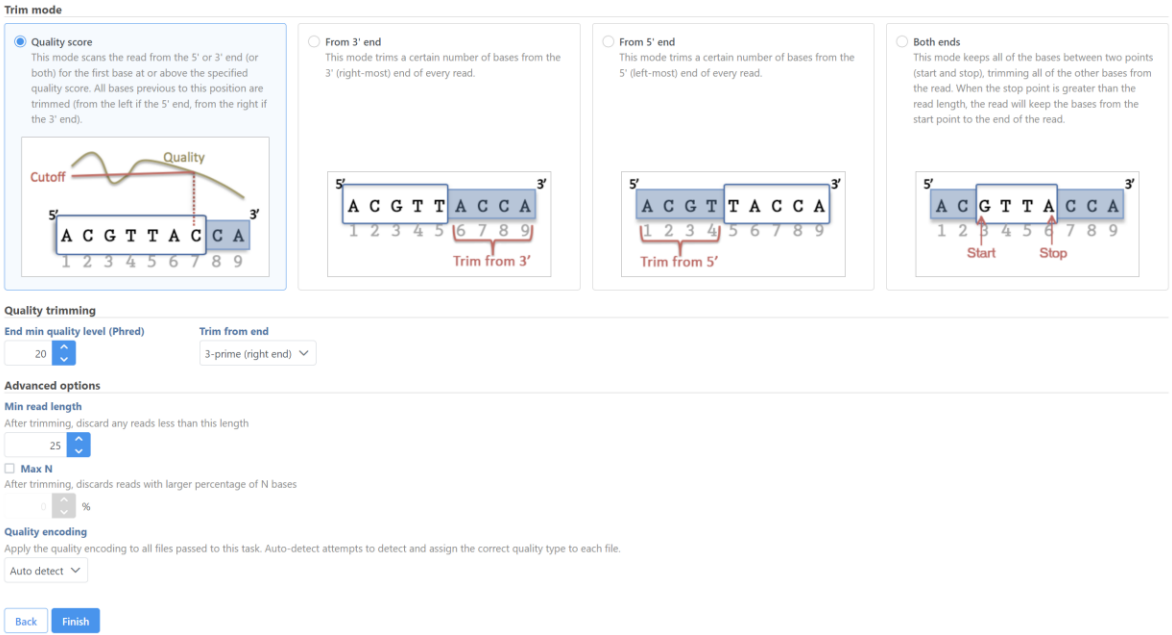


Trim Bases

- Click the project name to return to the Analyses tab
- Click **mRNA** data node
- Click **Pre-alignment tools** in the task menu
- Click **Trim bases**

Trim Bases

- Choose **Quality score** as the trim mode
- Edit End min quality level (Phred) to **20**
- Click **Finish**



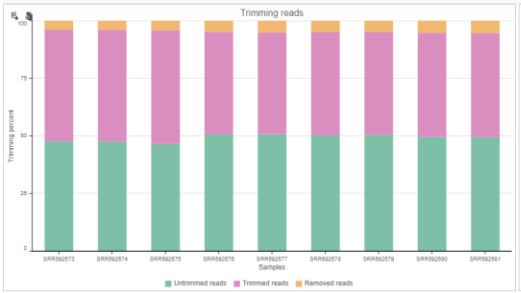
Trim Bases

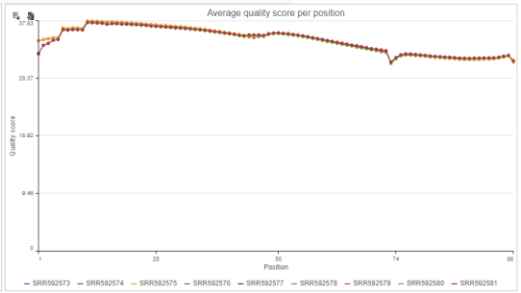
- Double-click **Trimmed reads** data node to open the task report

Sample name TF	Total reads	Reads trimmed	% Reads trimmed	Reads removed	% Reads removed	Average bases trimmed	Pre-trim quality	Post-trim quality
SRR592573	116,350	56,346	48.43%	4,623	3.97%	20.33	32.19	34.51
SRR592574	173,849	84,226	48.45%	7,003	4.03%	20.63	32.09	34.48
SRR592575	242,360	119,611	49.35%	10,109	4.17%	20.56	32.20	34.57
SRR592576	281,368	125,164	44.48%	13,722	4.88%	22.71	31.73	34.41
SRR592577	251,571	111,013	44.13%	12,924	5.14%	22.78	31.67	34.43
SRR592578	293,754	132,092	44.97%	14,518	4.94%	22.61	31.73	34.41
SRR592579	141,924	63,291	44.59%	7,011	4.94%	22.90	31.71	34.42
SRR592580	239,377	108,306	45.24%	12,605	5.27%	23.55	31.49	34.35
SRR592581	206,711	93,354	45.16%	11,179	5.41%	24.12	31.50	34.44

Rows per page 25 (1 of 1)

Download

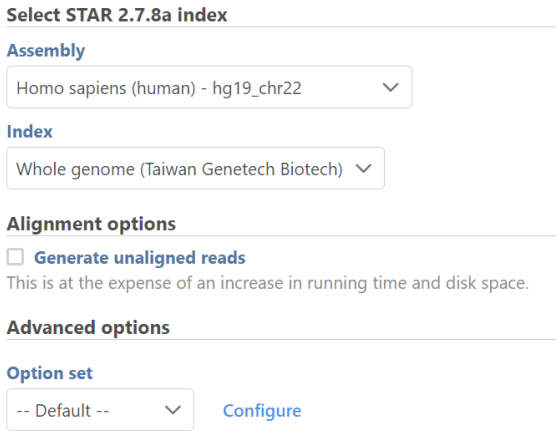




均泰生物科技股份有限公司
Genetech Biotech Co., Ltd.

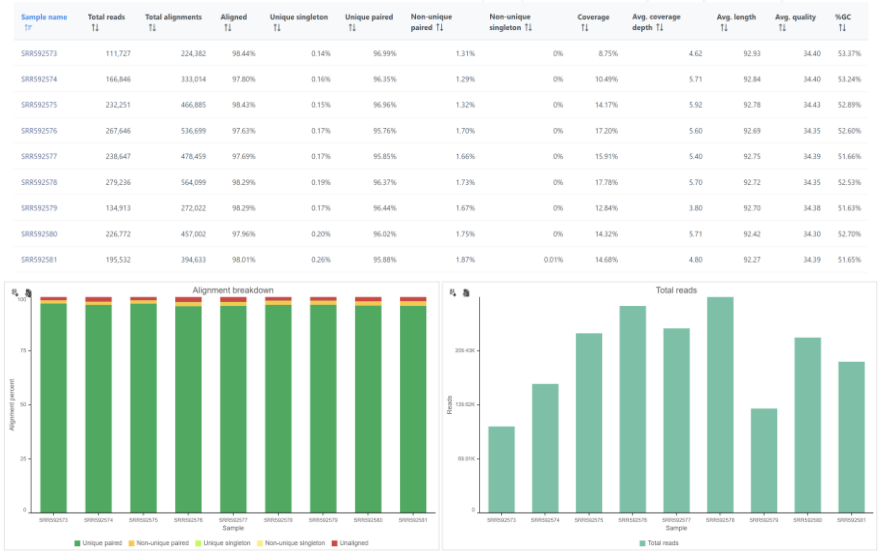
Alignment

- Click the project name to return to the Analyses tab
- Click **Trimmed reads**
- Click **STAR** from Aligners
- Choose **Homo sapiens - hg19_chr22** for Assembly and **Whole genome** for Index
- Click **Finish**



Post-alignment QA/QC

- Click **Aligned reads** data node
- Click **QA/QC** in the task menu
- Click **Post-alignment QA/QC**
- Double-click the **Post-alignment QA/QC** task node to view the task report

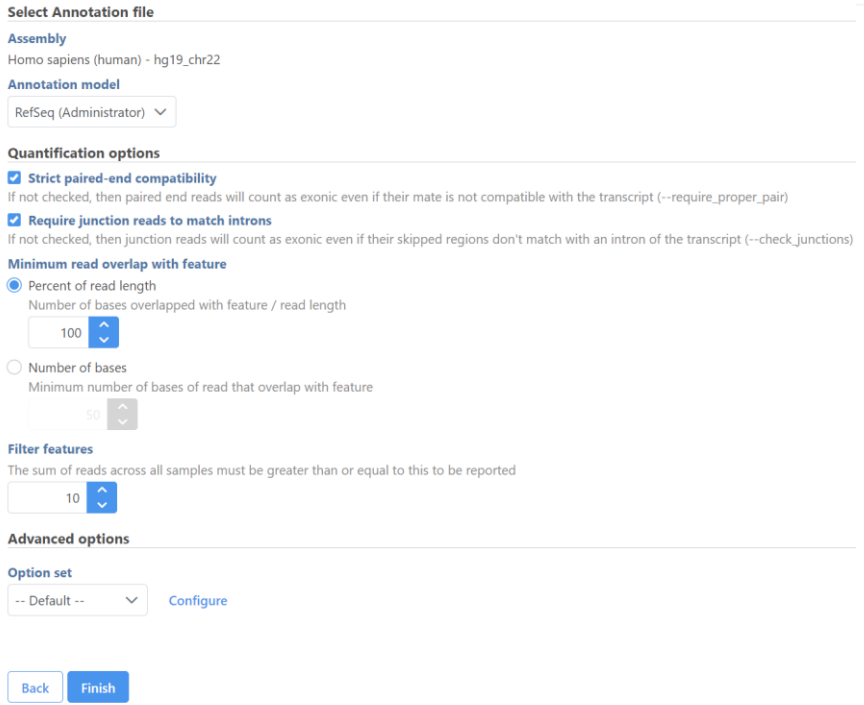


Quantification

- Click the project name to return to the Analyses tab
- Click **Aligned reads** data node
- Click **Quantification** in the task menu
- Click **Quantify to an annotation model (Partek E/M)**

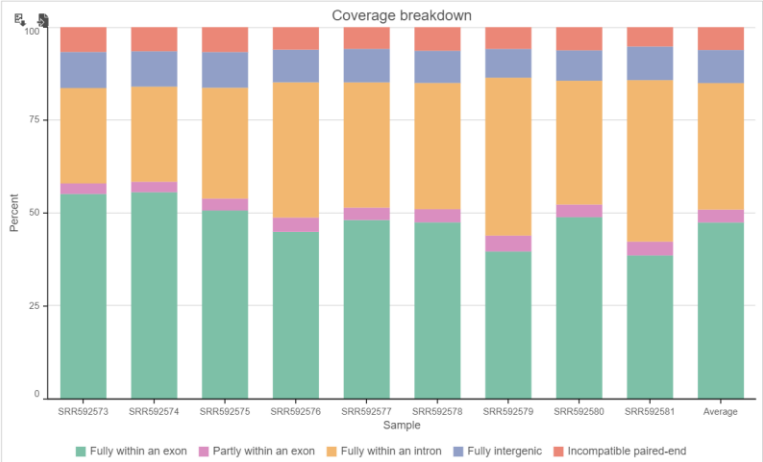
Quantification

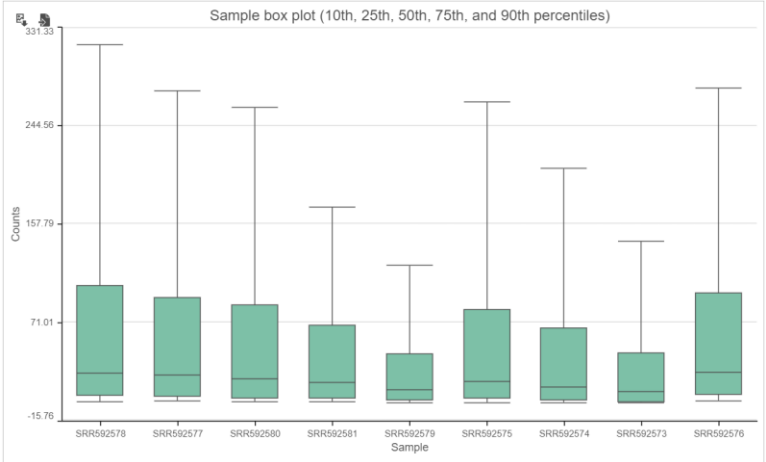
- Choose the **RefSeq** for Annotation model
  - You may need to download it first, via Library File Management
- Click **Finish**



