Gene Expression Analysis by CLC Genomic Workbench



創源生物科技公司
分子視算中心
Willis鄭耀瑋
資深工程專員

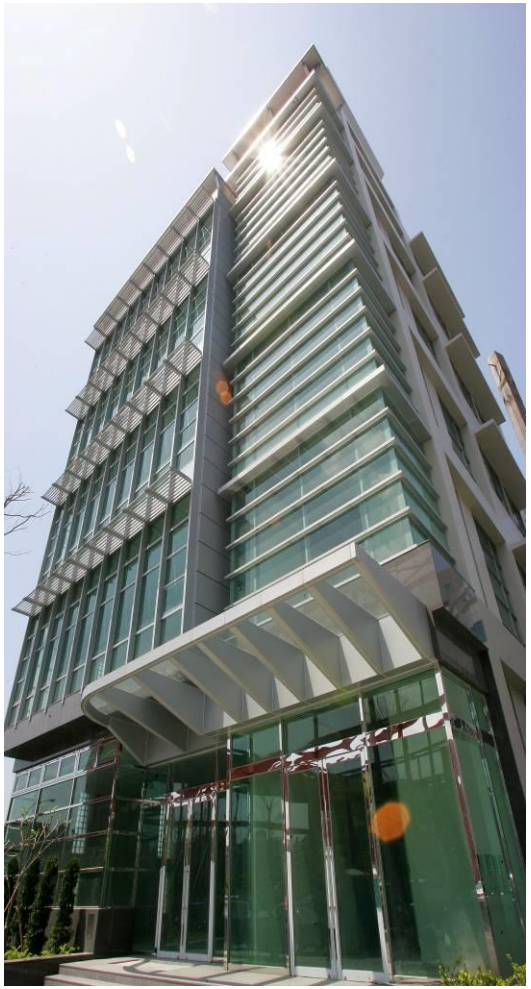
Copyright©2018 GGA Corp. All rights reserved.

Outline

- Introduction
  - Gene Expression Analysis by NGS
  - QIAGEN Solution
- Case Studies for NGS data by CLC
  - RNA-Seq Analysis
- Q&A

Copyright©2018 GGA Corp. All rights reserved.

About GGA



GGA is part of the BIONET Group (訊聯生物科技)

CEO: Christopher Tsai, Ph.D. 蔡政憲 博士

Established: Nov. 2008

Main Product & Service Areas:

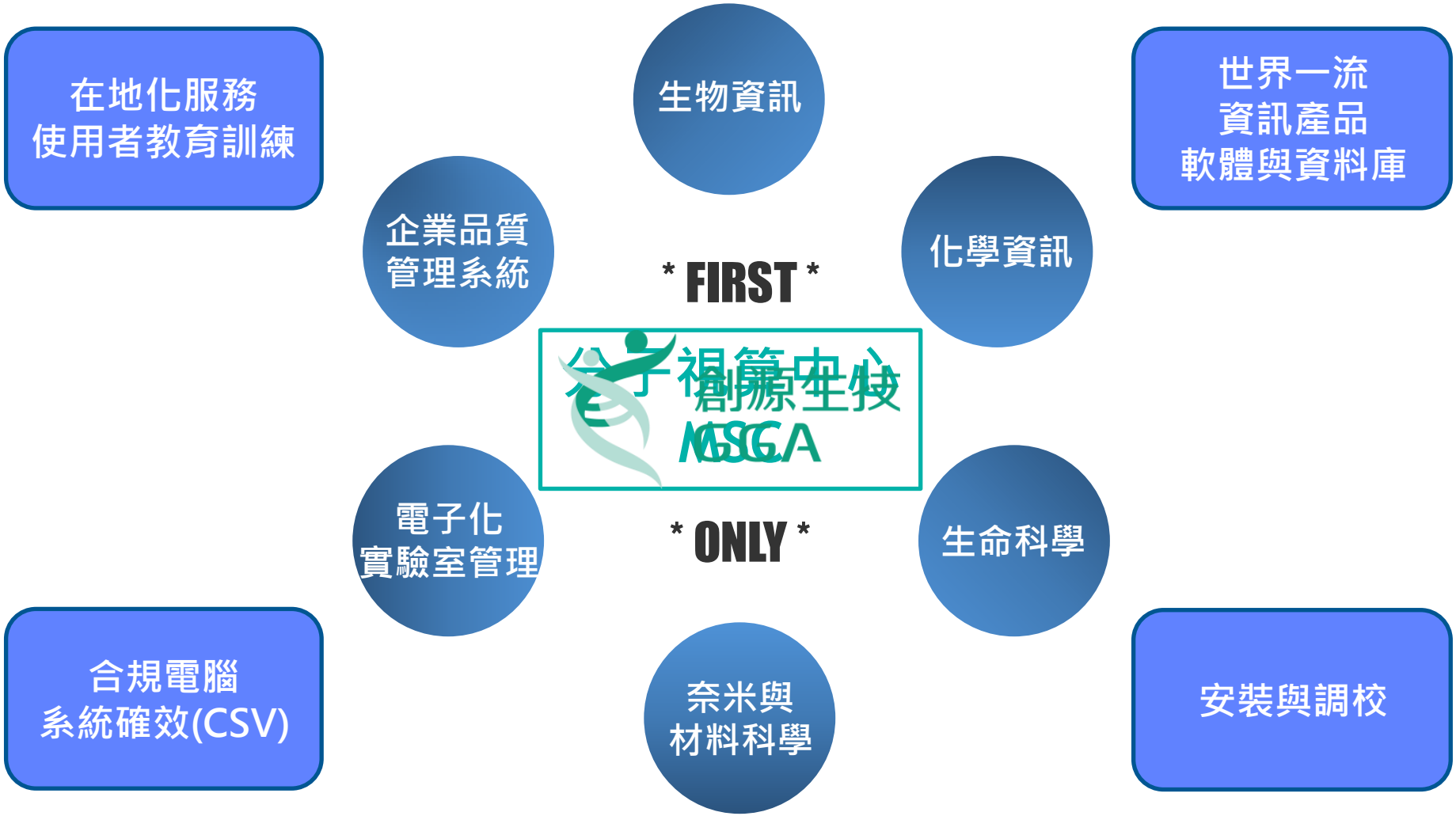
- Genetic Testing & Molecular Diagnosis
- **Scientific Informatics & Bio IT**

IPO Date: September 17, 2012

Stock Ticker: 4160 (Taiwan OTC)

Copyright©2018 GGA Corp. All rights reserved.

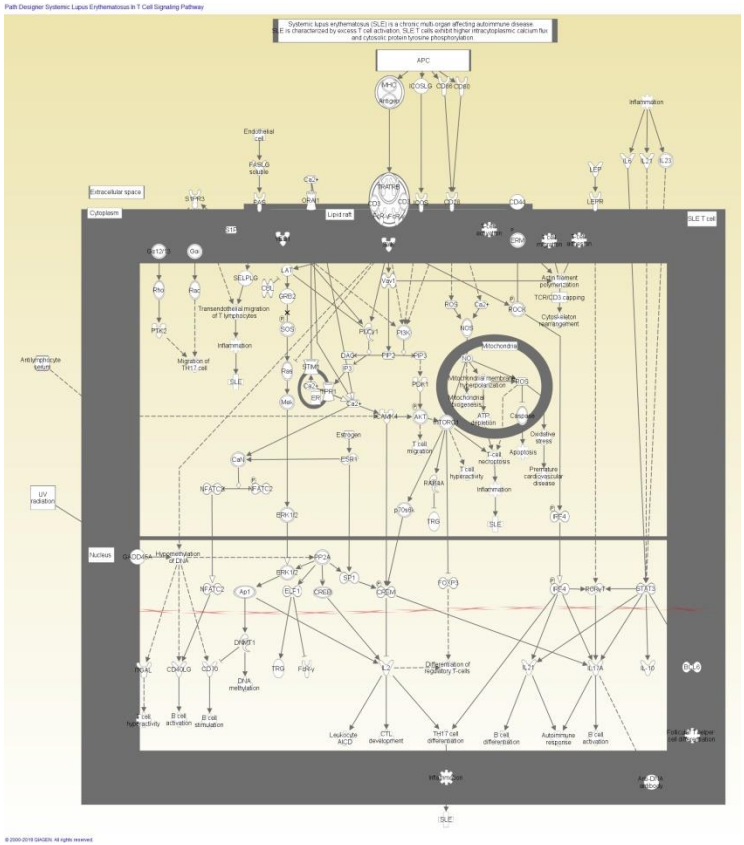
Molecular Science Center in GGA



Copyright©2018 GGA Corp. All rights reserved.

Pathway Analysis

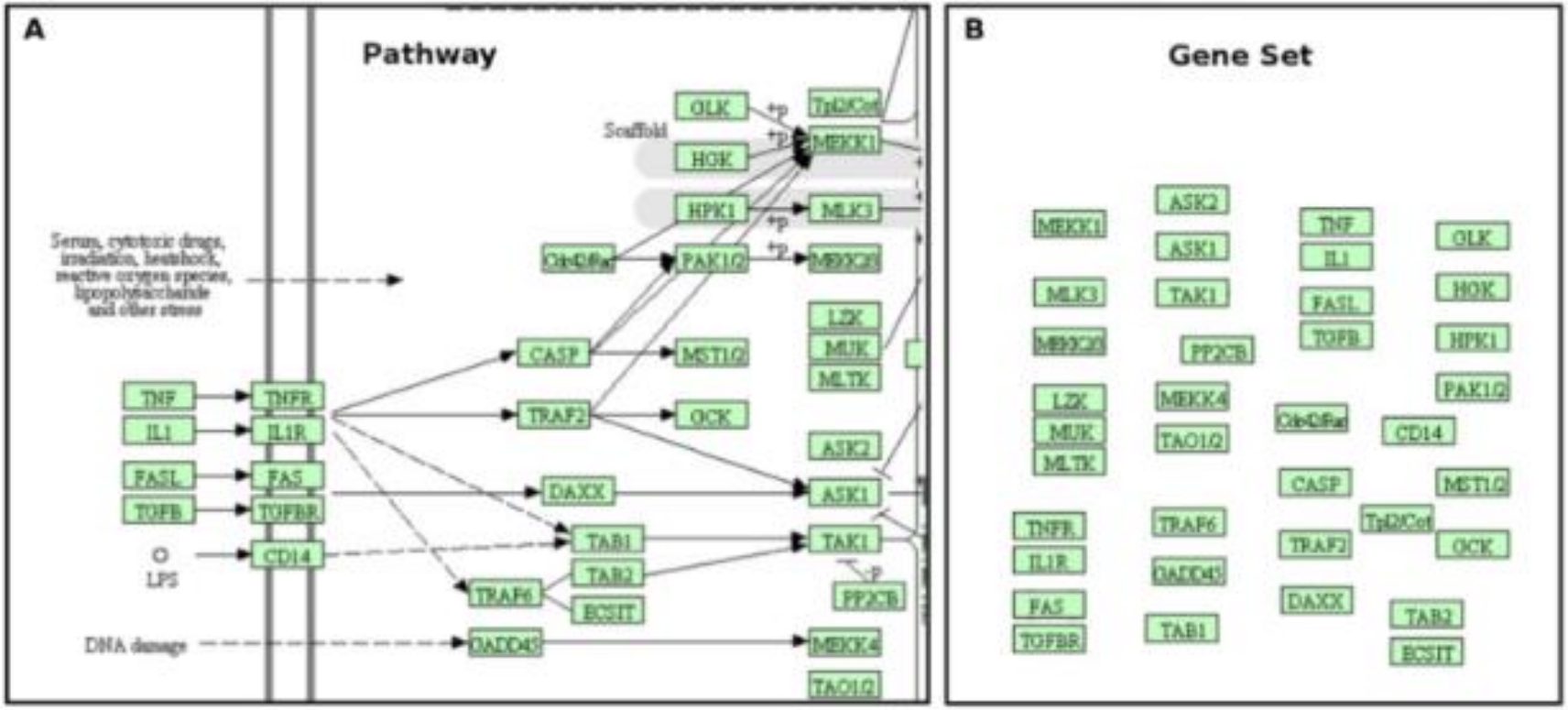
- Useful for finding out :
  - If the differentially expressed genes are associated with a certain biological process or molecular function.