Quantification

- Double-click **Gene counts** data node to view the task report





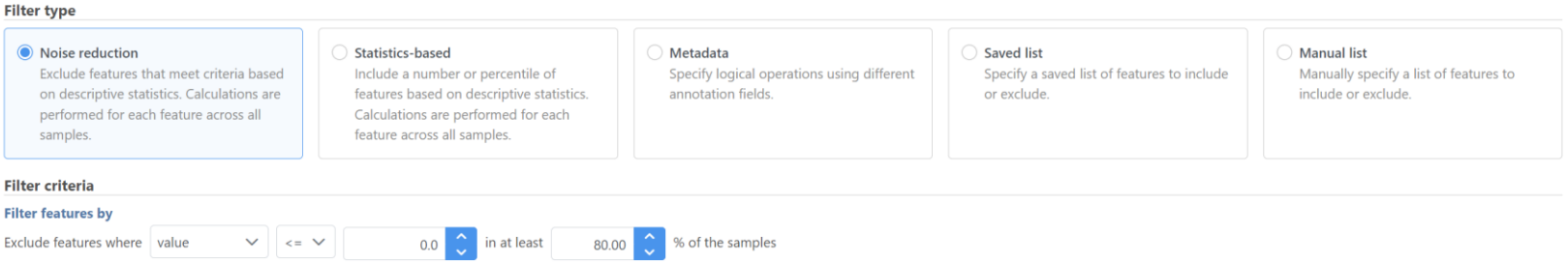
均泰生物科技股份有限公司
Genetech Biotech Co., Ltd.

Filter Features

- Click the project name to return to the Analyses tab
- Click **Gene counts** node
- Click **Filtering** in the task menu
- Click **Filter features**

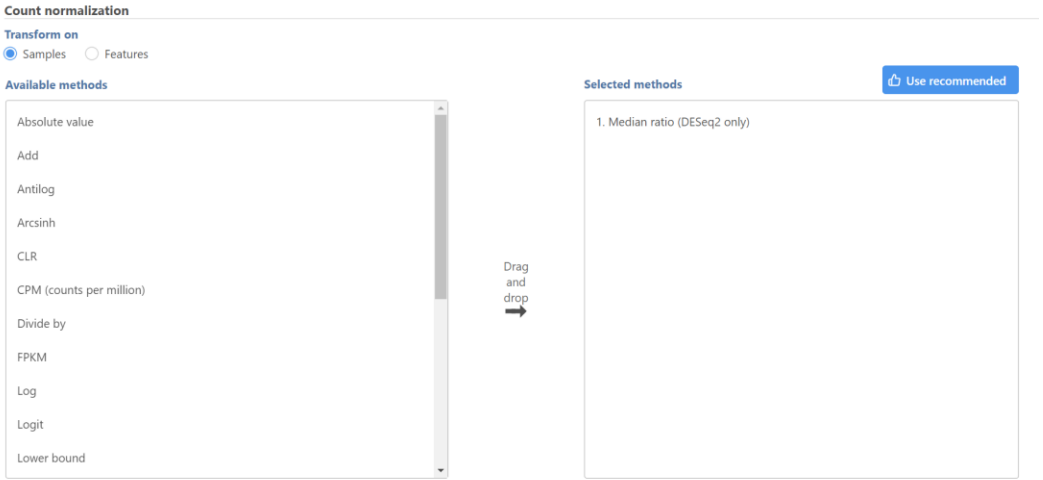
Filter Features

- Click the **Noise reduction filter** checkbox
- Set the filter to **Exclude features where value <= 0 in 80% of cells** using the drop-down menus and text boxes
- Click **Finish** to apply the filter



Normalization

- Click **Filtered counts** data node
- Click **Normalization and scaling** in the task menu
- Click **Normalization**
- Select **Recommended** if no preferred methods
- Click **Finish**



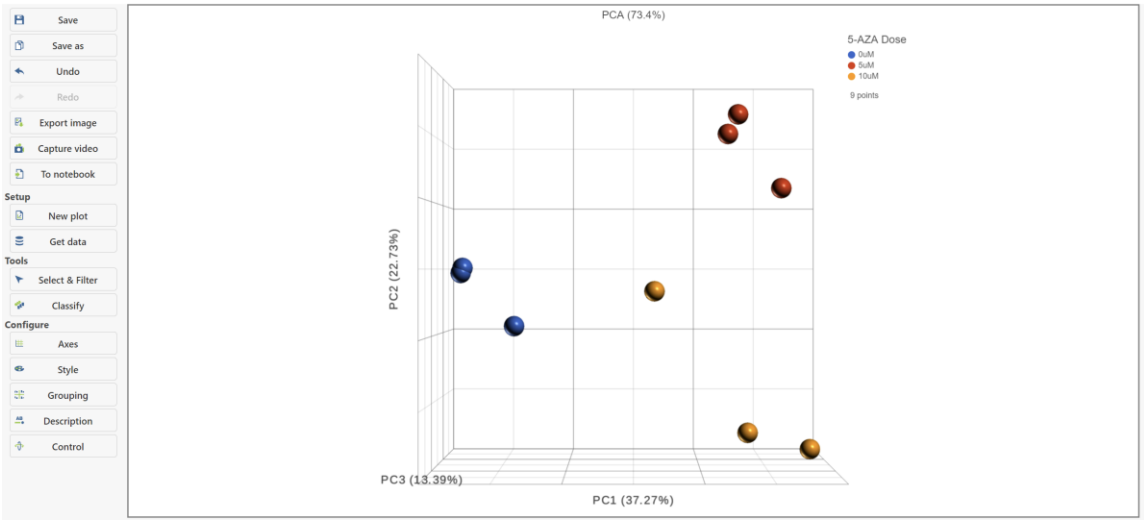
Principal Components Analysis (PCA)

The PCA scatter plot allows to visualize similarities and differences between the samples in a data set.

- Click **Normalized counts** data node
- Click **Exploratory analysis** in the task menu
- Click **PCA**
- Click **Finish** to run with the default options

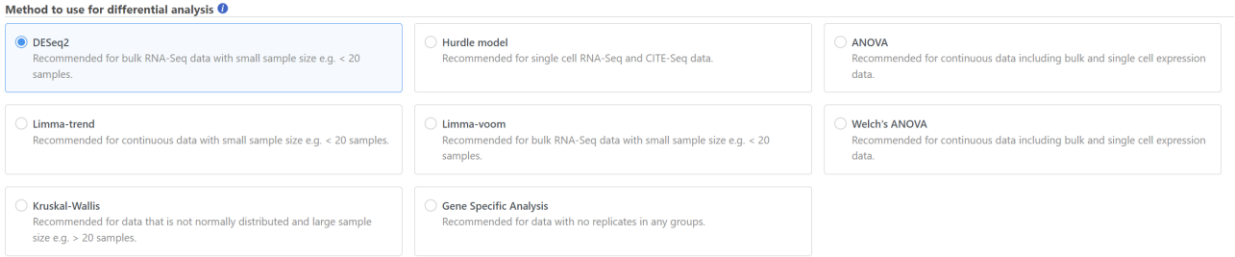
PCA

- Double click **PCA** data node to open the PCA scatter plot
- Click **Style** under Configure
- Set the **Color by** drop-down to **5-AZA Dose**



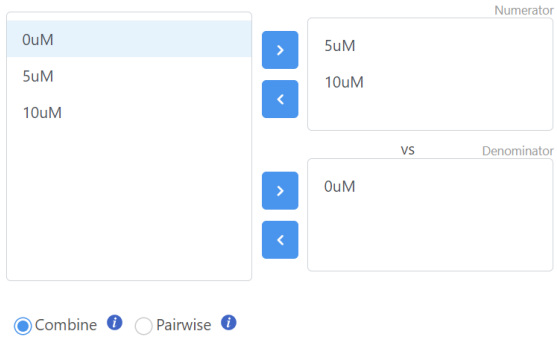
Differential Analysis

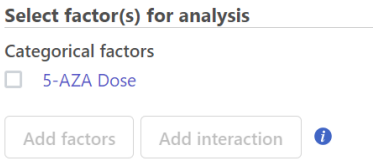
- Click **Normalized counts** node
- Click **Statistics** in the task menu
- Click **Differential analysis** in the task menu
- Select a differential analysis method
- In this tutorial we are going to use **DESeq2**

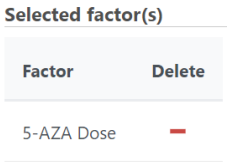


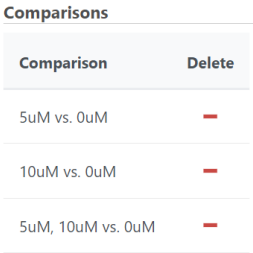
Differential Analysis

- Select the attribute **5-AZA Dose**
- Click **Next**
- Select comparison pairs and click **Add comparison**
- Click **Finish**





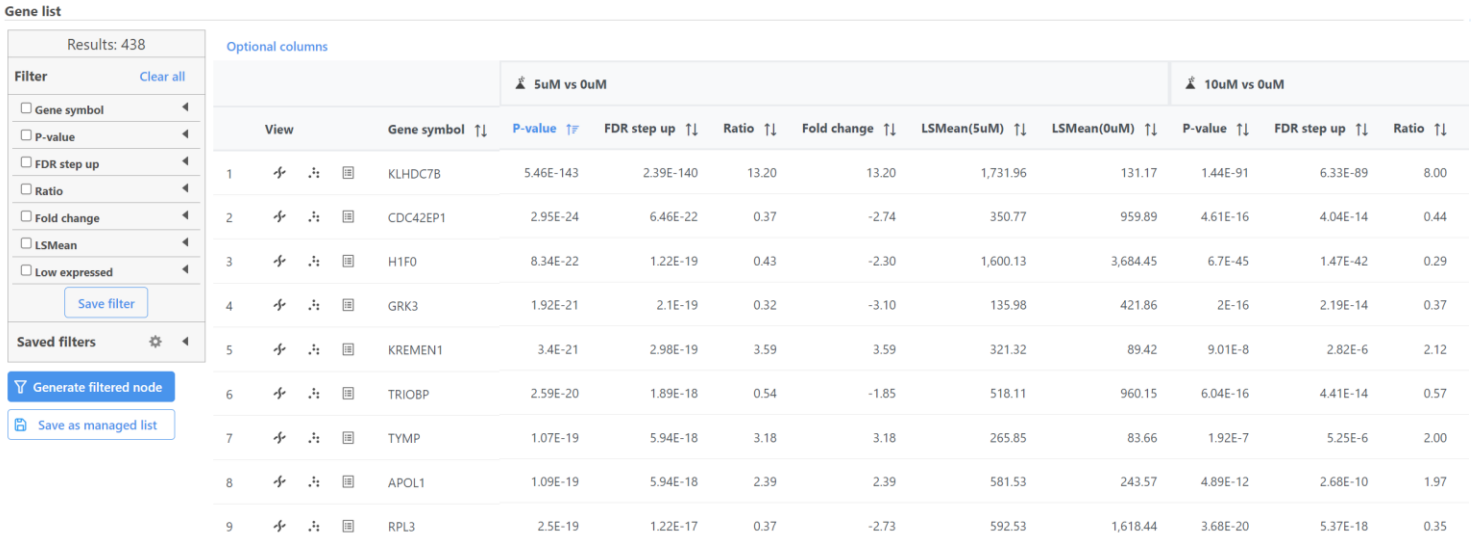




均泰生物科技股份有限公司
Genetech Biotech Co., Ltd.