Copyright©2018 GGA Corp. All rights reserved.

Gene set analysis v.s. Pathway analysis

- Pathway analysis can describe each gene set can affect with each others.



https://advaitabio.com/ipathwayguide/pathway-analysis-vs-gene-set-analysis/

Copyright©2018 GGA Corp. All rights reserved.

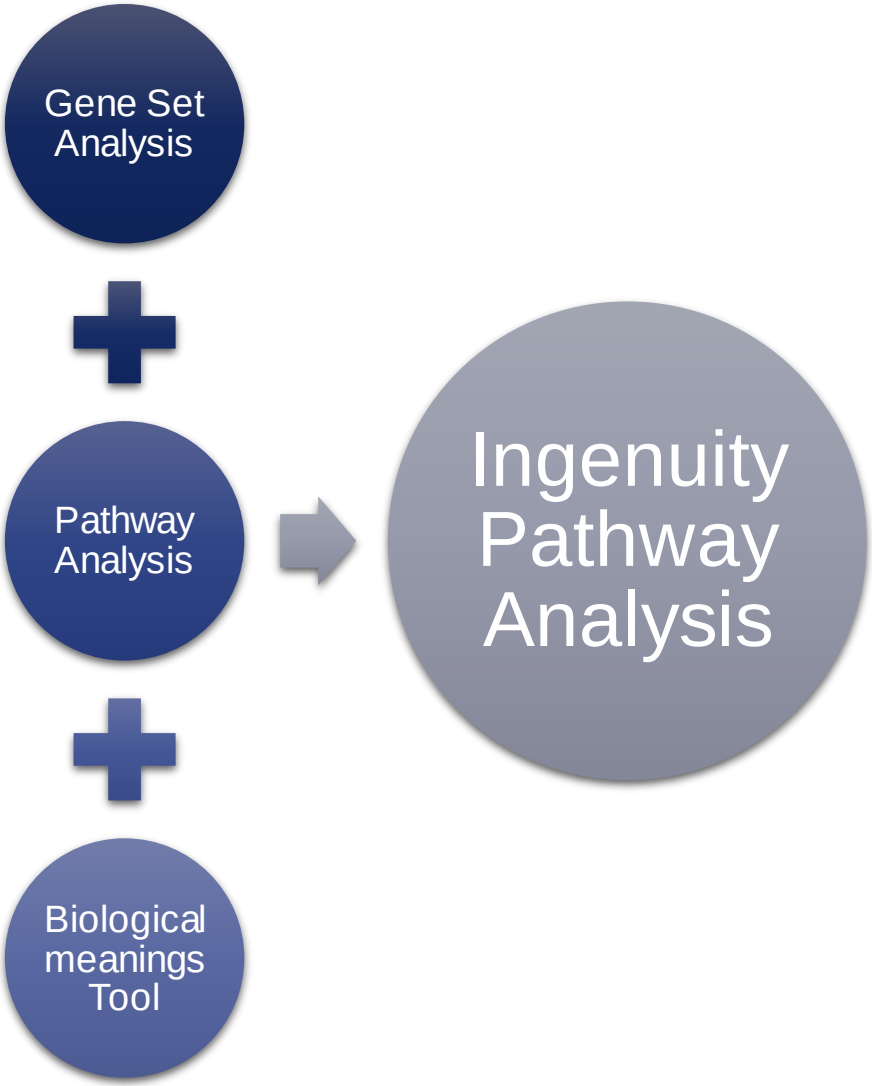
Database for these analysis

- Gene set
  - MsigDB, Gene Ontology, QIAGEN-Knowledge base database
- Pathway set
  - KEGG, QIAGEN-Knowledge base database
- Others

GO term	Description	Detected...	DE Genes	P-values	FDR p-value	Bonferroni
0042711	maternal behavior	14	3	4.17E-9	3.41E-5	5.45E-5
0060746	parental behavior	15	3	5.21E-9	3.41E-5	6.82E-5
0019098	reproductive behavior	38	3	9.65E-8	4.20E-4	1.26E-3
0045924	regulation of female receptivity	9	2	2.39E-6	7.80E-3	0.03
0060180	female mating behavior	12	2	4.37E-6	0.01	0.06
0007625	grooming behavior	15	2	6.96E-6	0.02	0.09
0035813	regulation of renal sodium excretion	20	2	1.26E-5	0.02	0.16
0044062	regulation of excretion	24	2	1.83E-5	0.03	0.24
0007617	mating behavior	28	2	2.50E-5	0.04	0.33
0098801	regulation of renal system process	33	2	3.49E-5	0.04	0.46
0032892	positive regulation of organic acid transport	34	2	3.71E-5	0.04	0.48
0051954	positive regulation of amine transport	34	2	3.71E-5	0.04	0.48
0035176	social behavior	36	2	4.16E-5	0.04	0.54
0051703	intraspecies interaction between organisms	36	2	4.16E-5	0.04	0.54
0045777	positive regulation of blood pressure	38	2	4.65E-5	0.04	0.61
0048609	multicellular organismal reproductive process	303	3	5.13E-5	0.04	0.67
1903793	positive regulation of anion transport	44	2	6.25E-5	0.05	0.82
0051705	multi-organism behavior	54	2	9.44E-5	0.07	1.00
0032890	regulation of organic acid transport	56	2	1.02E-4	0.07	1.00

Copyright©2018 GGA Corp. All rights reserved.

Ingenuity Pathway Analysis



Copyright©2018 GGA Corp. All rights reserved.

Gene Expression Analysis

- Goal:
  - To measure the gene expression profile
  - To evaluate the differential expression gene



NGS data
Array data

Gene	A	B
KRT15	19,881.69	217.09
KRT23	15,766.60	618.04
RPL19	14,088.17	16,551.90
TOP2A	12,243.82	1,302.26

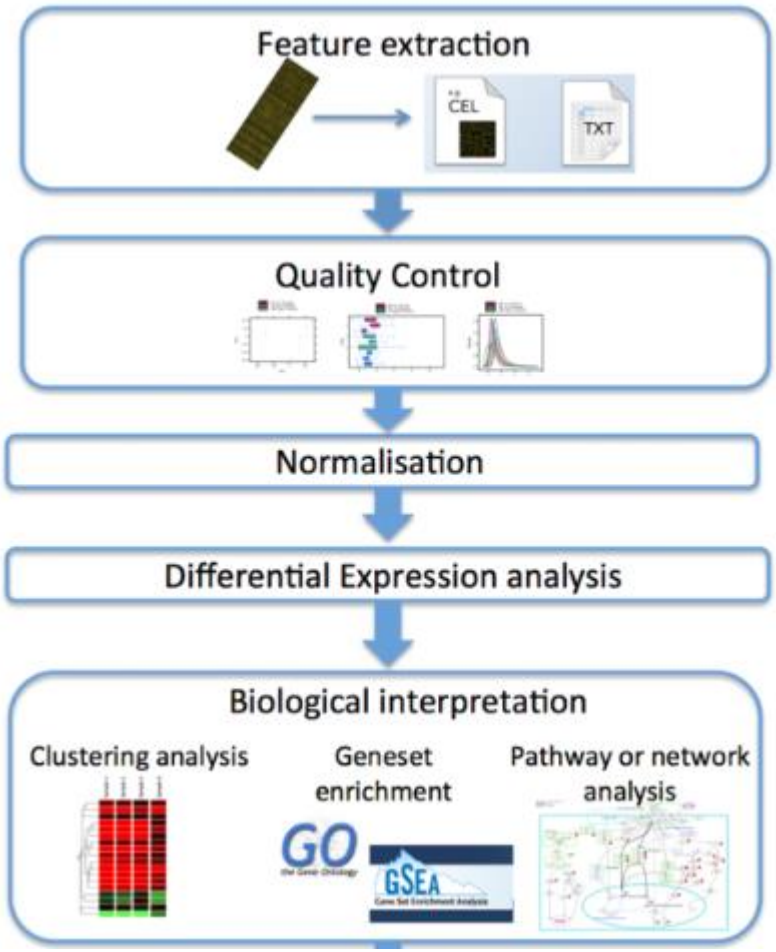
Expression profiles

Gene	Log_2 fold change	P-value	FDR p-value
ERBB2	4.17E+00	5.16E-07	1.50E-03
GRB7	3.55E+00	1.10E-05	0.02
STARD3	2.74	7.65E-05	0.07
STAC2	-3.90	1.28E-04	0.09

Differential expression profiles

Copyright©2018 GGA Corp. All rights reserved.

Gene expression analysis on array data



Example

— CEL, txt files

— PCA

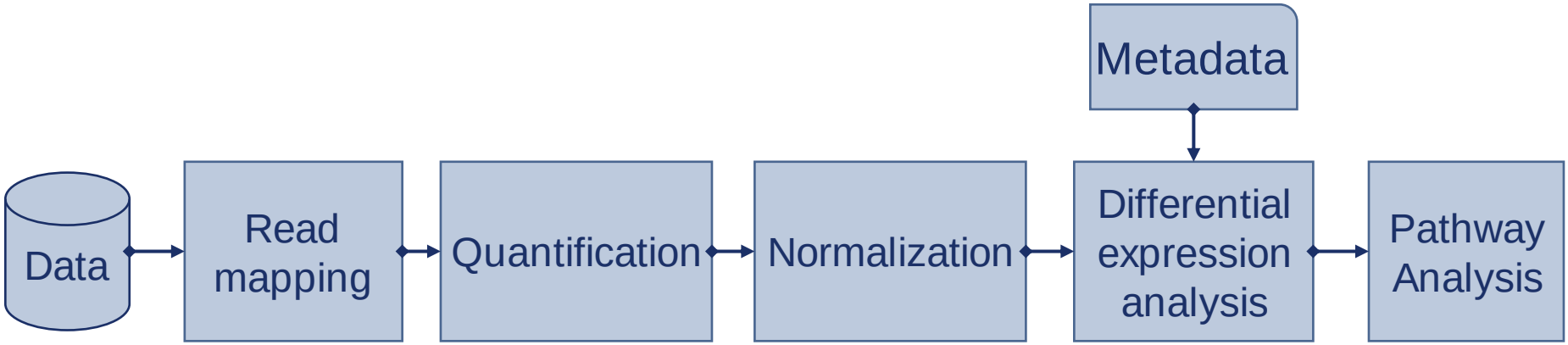
— Quantile

— Fold change, t-test, ajust p-value

https://www.ebi.ac.uk/training/online/course/functional-genomics-ii-common-technologies-and-data-analysis-methods/analysis-microarray-data

Copyright©2018 GGA Corp. All rights reserved.

General Pipeline (NGS)



Open Source Tools	HISATS	IDP	TC	DESeq	DAVID
	STAR	StringTie	UQ	edgeR	GSEA
	TopHat	Cufflinks	Med	limma	Reactome
		RSEM	TMM	sleuth	R
		eXpress	Q	CuffDiff	
			RPKM	Ballgown	
			Raw Count		
QIAGEN	CLC Genomic Workbench				IPA

http://toolsbiotech.blog.fc2.com/blog-entry-106.html

Copyright©2018 GGA Corp. All rights reserved.

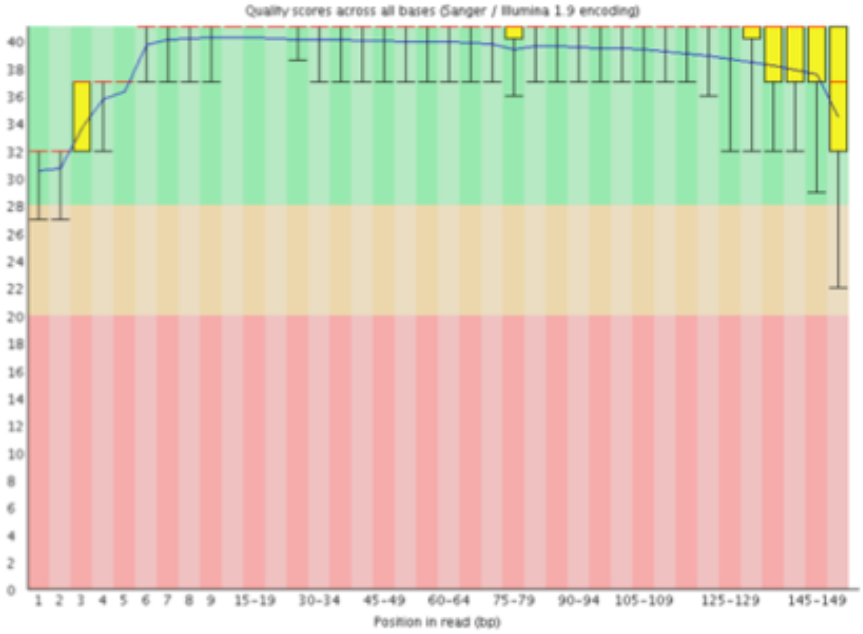
Quality Control

- To evaluate which data can be analyzed
- Criteria:
  - Read length
  - GC Content
  - Phred score
  - Per Base Sequence
  - Others

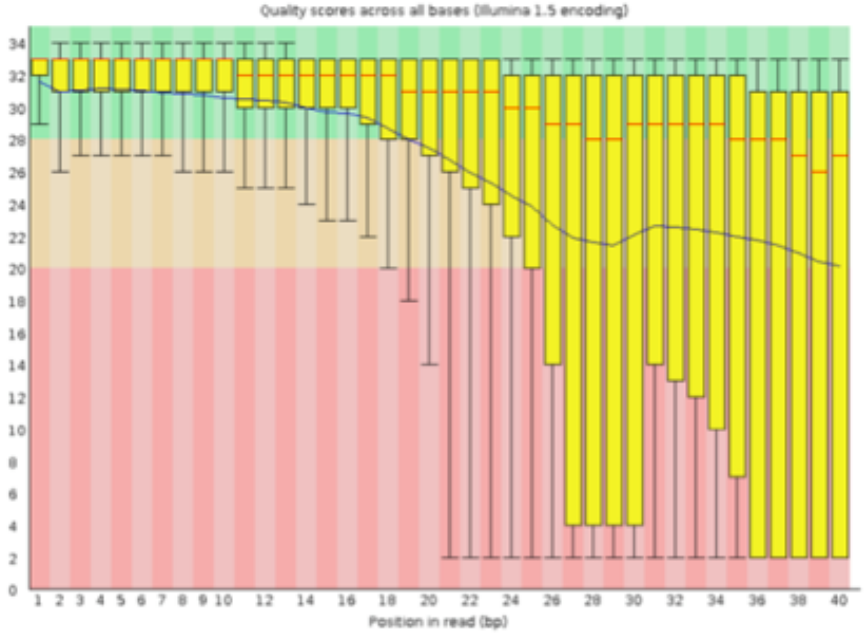
Copyright©2018 GGA Corp. All rights reserved.

Quality Control – Phred Quality Score

Good data



Bad data



$$Q = -10 \log_{10} P$$

$$P = 10^{-\frac{Q}{10}}$$

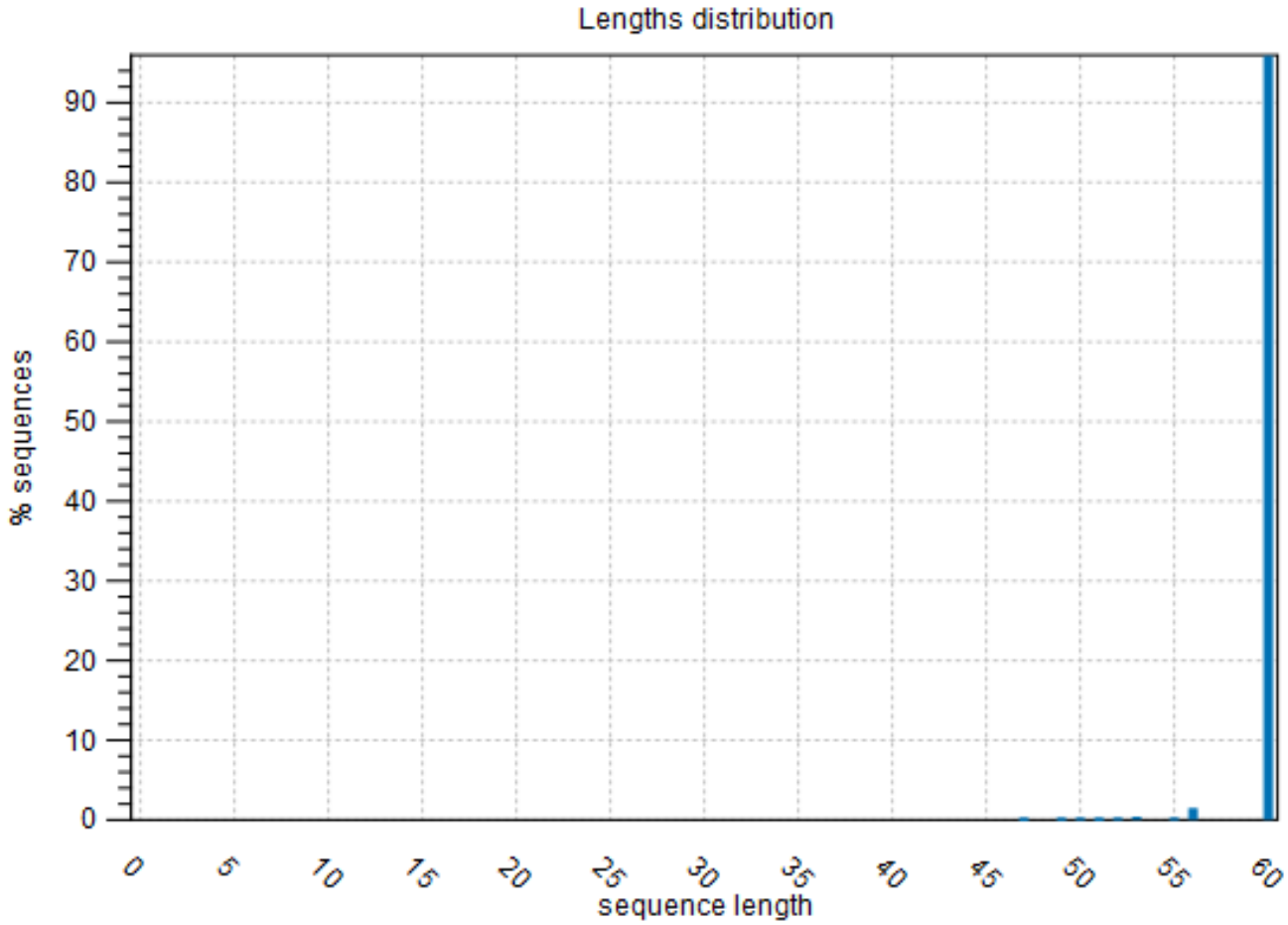
Phred Quality Score	Probability of incorrect base call	Base call accuracy
10	1 in 10	90%
20	1 in 100	99%
30	1 in 1000	99.9%
40	1 in 10,000	99.99%
50	1 in 100,000	99.999%

https://blog.horizondiscovery.com/diagnostics/the-5-ngs-qc-metrics-you-should-know

Copyright©2018 GGA Corp. All rights reserved.

Quality Control – Length distribution

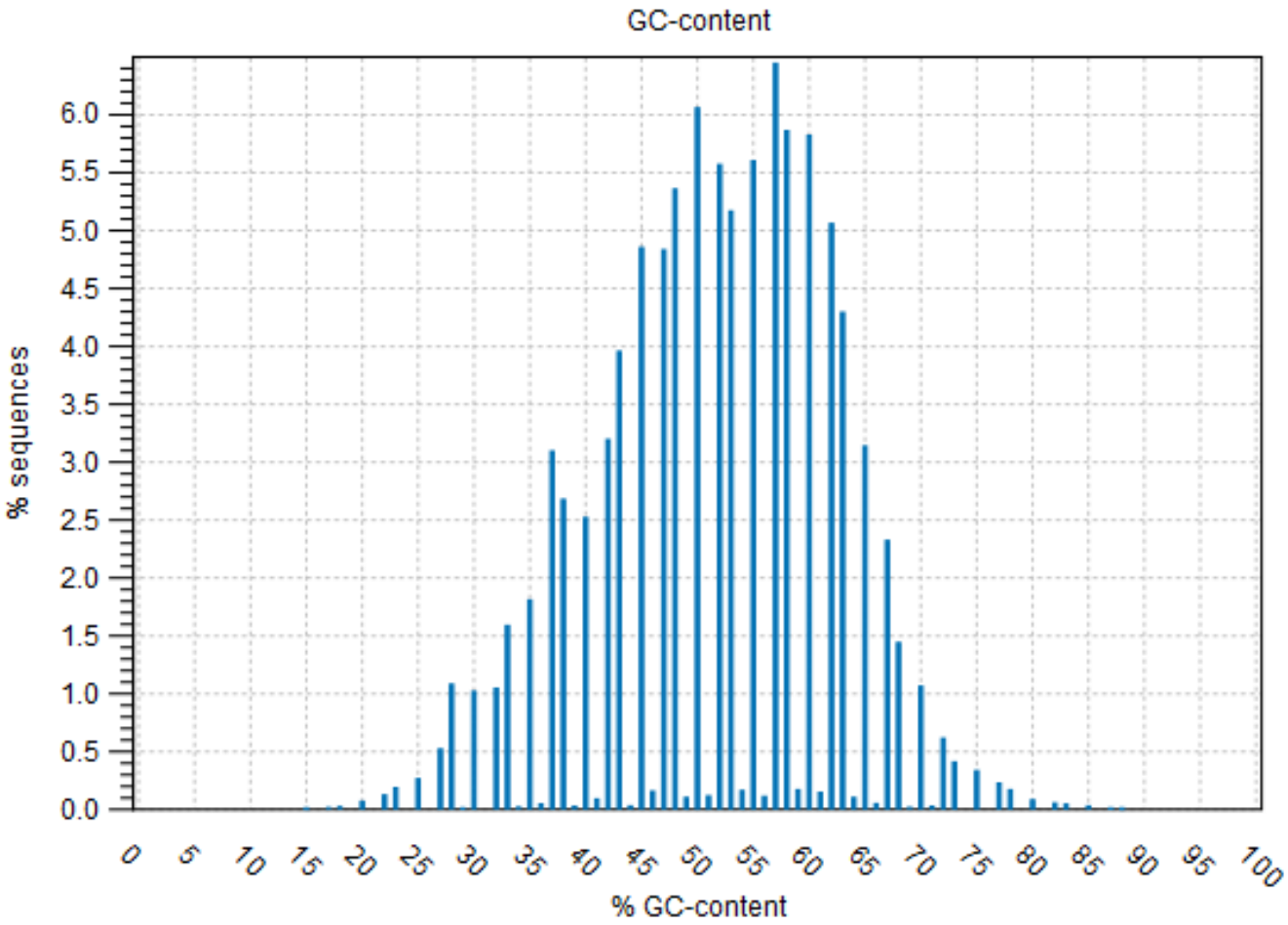
- Length distribution



Copyright©2018 GGA Corp. All rights reserved.