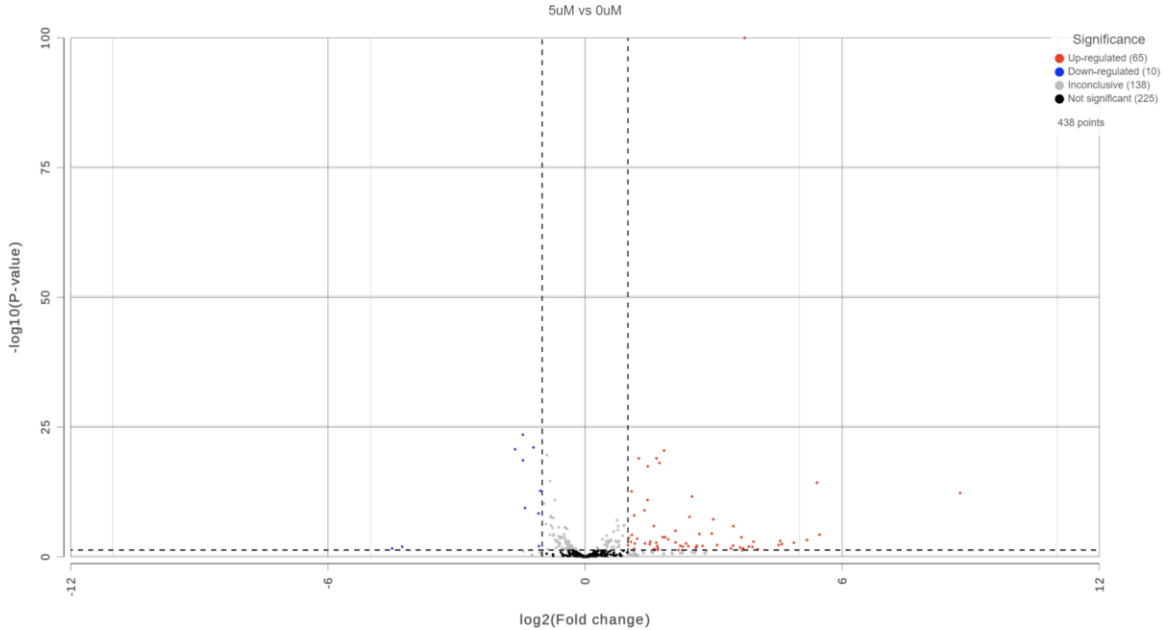
Differential Analysis

- Double click **DESeq2** data node to open the task report



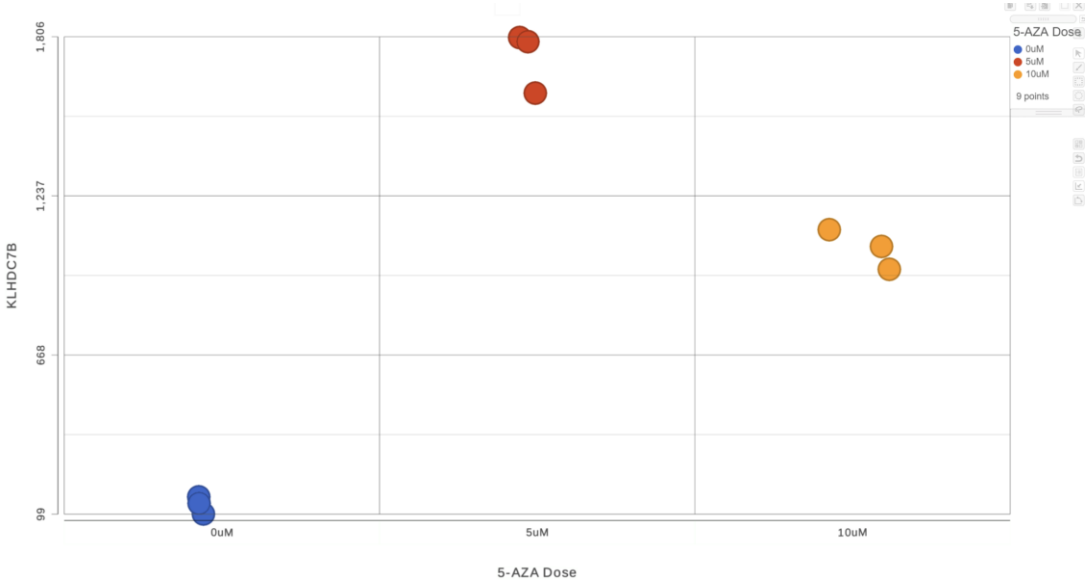
Volcano Plot for All Genes

- Click  next to the 5uM vs. 0uM comparison



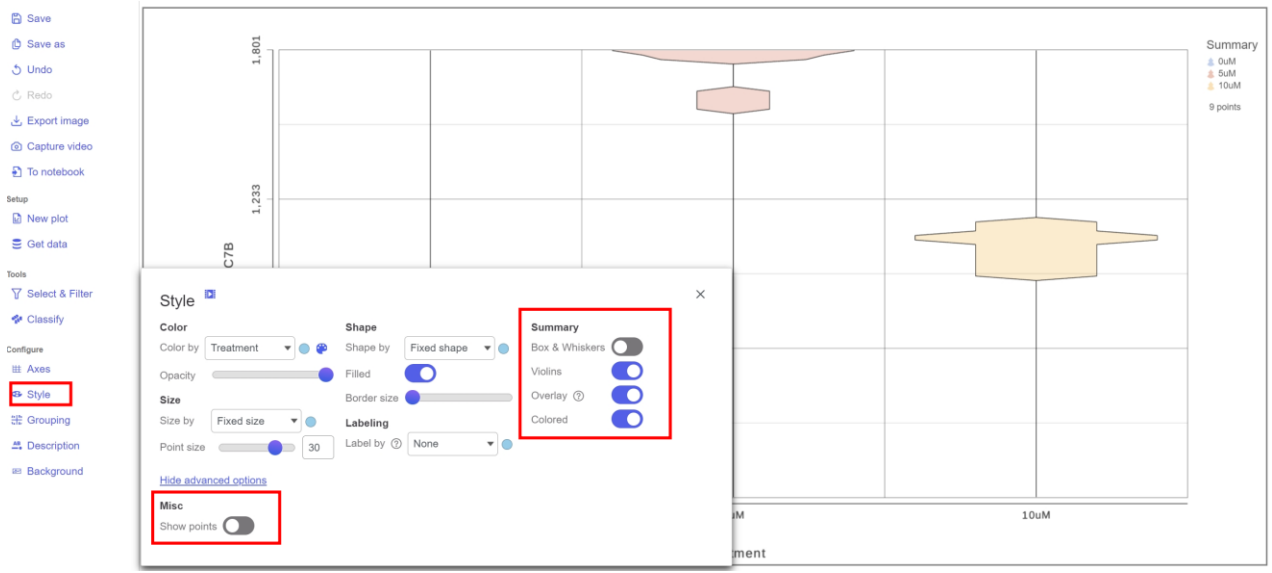
Dot Plot for One Gene

- Select  next to a gene symbol to open a dot plot for the gene



Violin Plot for One Gene

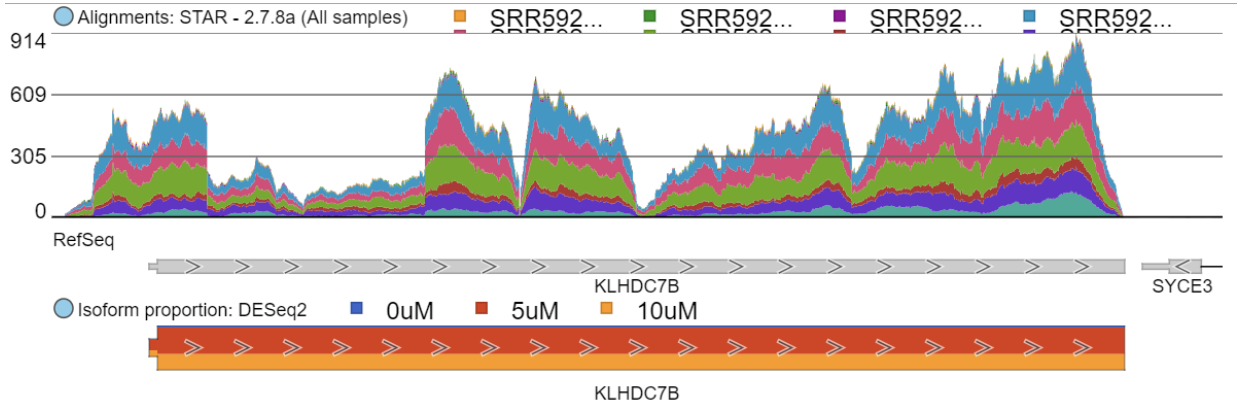
- The dot plot can be changed into a violin plot by some settings



均泰生物科技股份有限公司
Genetech Biotech Co., Ltd.

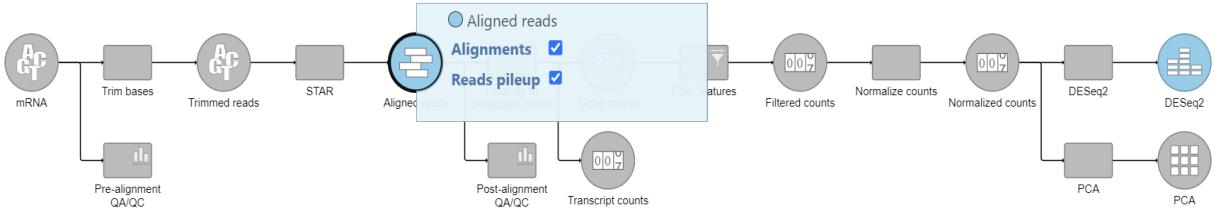
Browse Gene in Chromosome View

- Click next to a gene symbol to open Chromosome view



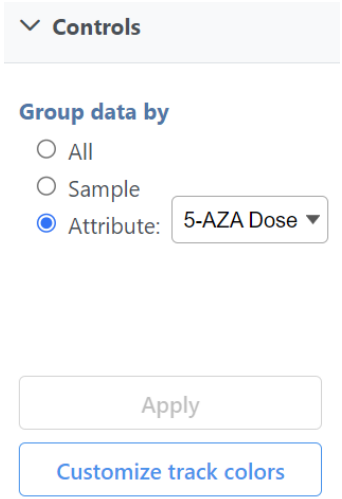
Browse Gene in Chromosome View

- Click **Select tracks**
- Choose **Aligned Reads** data node
- Click **Reads pileup** under Aligned reads
- Click **Display selection** to make the change



Browse Gene in Chromosome View

- Group data by **Attribute** and choose **5-AZA**
- Click **Apply**

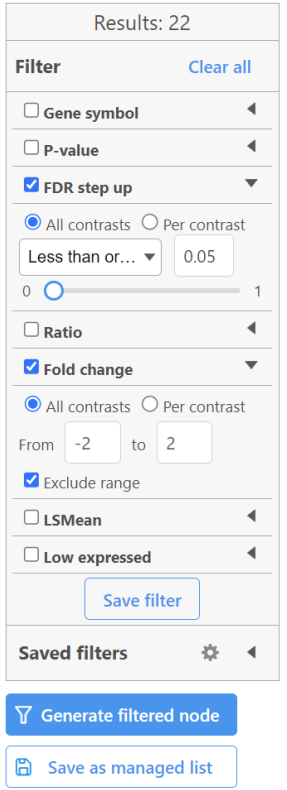




均泰生物科技股份有限公司
Genetech Biotech Co., Ltd.