Quality Control – GC Content

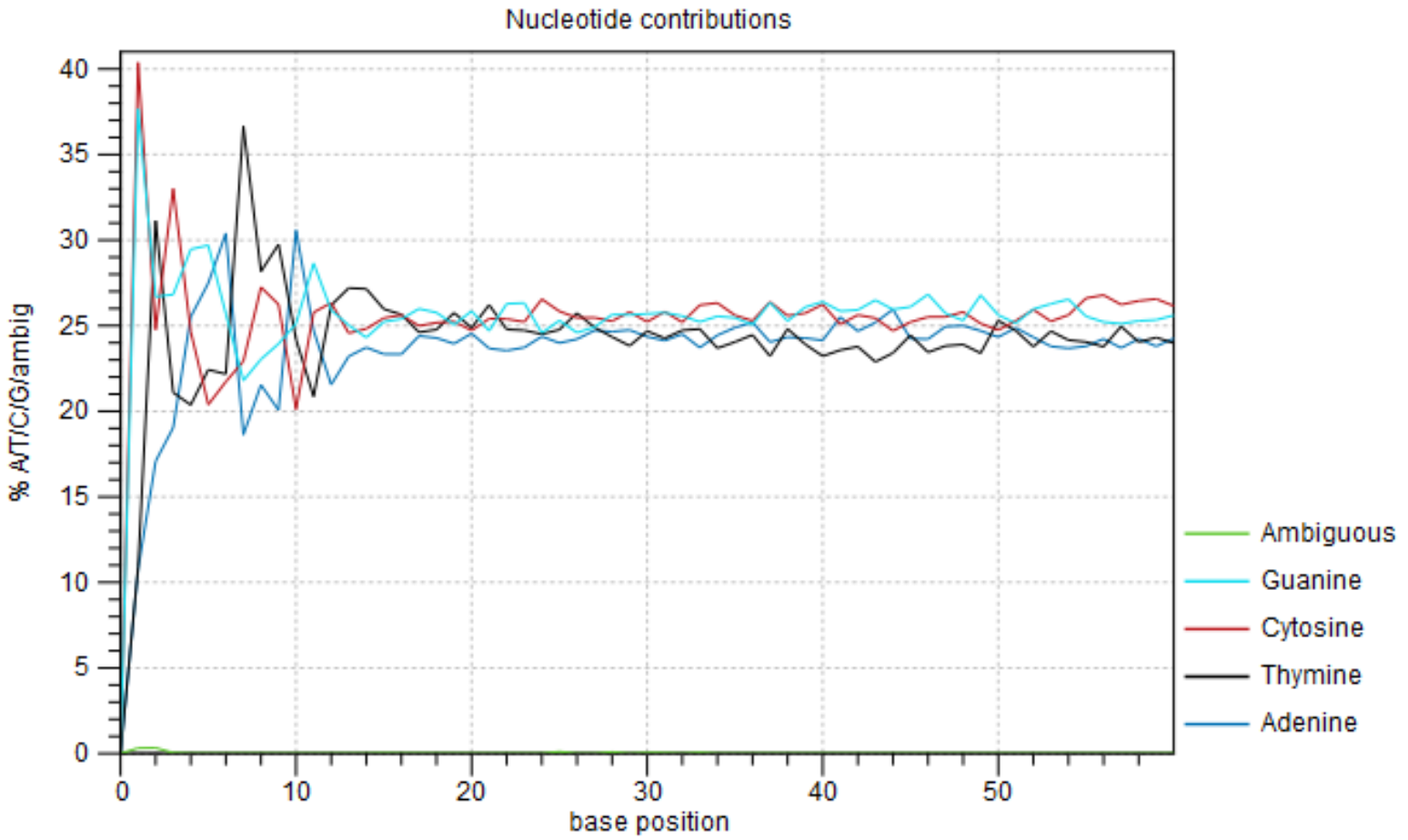
- GC Content



Copyright©2018 GGA Corp. All rights reserved.

Quality Control – Per base sequence

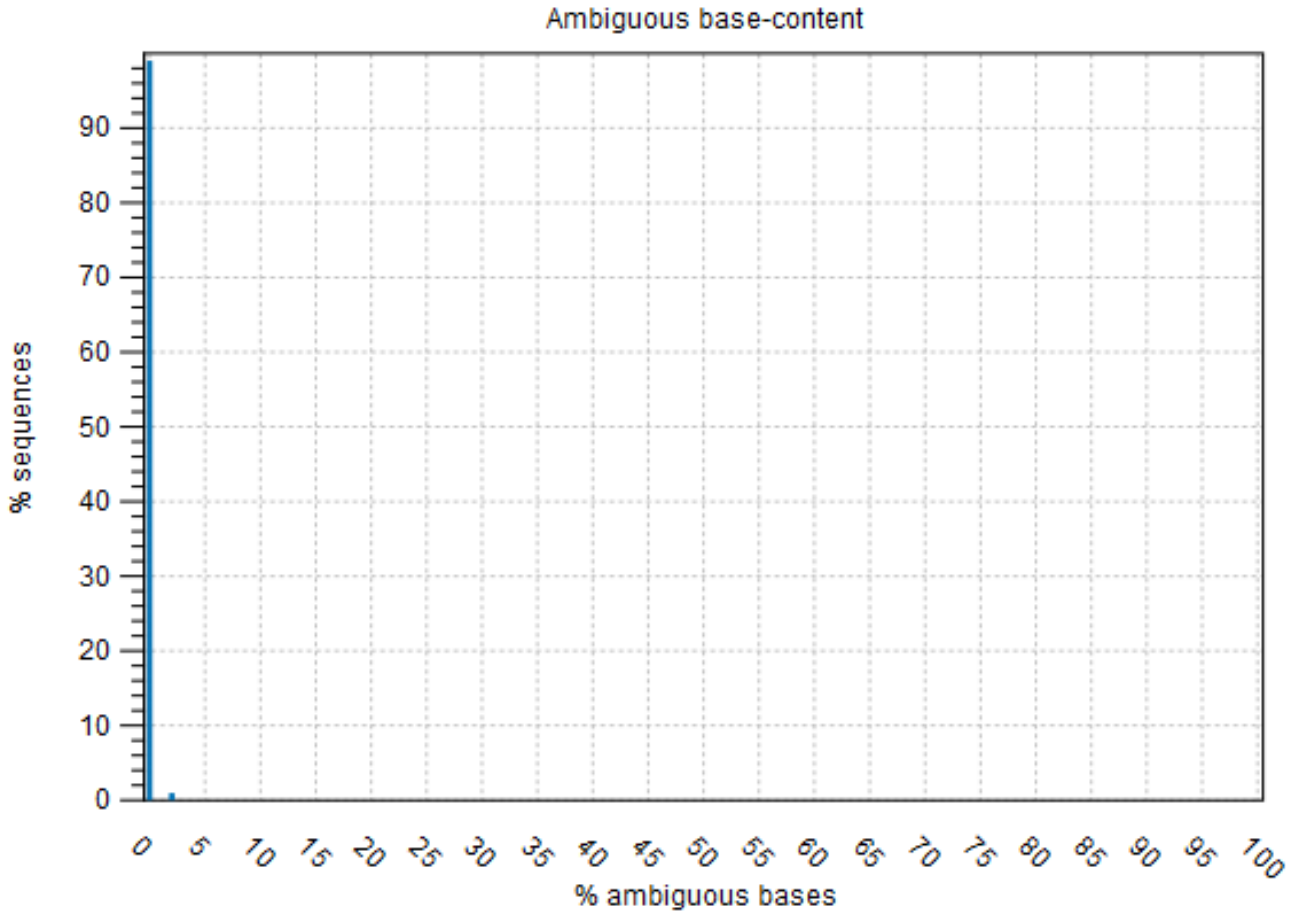
- Per base sequence



Copyright©2018 GGA Corp. All rights reserved.

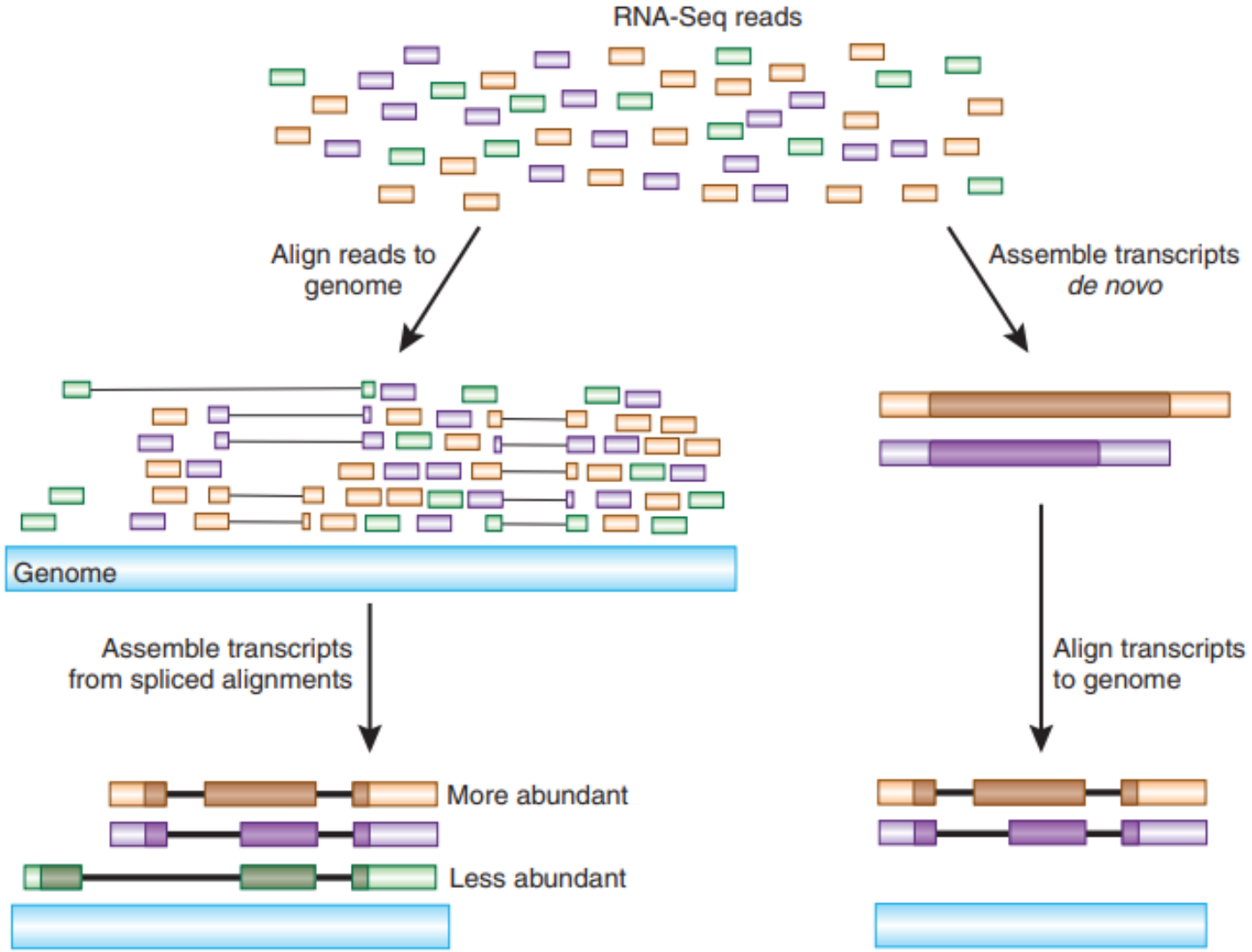
Quality Control – Ambiguous base content

- Ambiguous base content



Copyright©2018 GGA Corp. All rights reserved.