Identify DEG

- Switch to the browser tab showing the DESeq2 report
- Click **FDR step up**
- Set the cutoff value to **0.05**
- Click **Fold-change**
- Set to **From -2 to 2** with **Exclude range** selected
- Click **Generate filtered node** to create a data node with only the genes that pass the filter

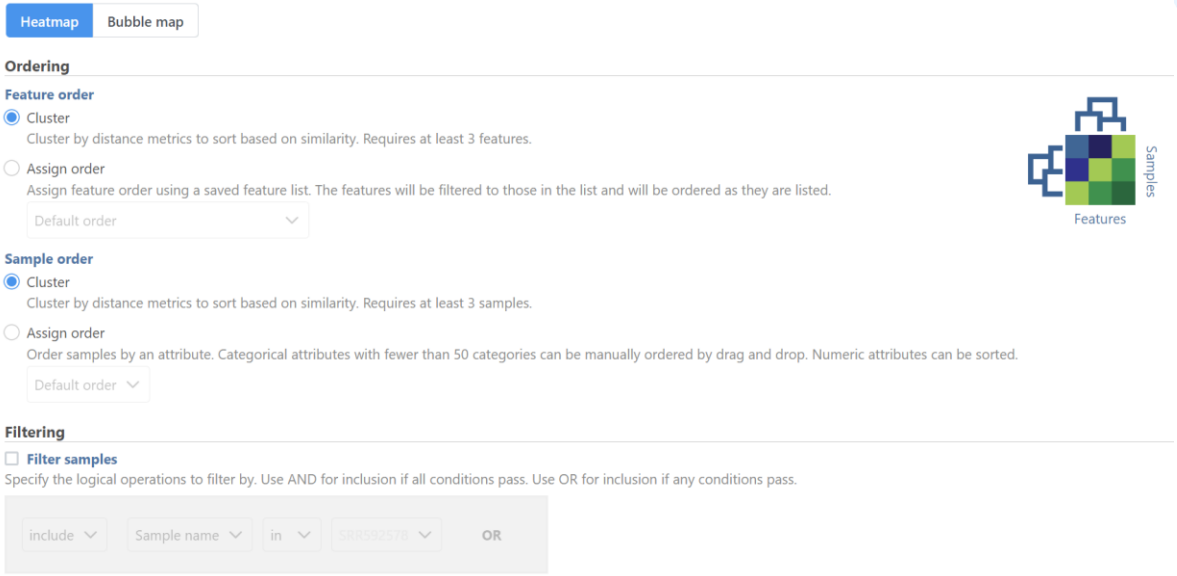


Hierarchical Clustering

- Click **Feature list** data node
- Click **Exploratory analysis** in the task menu
- Click **Hierarchical clustering / heatmap**

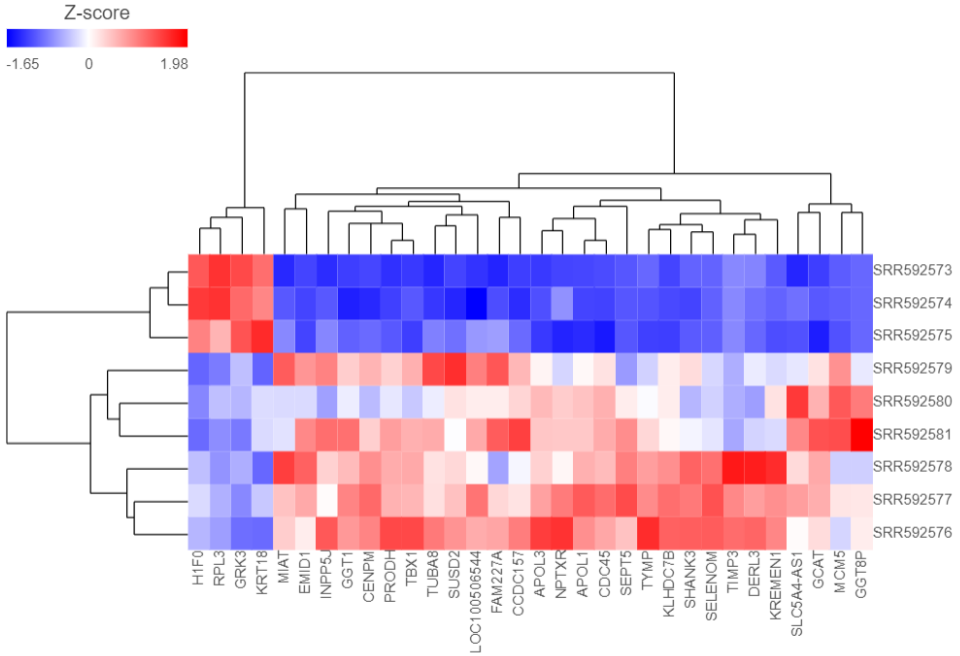
Hierarchical Clustering & Heatmap

- Click **Finish** to run with default settings



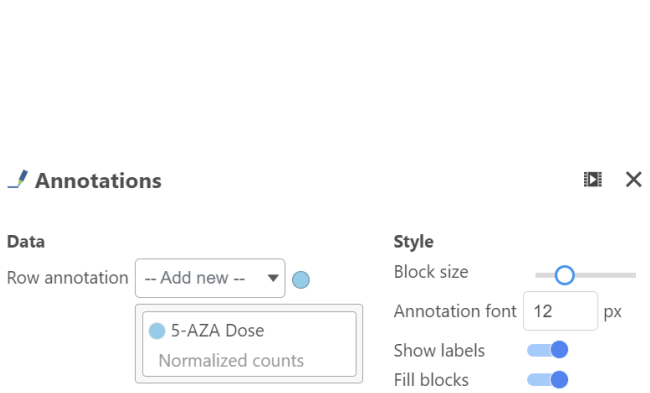
Hierarchical Clustering & Heatmap

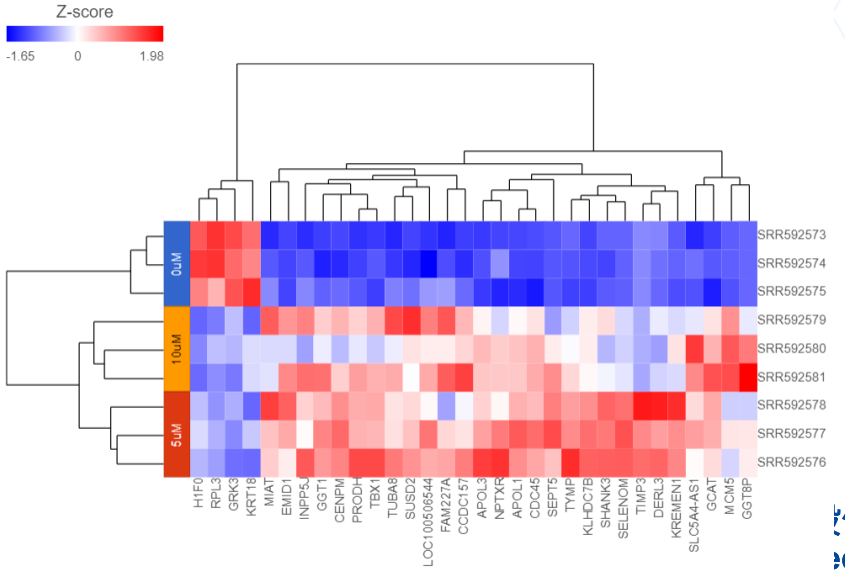
- Double-click the **Hierarchical clustering / heatmap** task node



Hierarchical Clustering & Heatmap

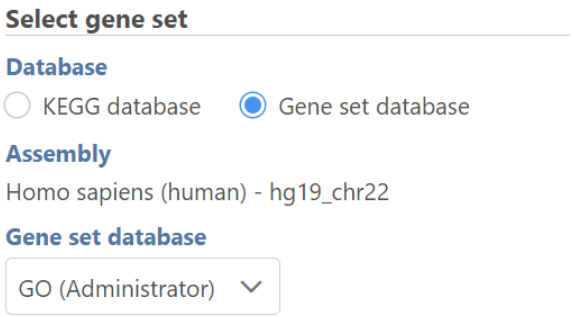
- Expand the **Annotations** card
- Set the Row annotation to **5-AZA Dose**





Enrichment Analysis

- Click the project name to return to the Analyses tab
- Click **Filtered feature list** data node
- Click **Biological interpretation** in the task menu
- Click **Gene set enrichment**
- Select **Gene set database** to perform GO enrichment analysis



Enrichment Analysis

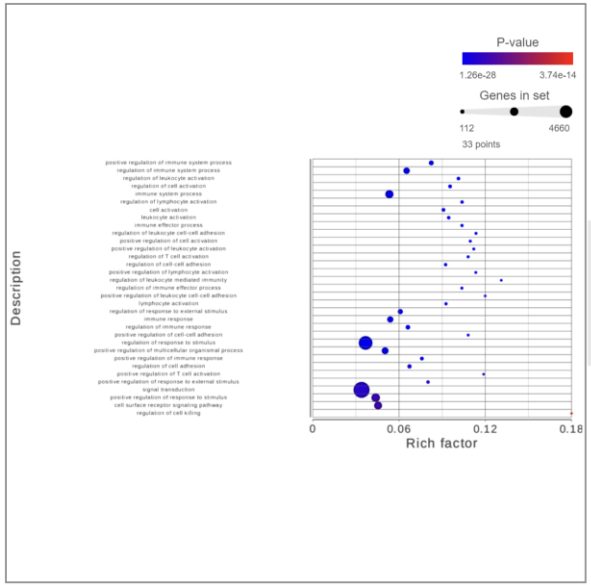
- Double-click **Gene set enrichment** task node to open the task report

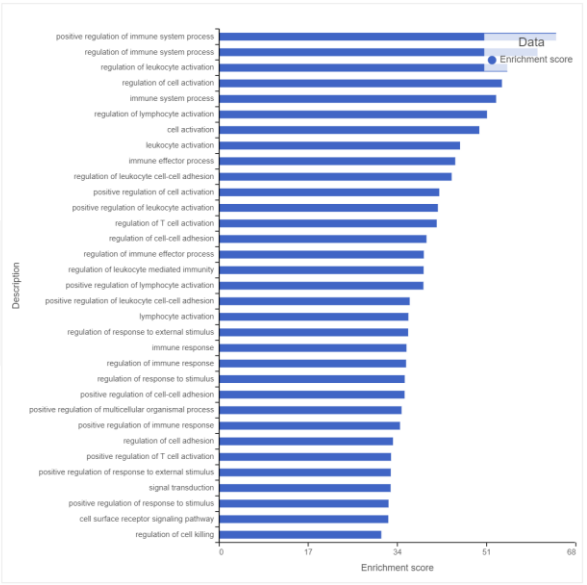
Gene set	Description	Type	Enrichment score	P-value	FDR step up	Rich factor	Genes in set	Genes in list	Genes not in list	Genes in list, not in set	Genes not in list, not in set
GO:1901605	alpha-amino acid metabolic process	biological process	9.30	9.18E-5	0.43	1.00	3	3	0	13	317
GO:0005198	structural molecule activity	molecular function	7.36	6.36E-4	0.68	0.40	10	4	6	12	311
GO:0006520	cellular amino acid metabolic process	biological process	6.39	1.68E-3	0.68	0.50	6	3	3	13	314
GO:0042219	cellular modified amino acid catabolic process	biological process	6.13	2.17E-3	0.68	1.00	2	2	0	14	317
GO:0043648	dicarboxylic acid metabolic process	biological process	6.13	2.17E-3	0.68	1.00	2	2	0	14	317
GO:0044843	cell cycle G1/S phase transition	biological process	6.13	2.17E-3	0.68	1.00	2	2	0	14	317
GO:0022616	DNA strand elongation	biological process	6.13	2.17E-3	0.68	1.00	2	2	0	14	317
GO:1901606	alpha-amino acid catabolic process	biological process	6.13	2.17E-3	0.68	1.00	2	2	0	14	317
GO:0000082	G1/S transition of mitotic cell cycle	biological process	6.13	2.17E-3	0.68	1.00	2	2	0	14	317
GO:0003688	DNA replication origin binding	molecular function	6.13	2.17E-3	0.68	1.00	2	2	0	14	317
GO:0006270	DNA replication initiation	biological process	6.13	2.17E-3	0.68	1.00	2	2	0	14	317

Plots of Enrichment Analysis

Data Viewer plots available when table contains 100 or fewer rows. Operators such as >= or <= can be used in the numeric columns' filters.

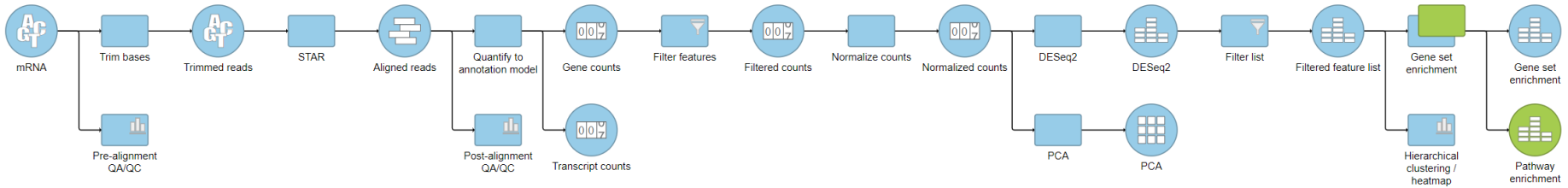
Download table Download gene sets View plots in Data Viewer





均泰生物科技股份有限公司
Genetech Biotech Co., Ltd.

Pipeline Overview



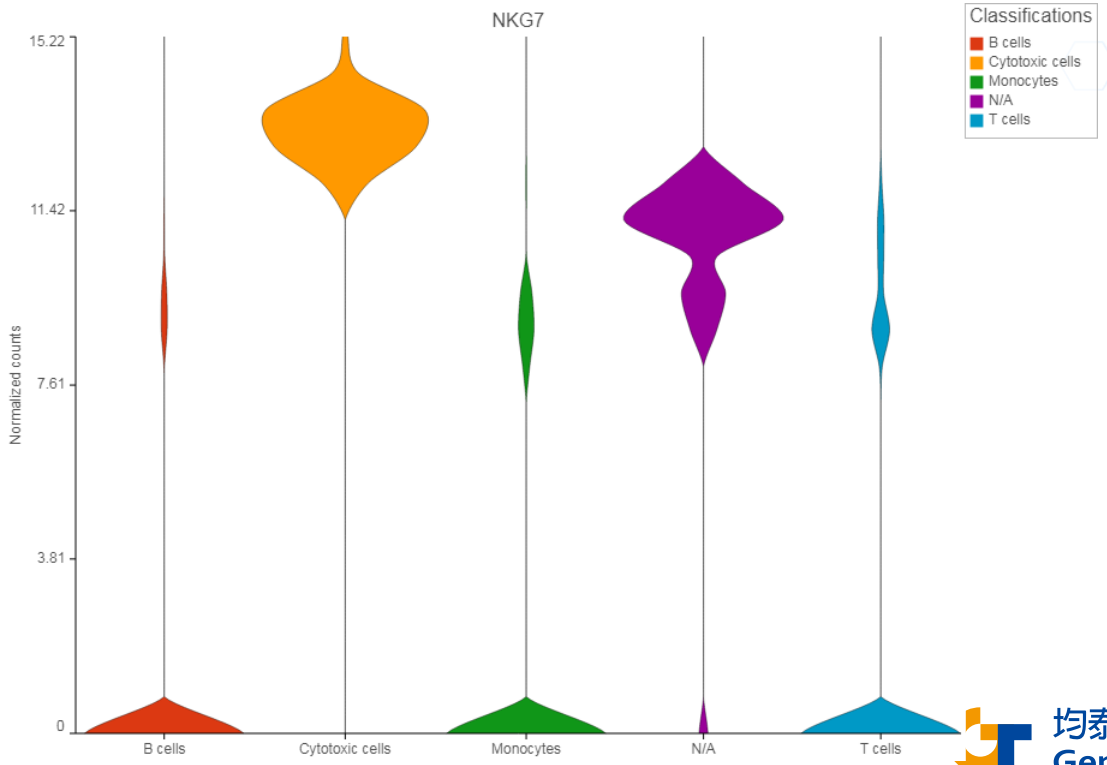
均泰生物科技股份有限公司
Genetech Biotech Co., Ltd.

Plot Interpretation

均泰生物科技股份有限公司
Genetech Biotech Co., Ltd.

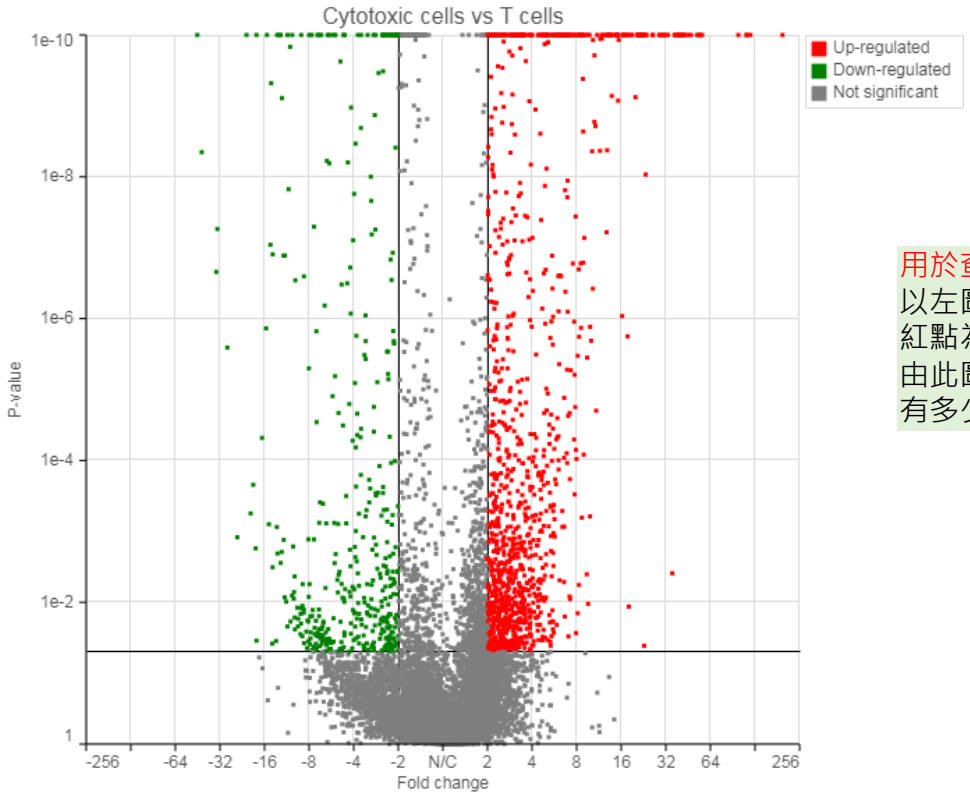
Violin Plot

X軸為不同的細胞類別，Y軸為Normalized後的 Read count數；客戶可自行將細胞分類，
並透過Feature Plot了解特定基因在不同類別中的RNA表現量



均泰生物科技股份有限公司
Genetech Biotech Co., Ltd.

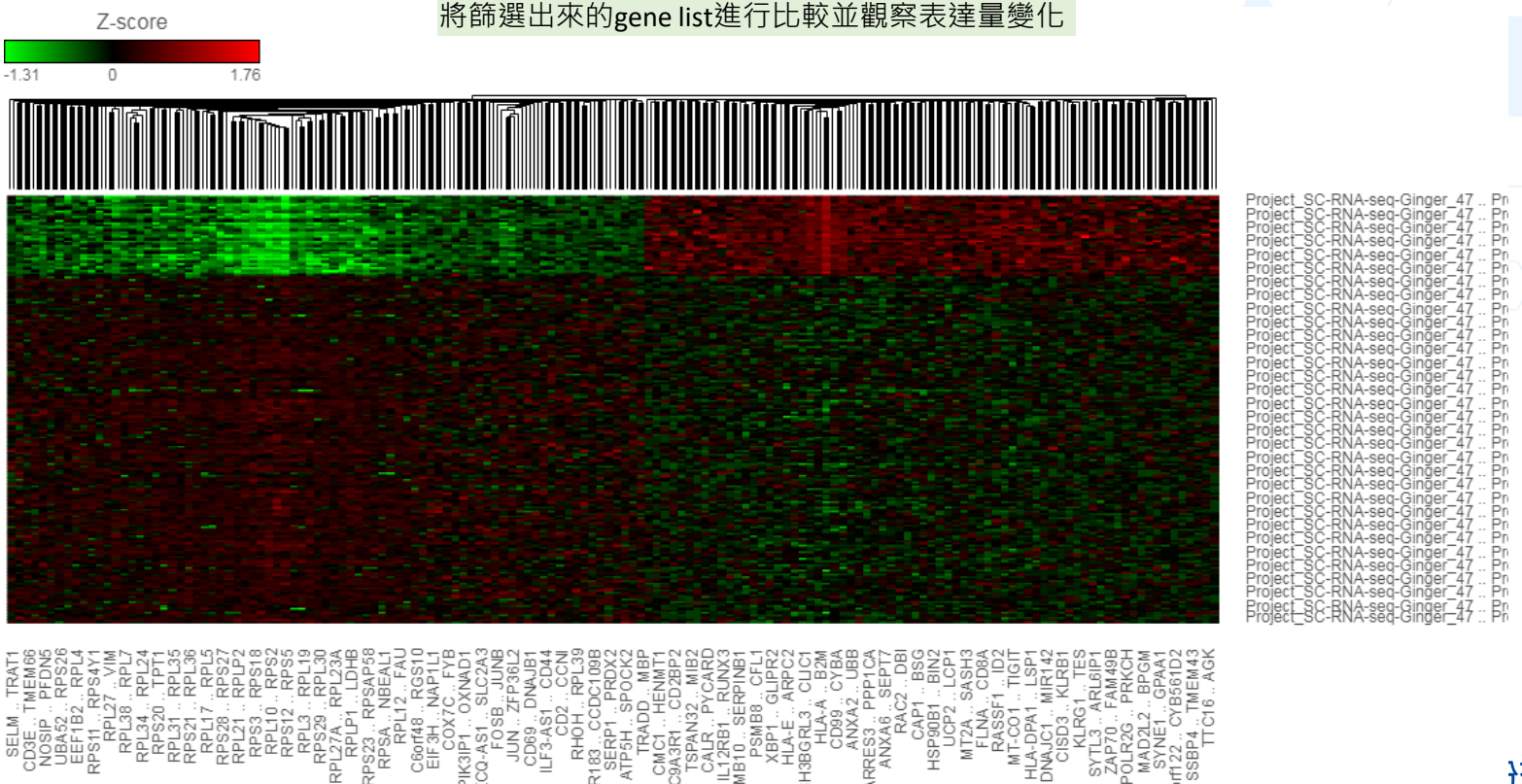
Volcano Plot



用於查看特定細胞群中高表達基因及低表達基因的數量
以左圖說明，X軸為Fold change，Y軸為P-value；
紅點為Up-regulated gene，綠點為Down-regulated gene
由此圖可看出cytotoxic cells 和 T-cells 這兩個種類的細胞群相比之後，
有多少up-regulated, down-regulated 及 un-change 的基因