Read mapping – Strategy

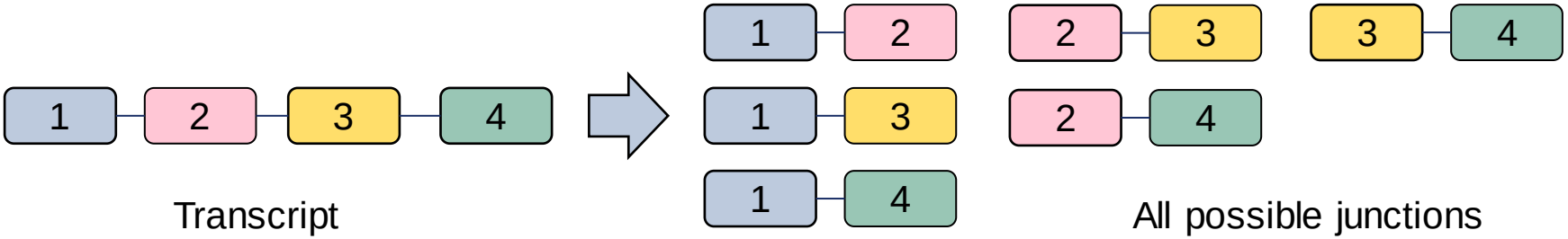


Nature Biotechnology. 2010(28):421-423

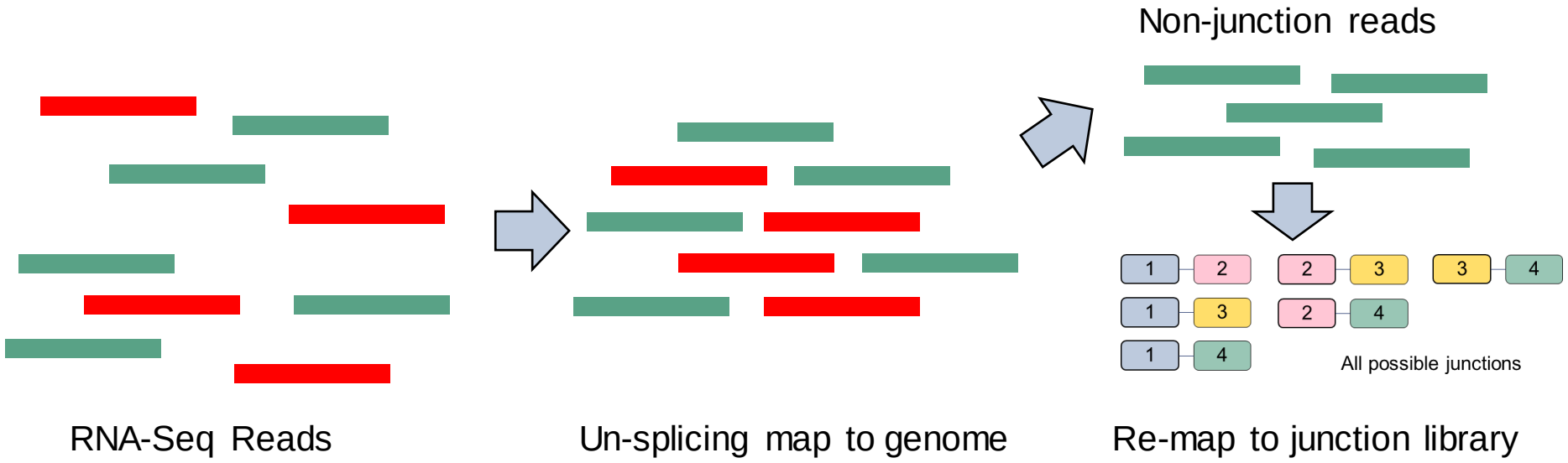
Copyright©2018 GGA Corp. All rights reserved.

Read mapping – Join junction

1. Build junction library for each known transcript



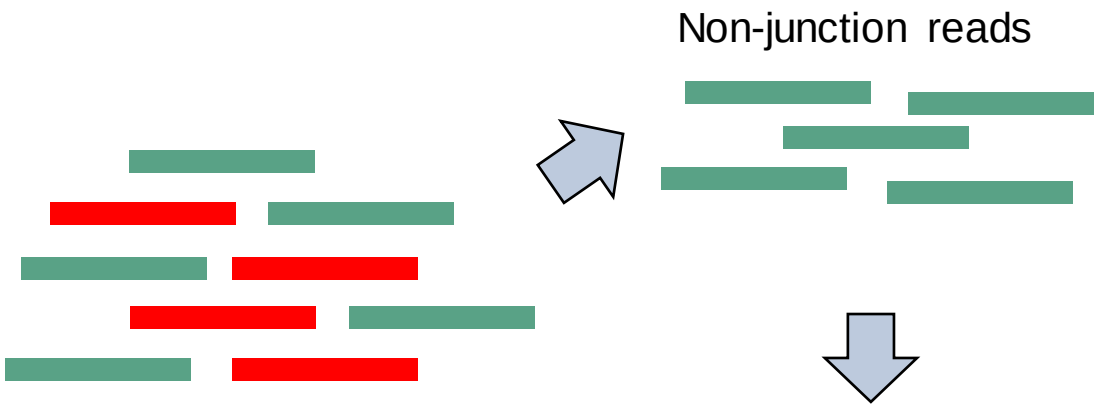
2. Map RNA-Seq reads to the genome and junction library

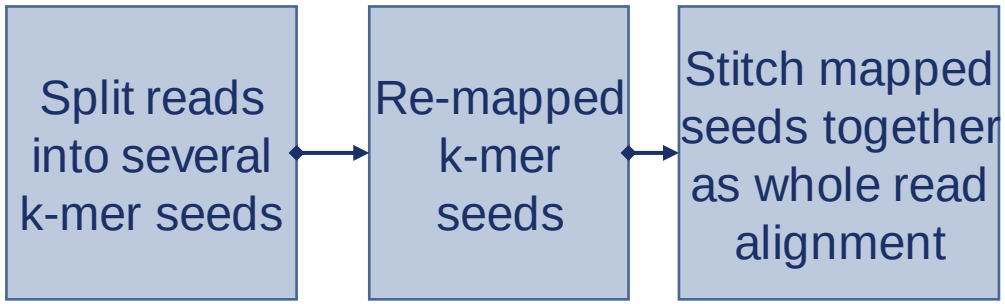


https://zh-tw.coursera.org/lecture/bioinformatics-pku/rna-seq-mapping-assembling-kSa63

Copyright©2018 GGA Corp. All rights reserved.

Read mapping – Split reads



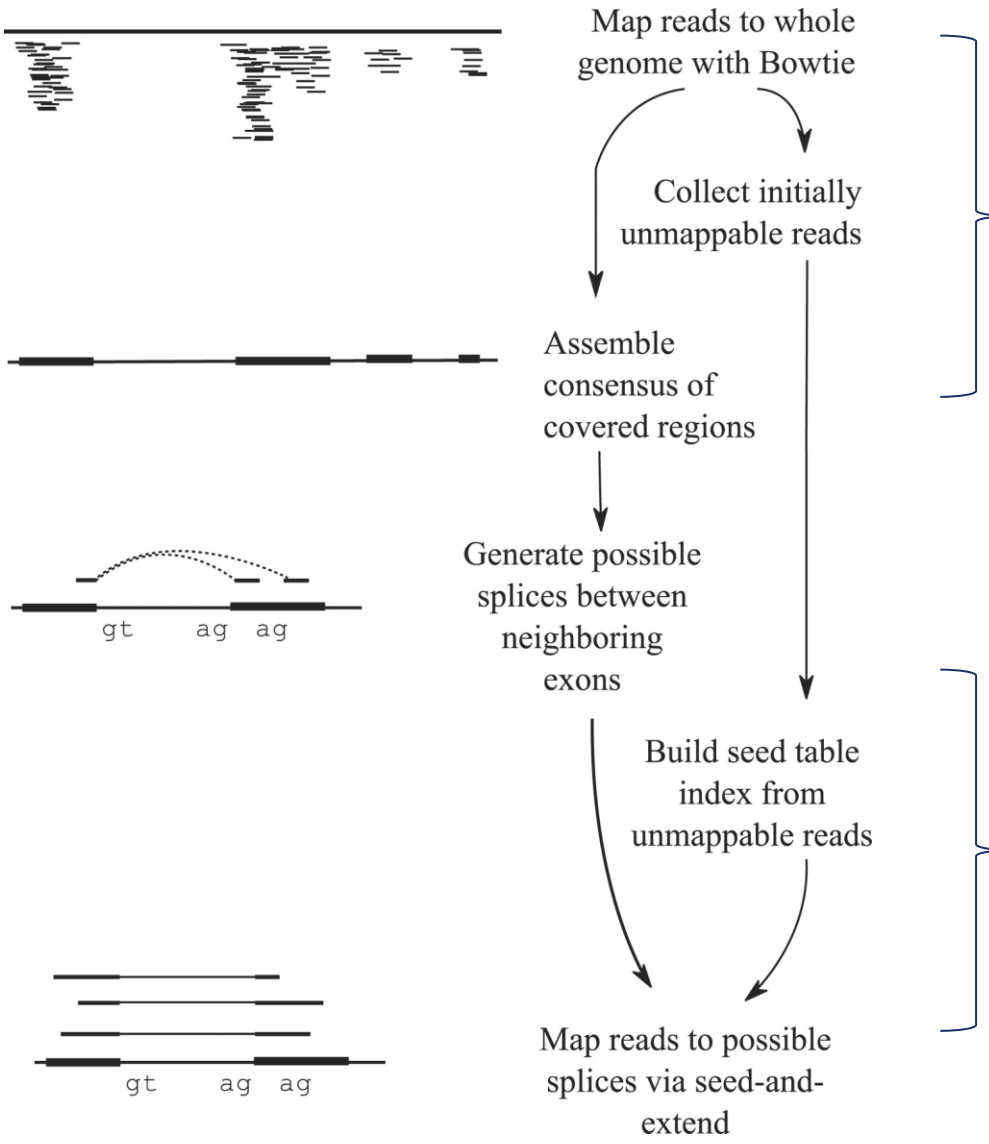




Copyright©2018 GGA Corp. All rights reserved.

Read mapping – TopHat (hybrid methods)

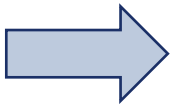
- TopHat



Copyright©2018 GGA Corp. All rights reserved.

Quantification

- Index:
  - Coverage
  - Gene length
  - Reads in exons



Gene	A	B
KRT15	19,881.69	217.09
KRT23	15,766.60	618.04
RPL19	14,088.17	16,551.90
TOP2A	12,243.82	1,302.26

Copyright©2018 GGA Corp. All rights reserved.

Quantification

- CLC Model (The EM estimation algorithm)
  - Like RSEM and eXpress method
  - It iteratively estimates the abundance of transcripts/genes, and assigns reads to transcripts/genes according to these abundances.



Copyright©2018 GGA Corp. All rights reserved.

Normalization

- TMM
  - TMM normalization adjusts library sizes based on the assumption that most genes are not differentially expressed.
    - 'log CPM' (Counts per Million) values are calculated for each gene. The CPM calculation uses the effective library sizes as calculated by the TMM normalization.
    - After this first normalization, a second one is performed across samples for each gene: the counts for each gene are mean centered, and scaled to unit variance.
    - Genes or transcripts with zero expression across all samples or invalid values (NaN or +/- Infinity) are removed.

Copyright©2018 GGA Corp. All rights reserved.