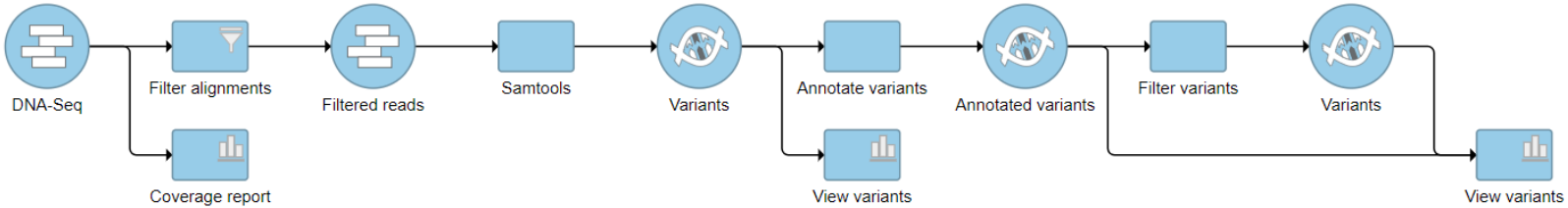
Heatmap

將篩選出來的gene list進行比較並觀察表達量變化

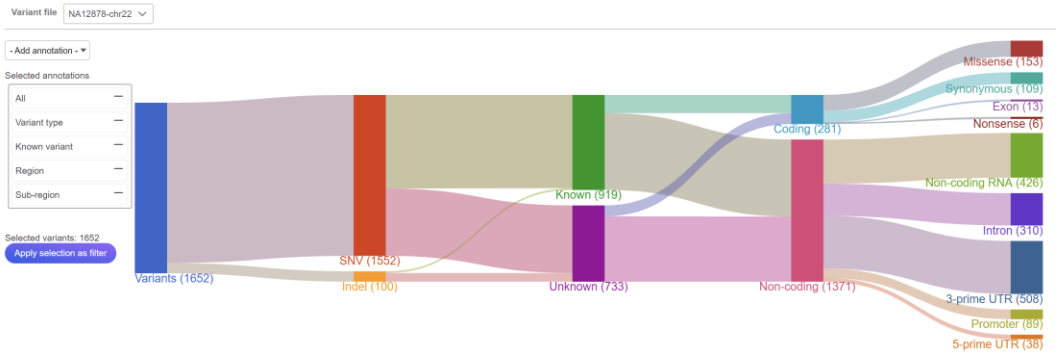


Demo Results of DNA-Seq

DNA-Seq Demo Pipeline



Variants Report



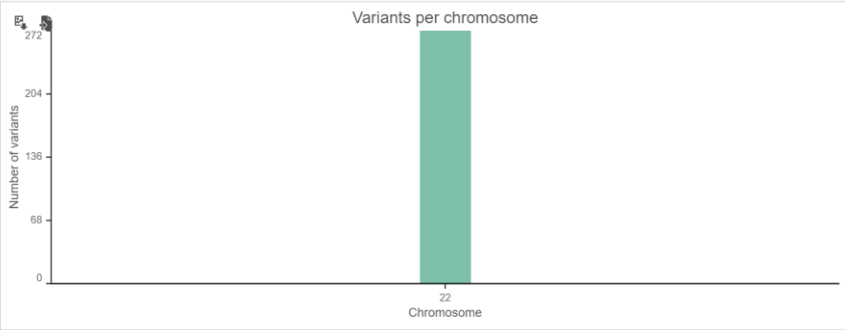
View	Chr	Position	ID	Ref	Alts	VarQual	Type	Gene ID	Transcript ID	Strand	Gene section	Nt change	AA change	Genotype
1	22	16150589	rs2844966	C	A	3.22	Non-coding RNA	DUXAP8	NR_122113	-	Exon 8		None	AA
2	22	16151138	rs11702921	C	T	36.37	Non-coding RNA	DUXAP8	NR_122113	-	Exon 8		None	CT
3	22	16151692	rs2019757	T	C	13.30	Non-coding RNA	DUXAP8	NR_122113	-	Exon 8		None	TC
4	22	16151782	rs7285270	C	T	11.61	Non-coding RNA	DUXAP8	NR_122113	-	Exon 8		None	CT
5	22	16186885	.	A	G	62.00	Non-coding RNA	DUXAP8	NR_122113	-	Exon 6		None	AG
6	22	16186937	.	T	C	6.00	Non-coding RNA	DUXAP8	NR_122113	-	Exon 6		None	TC
7	22	16186939	.	C	G	29.42	Non-coding RNA	DUXAP8	NR_122113	-	Exon 6		None	GG
8	22	16269934	rs2845206	A	G	72.00	Missense	POTEH	NM_001136213	-	Exon 7	c.1247T>C	L416S	GG
9	22	16276569	.	A	G	18.01	Intron	POTEH	NM_001136213	-	Intron 5	c.1166+1179T>C	None	AG
10	22	16276820	rs376561434	G	GT	29.39	Intron indel	POTEH	NM_001136213	-	Intron 5	c.1166+928C>C	None	REF/ALT
11	22	16276889	rs7292440	A	G	49.32	Intron	POTEH	NM_001136213	-	Intron 5	c.1166+859T>C	None	AG
12	22	16277059	rs9680617	C	T	11.16	Intron	POTEH	NM_001136213	-	Intron 5	c.1166+689G>A	None	CT
13	22	16282259	rs3931625	C	T	26.32	Synonymous	POTEH	NM_001136213	-	Exon 3	c.807G>A	None	CT
14	22	16449491	rs202124975	T	G	13.44	Missense	OR11H1	NM_001005239	-	Exon 1	c.314A>C	K105T	TG
15	22	17072483	rs2236639	A	G	103.00	Missense	CCT8L2	NM_014406	-	Exon 1	c.958T>C	W320R	GG

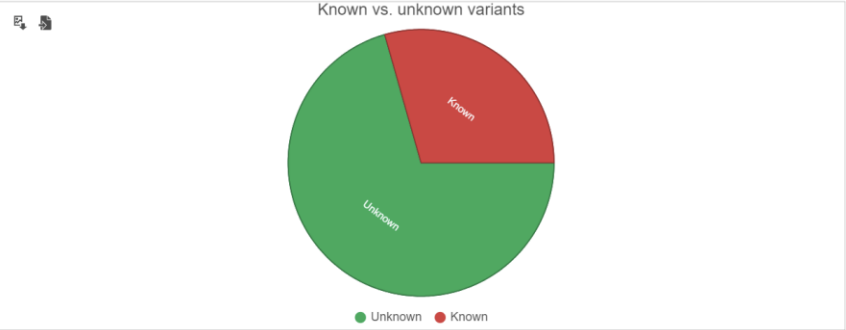
Inspect Variants

Sample name	Average variant quality	Variants	SNVs	Indels	Known variants	% known variants	Ti/Tv ratio	Het/Hom ratio
NA12878-chr22	83.20	272	263	9	80	29.41%	2.05814 (177/86)	23.73 (261/11)

1-1 of 1

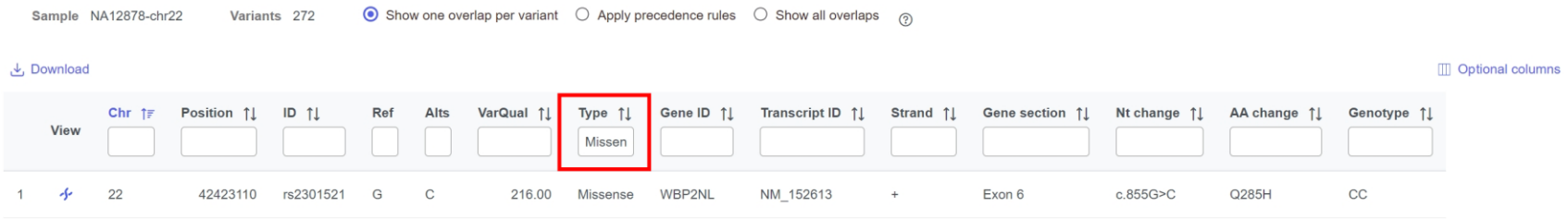
Show 25





Filter Results Table

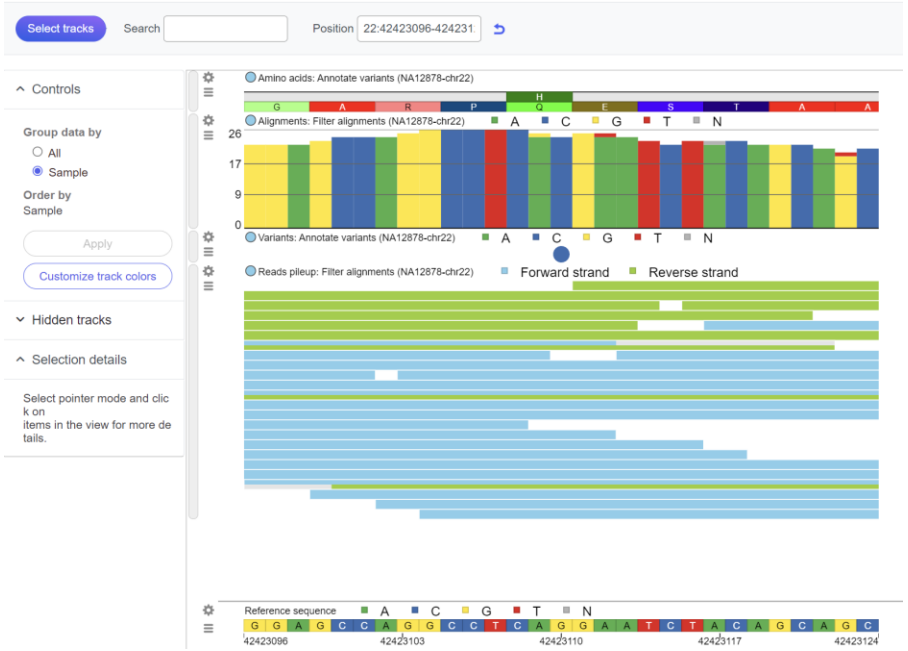
- Click a sample name for a sample-level table
- To select missense variants only, type **Missense** under Type



均泰生物科技股份有限公司
Genetech Biotech Co., Ltd.

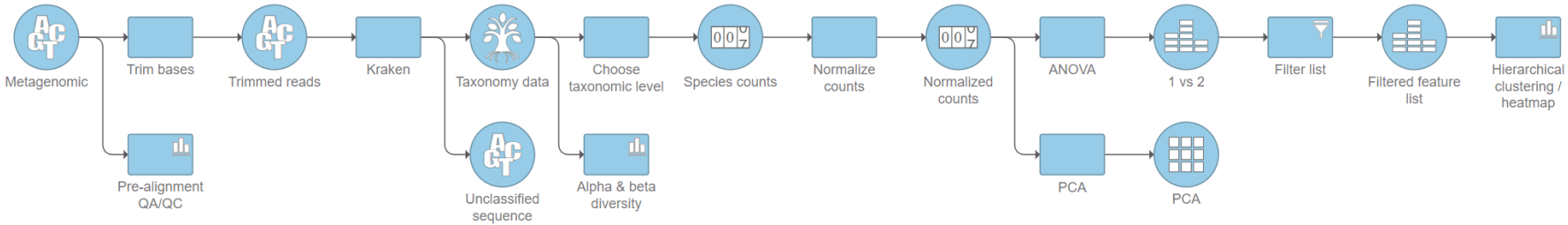
Exploring Browser View

- Click on  to browse in chromosome view
- Click on **Select tracks** to add or remove tracks



Demo Results of Metagenomics

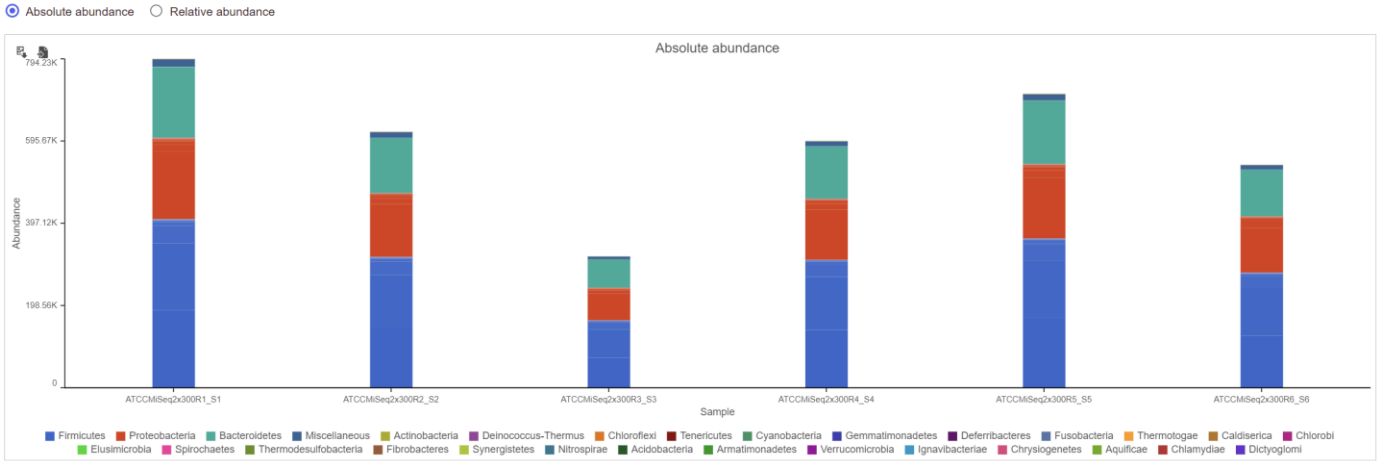
Metagenomics Demo Pipeline



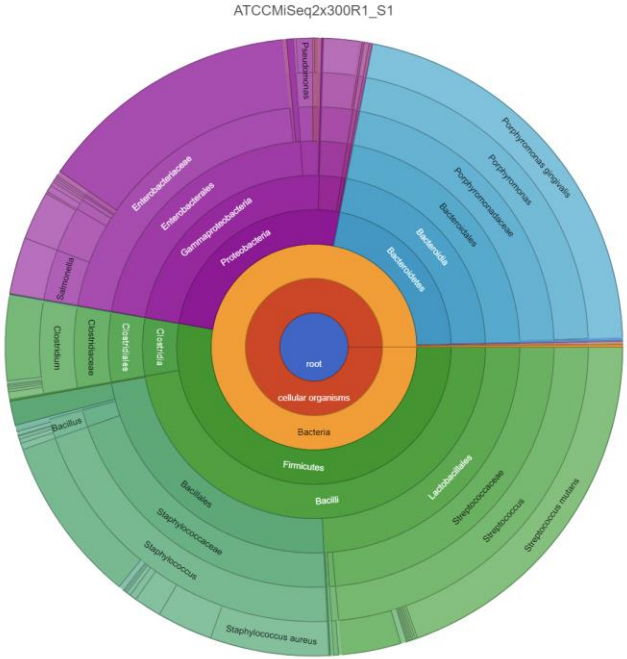
均泰生物科技股份有限公司
Genetech Biotech Co., Ltd.

Kraken Report

Sample name	View	Classified reads	Unclassified reads	Superkingdoms	Kingdoms	Phyla	Classes	Orders	Families	Genera	Species	No rank
ATCCMiSeq2x300R1_S1		794224	891	1	0	28	46	103	212	447	814	2
ATCCMiSeq2x300R2_S2		618119	651	1	0	30	48	95	192	418	748	2
ATCCMiSeq2x300R3_S3		317325	40	1	0	23	38	87	170	352	616	2
ATCCMiSeq2x300R4_S4		595955	79	1	0	24	36	88	172	364	658	2
ATCCMiSeq2x300R5_S5		709840	92	1	0	25	44	97	195	416	742	2
ATCCMiSeq2x300R6_S6		538420	86	1	0	26	43	93	177	371	673	2



Taxanomy Pie Chart



均泰生物科技股份有限公司
Genetech Biotech Co., Ltd.

Alpha Diversity Report

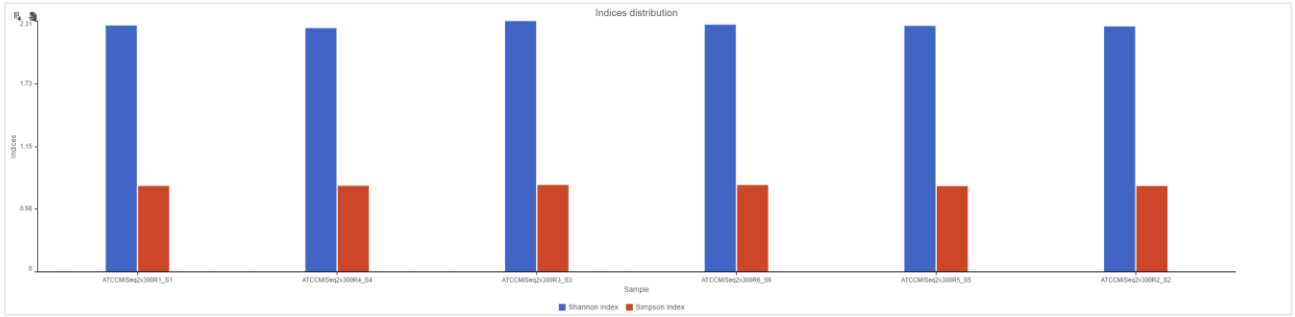
Alpha diversity report | Beta diversity report

Alpha diversity report

Download | Optional columns

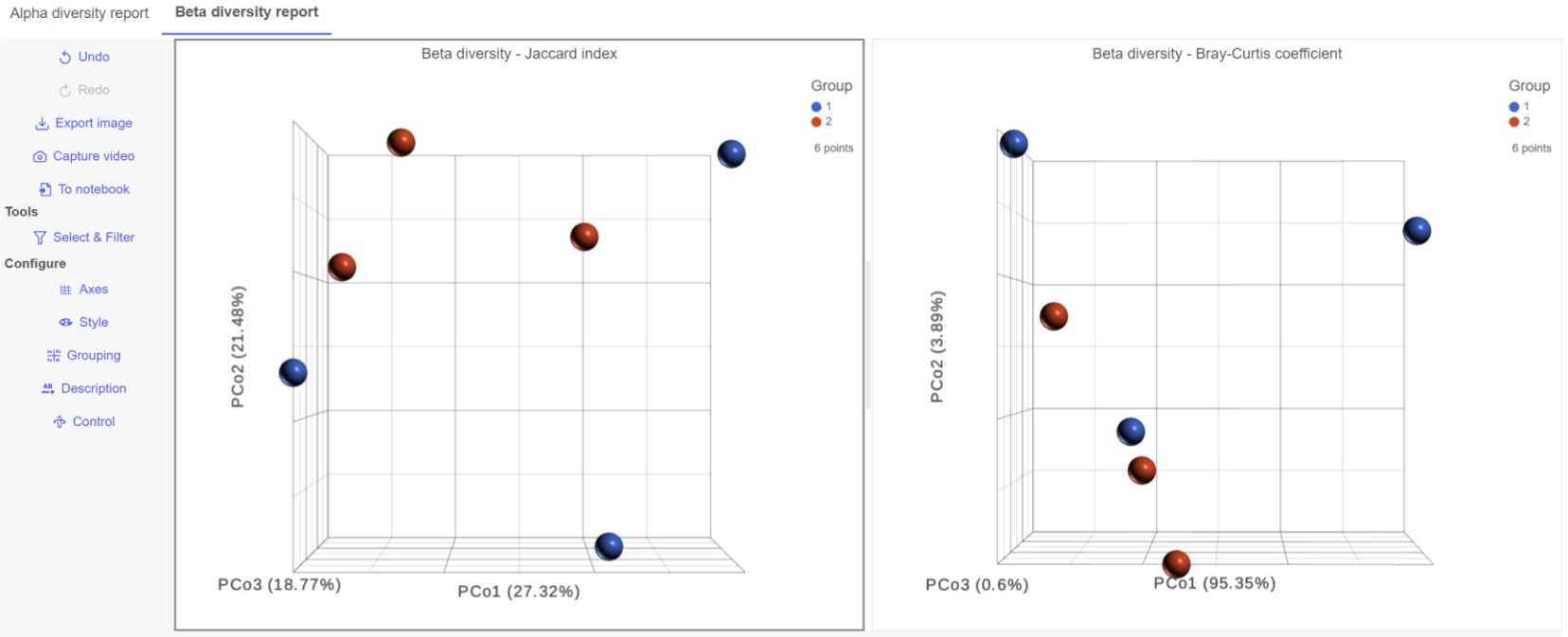
Sample name	Shannon index	Simpson index
ATCCMiSeq2x300R1_S1	2.27	0.79
ATCCMiSeq2x300R2_S2	2.26	0.79
ATCCMiSeq2x300R3_S3	2.31	0.80
ATCCMiSeq2x300R4_S4	2.24	0.79
ATCCMiSeq2x300R5_S5	2.26	0.79
ATCCMiSeq2x300R6_S6	2.27	0.80