Differential Expression

- Generalized Linear Model (GLM)
  - assume that the read counts follow a **Negative Binomial distribution**
    - Similarly to EdgeR and DESeq
  - Calculate log2 fold change
  - Fit distribution and test the raw p-value
  - Transform to adjust p-value

Copyright©2018 GGA Corp. All rights reserved.

How to prepare final data for IPA

- Any text data include gene symbol and values
  - Gene expression profiles
  - Differential expression profiles
  - Others
- You can use EXCEL to manipulate your data

Gene	A	B
KRT15	19,881.69	217.09
KRT23	15,766.60	618.04
RPL19	14,088.17	16,551.90
TOP2A	12,243.82	1,302.26

Expression profiles

Gene	Log_2 fold change	P-value	FDR p-value
ERBB2	4.17E+00	5.16E-07	1.50E-03
GRB7	3.55E+00	1.10E-05	0.02
STARD3	2.74	7.65E-05	0.07
STAC2	-3.90	1.28E-04	0.09

Differential expression profiles

Copyright©2018 GGA Corp. All rights reserved.

How much genes you choose for IPA

- 95% confidence interval
  - Genes: 20000
  - After filter by distribution: about 1000 genes
- Too much genes is not good for IPA
  - It is easy to let pathway enriched

Copyright©2018 GGA Corp. All rights reserved.



CLC Genomic Workbench

Copyright©2018 GGA Corp. All rights reserved.

Sample to Insight Workflow

Sample

Sample Prep → Assay Data → Sequence-Level Statistics → Biology of Interest (Genes, Variants, etc) → Comparative Statistical Analysis → Biological Interpretation

Insight








Upstream Analysis | "primary" | "secondary" | "tertiary"

QIAGEN

QIAcube
GeneRead™ Panel

Sequencer

CLC bio A QIAGEN Company

Genomics
Workbench

POWERED BY INGENUITY

Variant/Pathway
Analysis

Streamline Integration and Workflow Embedding

Copyright©2018 GGA Corp. All rights reserved.

Powered by…



No background in bioinformatics required

Would you prefer to own both the experimental design and the data analysis?
We're bring
scientists, n

Xxxxxx

When data turns to insights

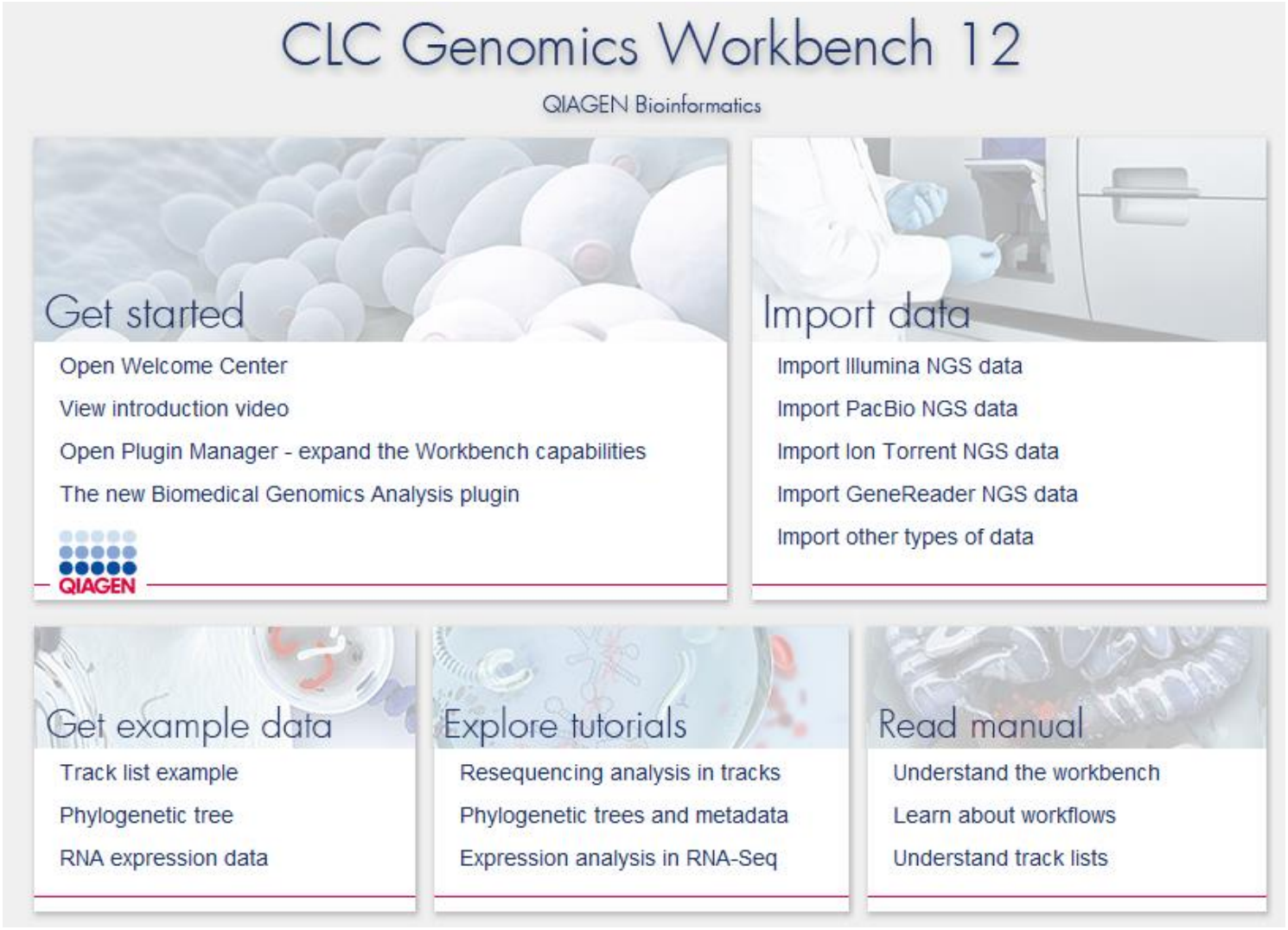
CLC Genomics Workbench

CLC Genomics Server



Copyright©2018 GGA Corp. All rights reserved.

The new CLC Genomics Workbench 12



Copyright©2018 GGA Corp. All rights reserved.

In CLC Workbenches you can..

Tools

- Genome Finishing Module
- GeneMark Gene Finding
- Microbial Genomics Module
- Classical Sequence Analysis
- Molecular Biology Tools
- BLAST
- Track Tools
- Prepare Sequencing Data
- Resequencing Analysis
- RNA-Seq Analysis
- Microarray and Small RNA Analysis
- Epigenomics Analysis
- De Novo Sequencing
- Installed Workflows
- Utility Tools
- QIAseq Panel Expert Tools
- Legacy Tools



QC & Reads Processing	●
Resequencing (whole genome, exome, targeted)	●
Transcriptomics (RNA-Seq)	●
De novo assembly	●
Epigenomics	●
Microbiome Analysis (Microbial Genomics Module)	●
Contigs Assembly (Genome Finishing Module)	●
Workflow (Pipeline)	●
Microarray Analysis	●
Phylogeny Tools	●
Blast, Sanger Sequencing, Cloning, Primer Design, …	●

Copyright©2018 GGA Corp. All rights reserved.

CLC Genomic Workbench 12

- Track viewing and navigation improvements
- A Reference Data Manager, for finding, downloading and managing reference data.
- Access to biomedical genomics and targeted panel data analyses through the new Biomedical Genomics Analysis plugin.
- Improved data storage efficiency

Copyright©2018 GGA Corp. All rights reserved.

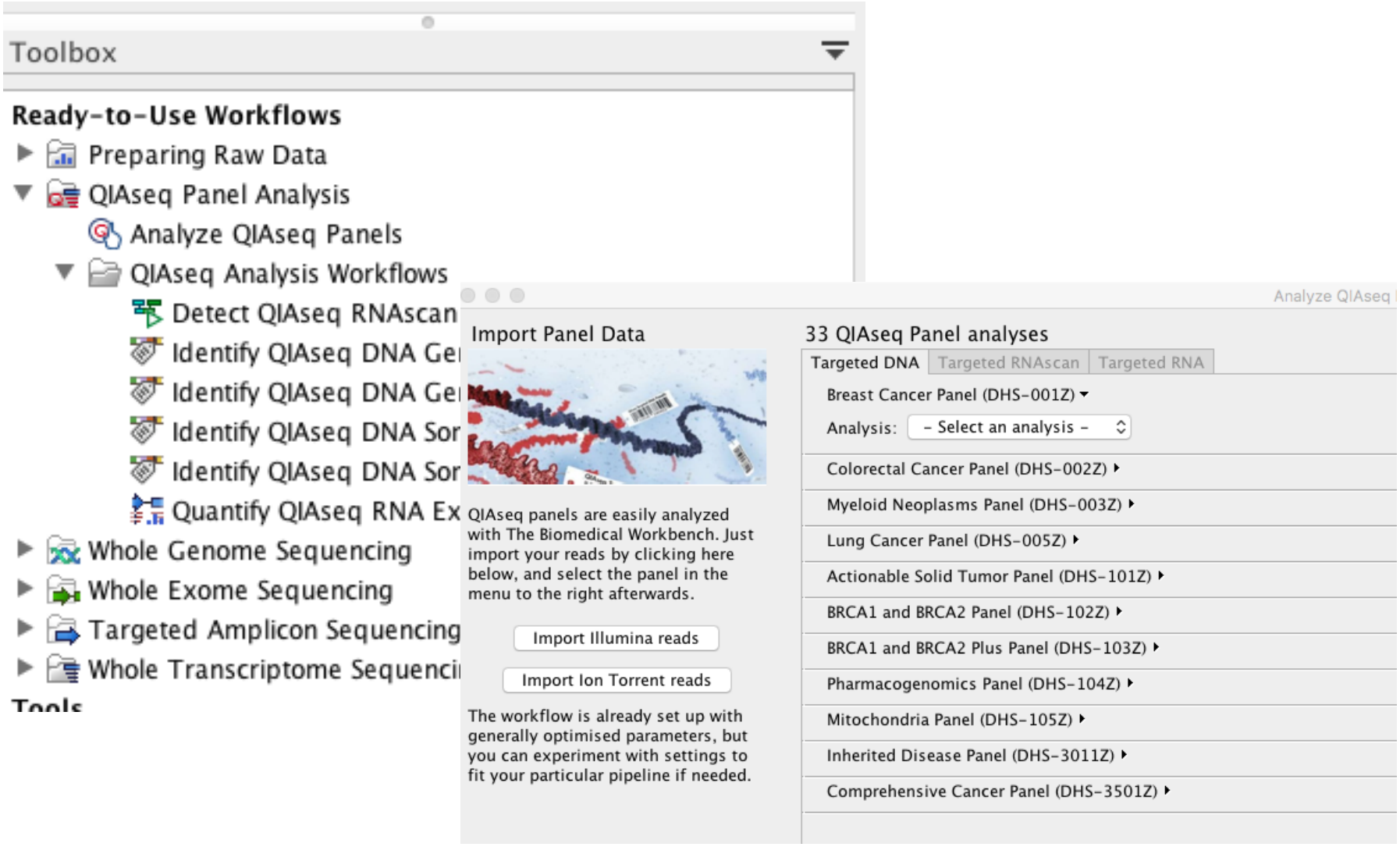
New: CLC file compression

- All new files generated by upgraded CLC GxWB 12 will be compressed
- - Saving 50% disk space

Use case	Output before compression	Output after compression
Identify Known Variants in One Sample workflow (Whole Exome Sequencing data)	16.0 GB	7.7 GB
De novo assembly of NGS reads	6.1 GB	2.1 GB
Imported read mapping	3.98 GB	1.77 GB
Imported NGS reads	9.71 GB	4.69 GB

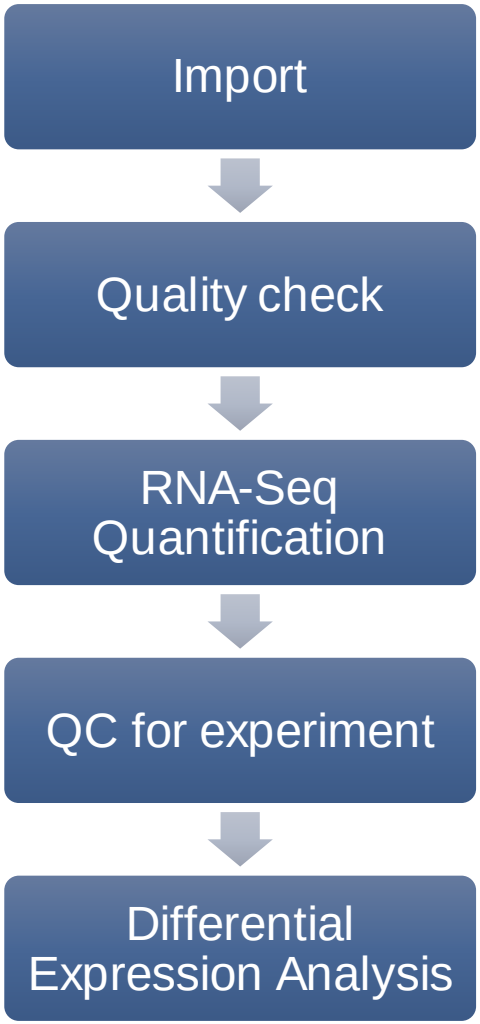
Copyright©2018 GGA Corp. All rights reserved.