1-6 of 6 | Show 25

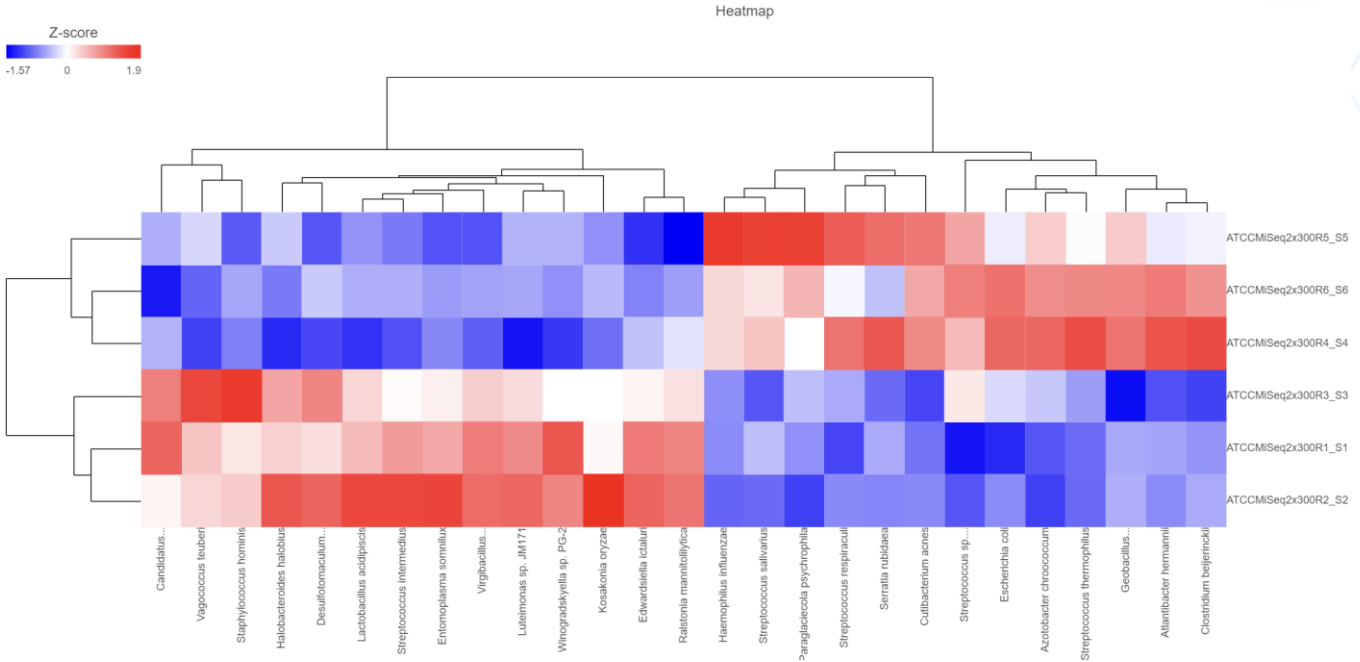


均泰生物科技股份有限公司
Genetech Biotech Co., Ltd.

Beta Diversity Report



Compare Abundance between Samples

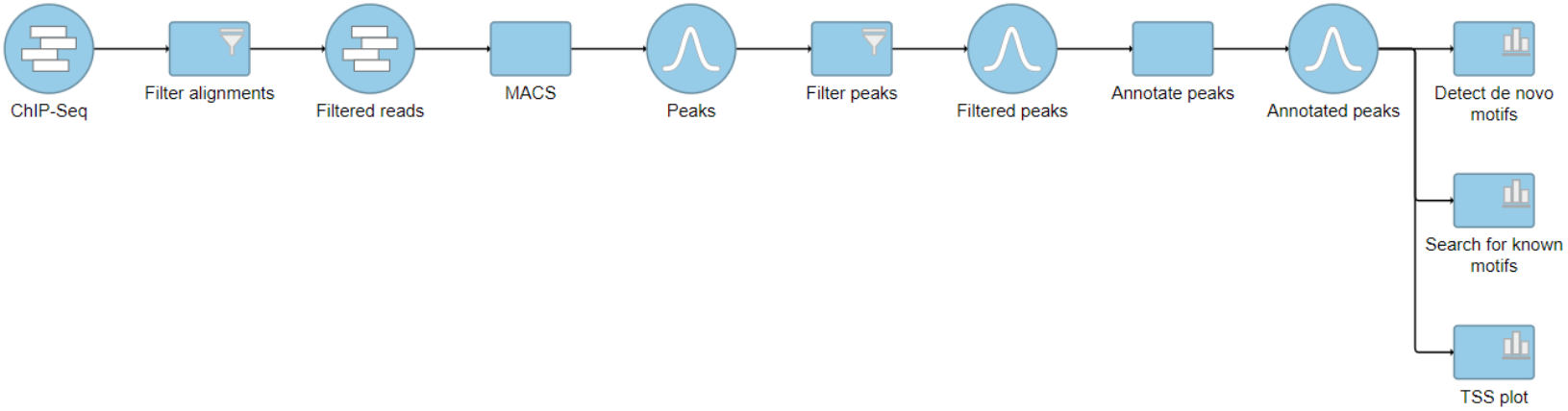


均泰生物科技股份有限公司
Genetech Biotech Co., Ltd.

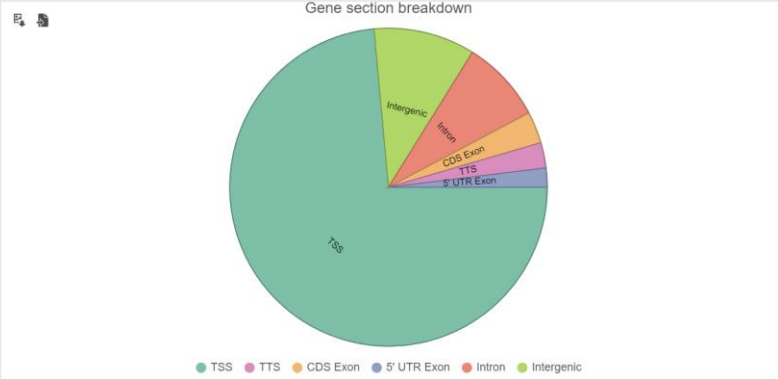
Demo Results of ChIP-Seq

均泰生物科技股份有限公司
Genetech Biotech Co., Ltd.

ChIP-Seq Demo Pipeline



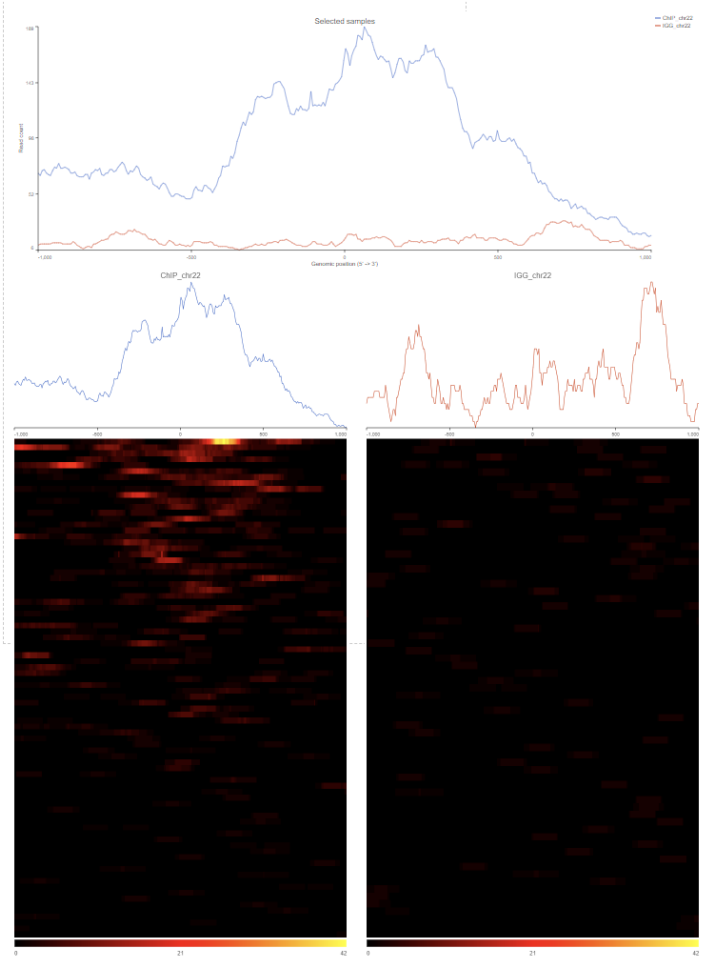
Peak Annotation



Optional columns

	Chromosome	Start	End	Length	-log10(pvalue)	Fold enrichment	Sample name	Gene IDs	Transcript IDs	Gene section	Distance to TSS	Distance to TTS
	22	50,638,871	50,640,269	1,398	45.12	20.36	ChIP_chr22_vs_IGG_chr22	SELENOO	NM_031454	TSS	0	-15,777
	22	31,884,921	31,886,689	1,768	40.55	11.97	ChIP_chr22_vs_IGG_chr22	EIF4ENIF1	NM_001164501	TSS	0	-49,576
	22	36,783,579	36,783,989	410	40.27	18.50	ChIP_chr22_vs_IGG_chr22	MYH9	NM_002473	TSS	124	-106,256

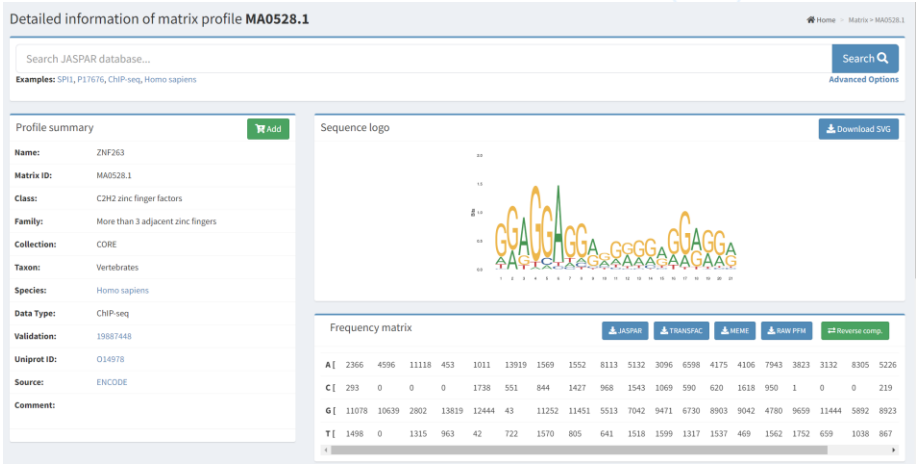
TSS plot



Search for Known Motifs

Summary | Detail

Motif name	Consensus sequence	p-value
ZNF263 (MA0528.1)	RRRGGAGGRNDRDVDDRRRRR	2.89E-291
BPC1 (MA1404.1)	RAGAGAGAGAGAGAGAGAGAGA	9.78E-177
BPC5 (MA1403.1)	AGAGAGAGAGAGAGAGAGAGAGAGAGAG	3.49E-163
AT1G71450 (MA1233.1)	HNNCDHCDHHDYCDCCGHCD	7.39E-150
ZNF384 (MA1125.1)	NNNNAAAAAANN	2.13E-114
HMG-I/Y (MA0045.1)	NDWVRRRNRVMRDMRH	1.45E-100



均泰生物科技股份有限公司
Genetech Biotech Co., Ltd.

Detect de novo Motifs

Assembly

Homo sapiens (human) - hg19_chr22

Number of motifs

The number of de novo motifs to find.

1

Motif length

The range of motif lengths to examine.

6 bp to 16 bp

Motif name	Consensus sequence	Log likelihood ratio	Sequence logo
motif1	YYTYYTYY	2,109.03	

均泰生物科技股份有限公司
Genetech Biotech Co., Ltd.