QIAseq Support



Copyright©2018 GGA Corp. All rights reserved.

Expression Analysis using RNA-Seq



Copyright©2018 GGA Corp. All rights reserved.

Expression Analysis using RNA-Seq

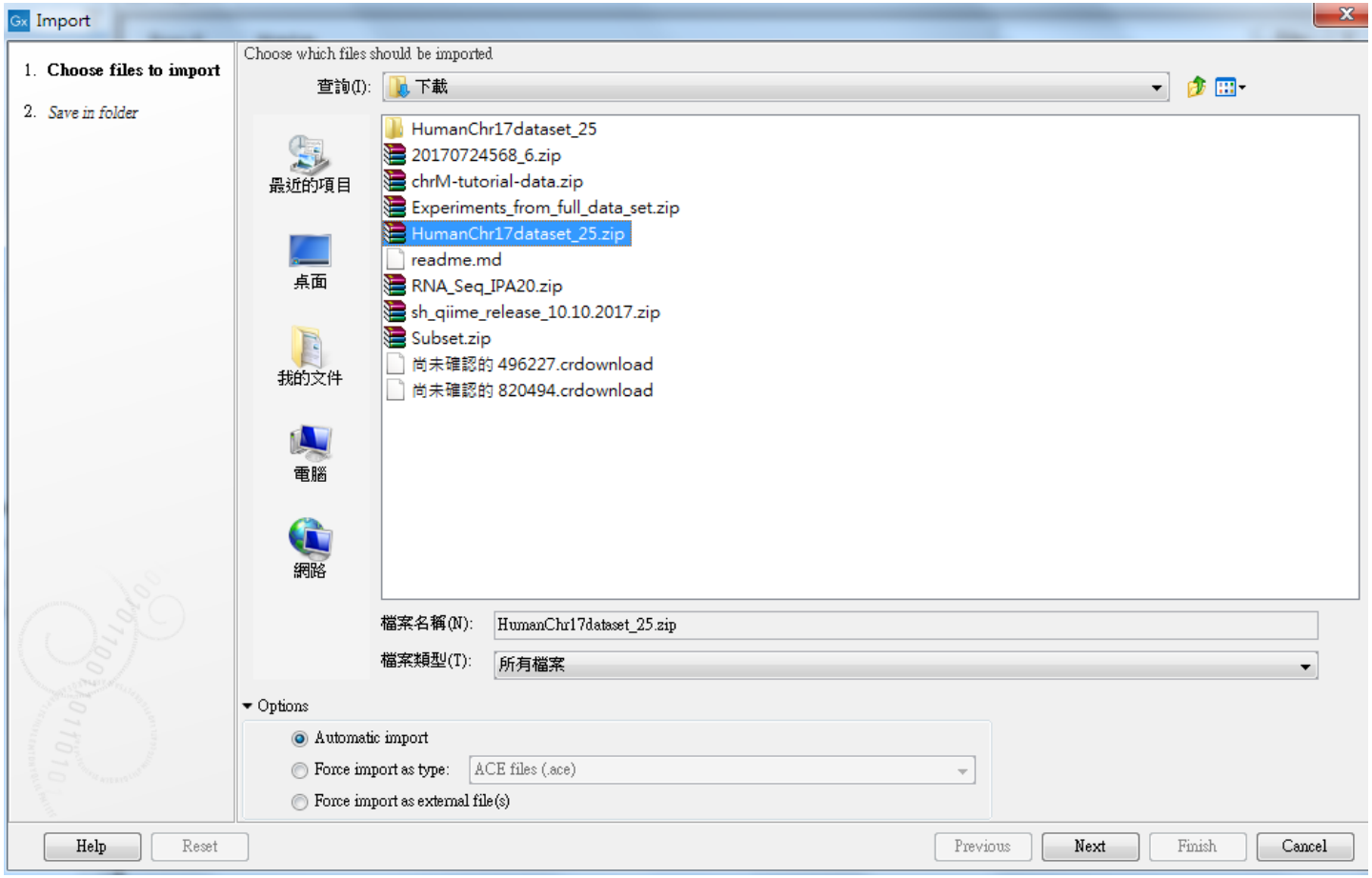
- Example data:
  - Import:
    - Study – SRP032789 (http://www.ebi.ac.uk/ena/data/view/SRP032789)
    - RNA-seq of breast cancer subtype and normal tissue – Chr17 only
    - Use **Standard Import** to import .clc file
    - normal samples –TNBC x 3
    - breast cancer samples –HER2 x 3
- In this section:
  - RNA-Seq Analysis
  - Differential Expression

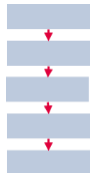
http://resources.qiagenbioinformatics.com/testdata/HumanChr17dataset_25.zip

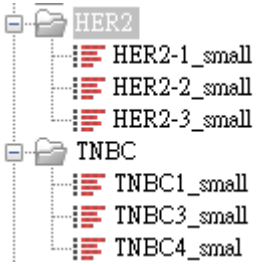
Copyright©2018 GGA Corp. All rights reserved.

RNA-Seq Analysis – Breast Cancer
Import Example data (.clc file)

- **Import > Standard Import**





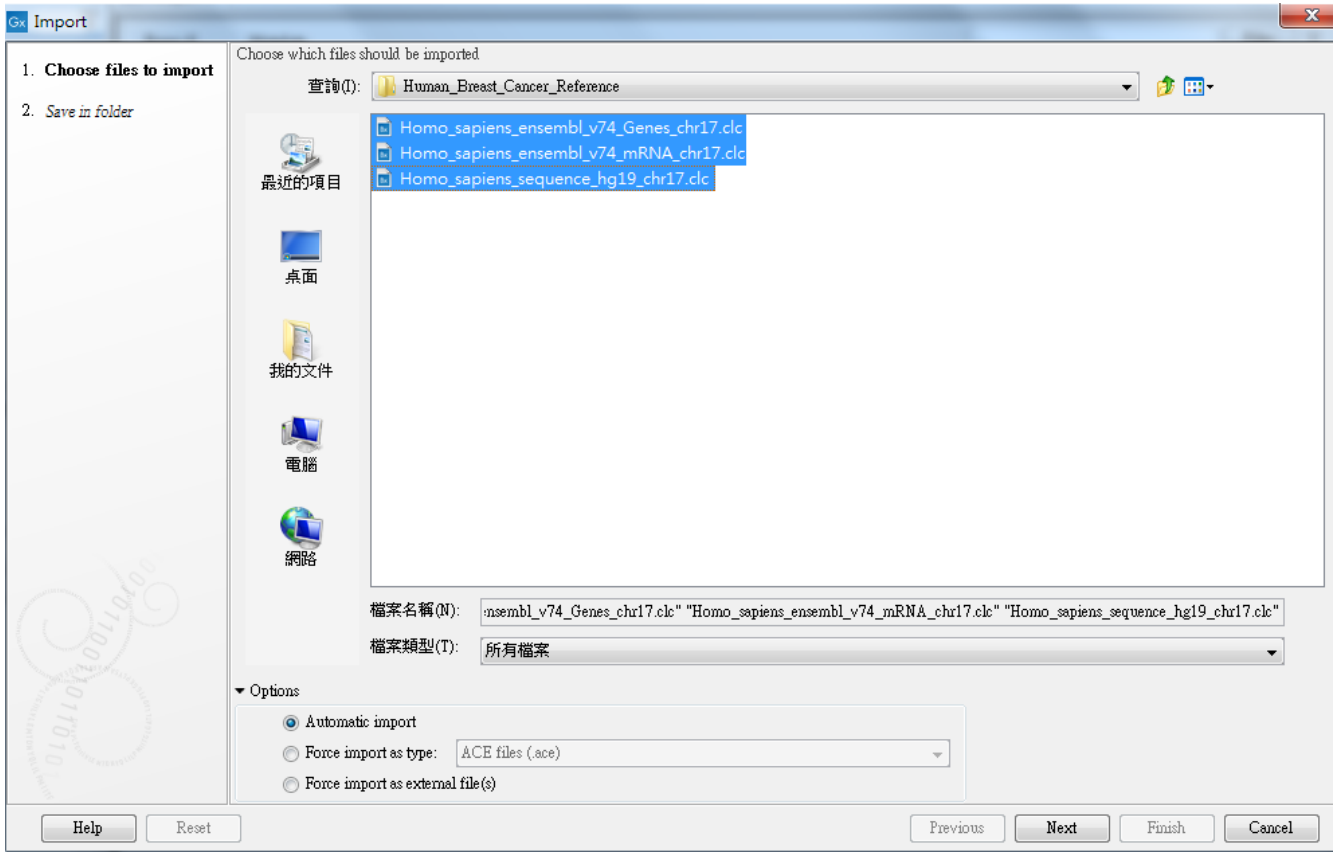


Copyright©2018 GGA Corp. All rights reserved.

RNA-Seq Analysis – Breast Cancer Import Reference data (.clc file)

Homo_sapiens_ensembl_v74_Genes_chr17.clc	2018/3/30 上午 1...	CLC Files	158 KB
Homo_sapiens_ensembl_v74_mRNA_chr17.clc	2018/3/30 上午 1...	CLC Files	4,086 KB
Homo_sapiens_sequence_hg19_chr17.clc	2018/3/30 上午 1...	CLC Files	19,024 KB

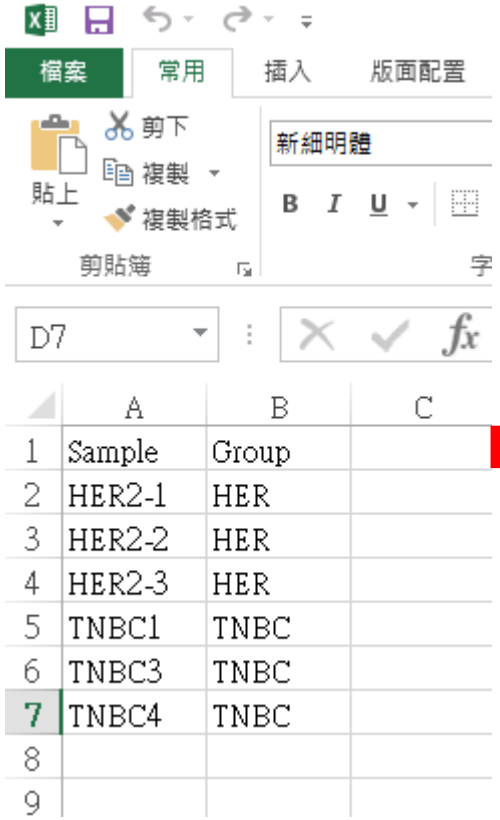
Original from CLC-Bx



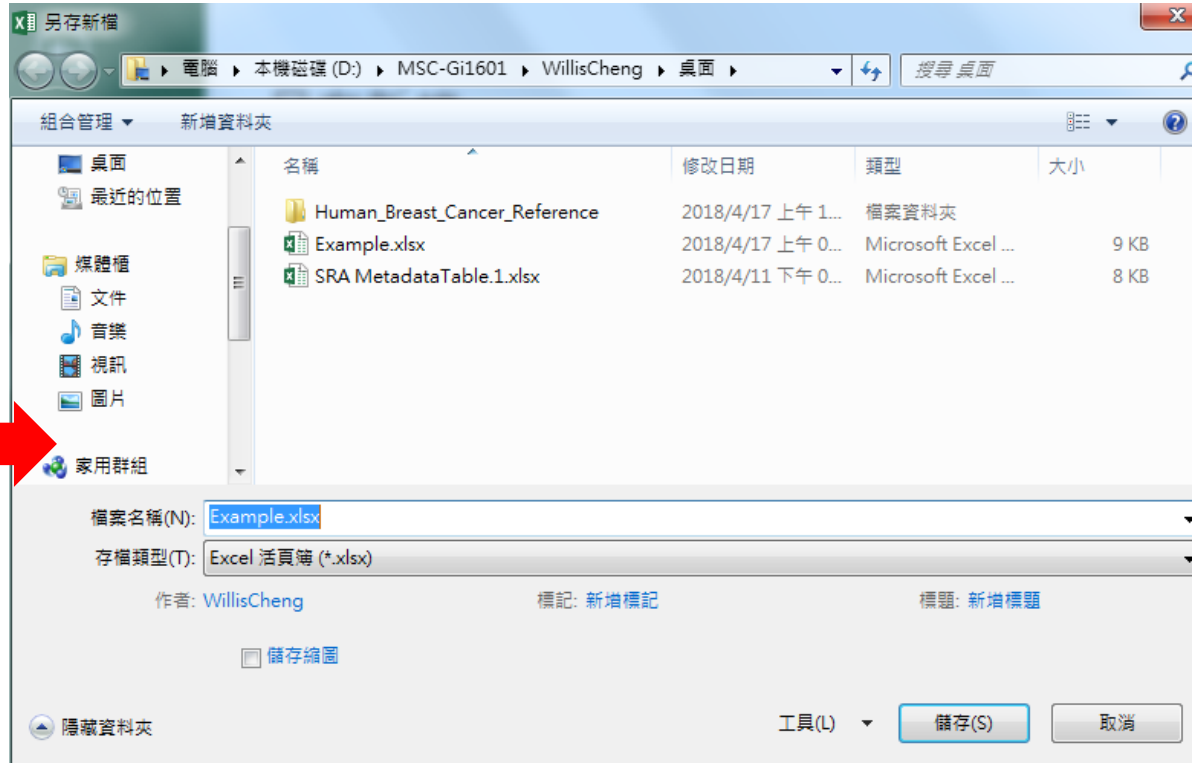
Copyright©2018 GGA Corp. All rights reserved.

RNA-Seq Analysis – Breast Cancer Build Metadata

- Use EXCEL, and save as xlsx file



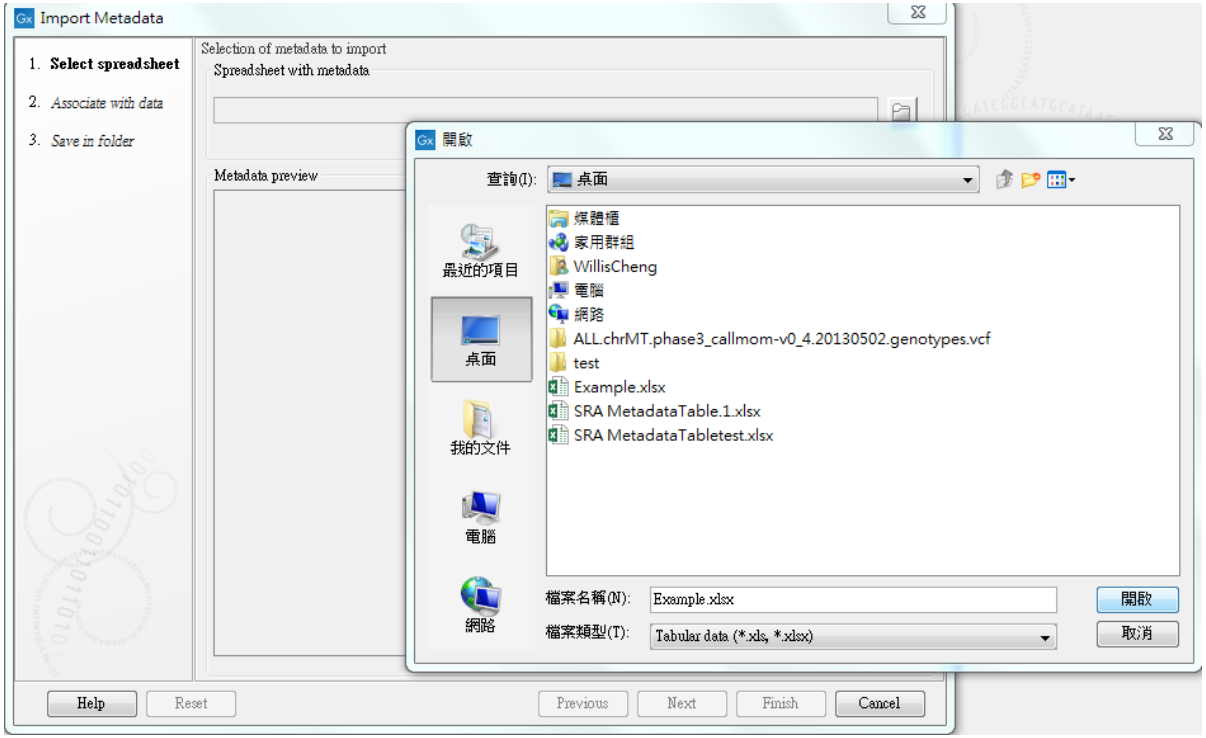
	A	B	C
1	Sample	Group	
2	HER2-1	HER	
3	HER2-2	HER	
4	HER2-3	HER	
5	TNBC1	TNBC	
6	TNBC3	TNBC	
7	TNBC4	TNBC	
8			
9			



Copyright©2018 GGA Corp. All rights reserved.

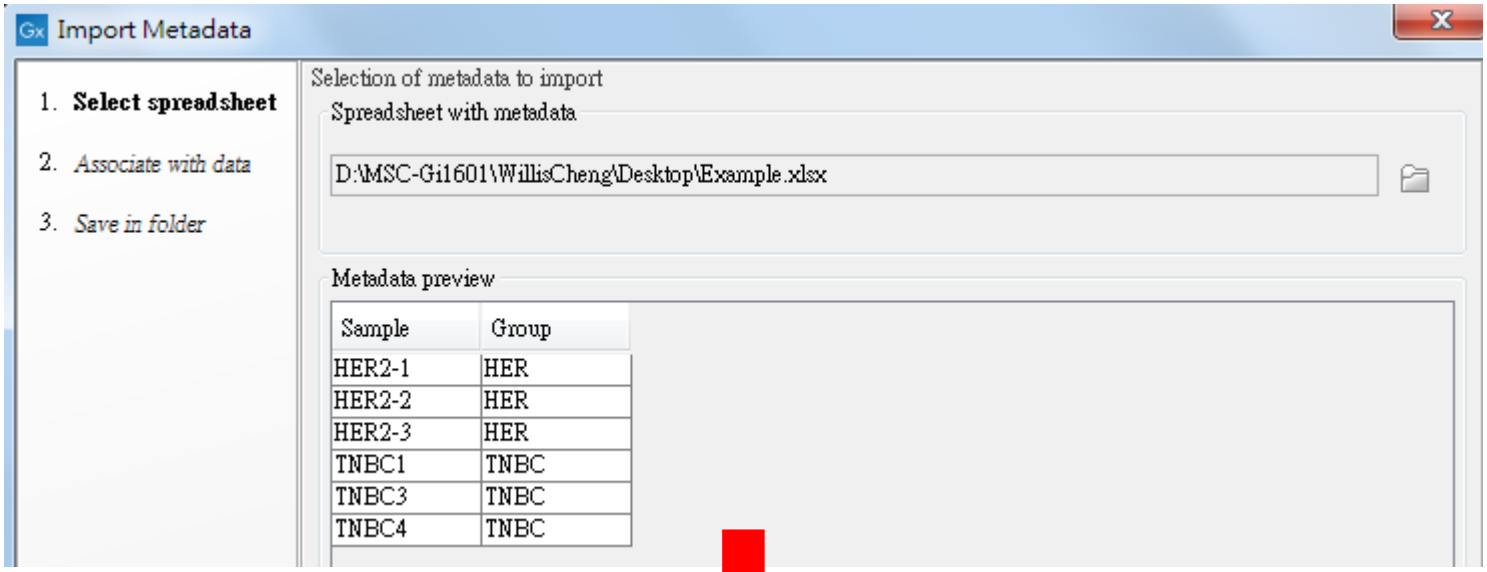
RNA-Seq Analysis – Breast Cancer Import Metadata

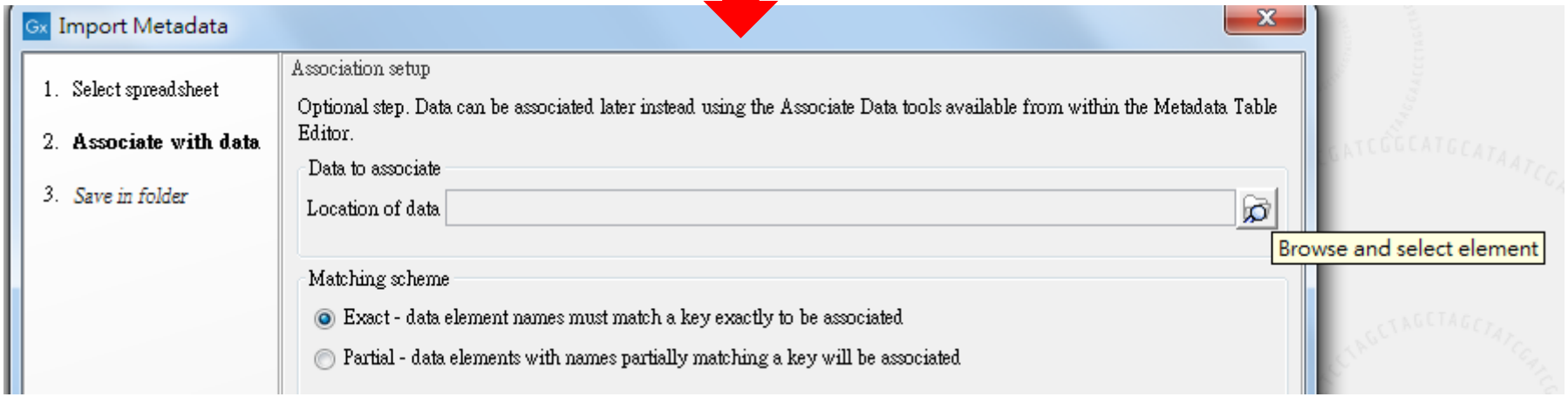
- **Import > Import Metadata**
  - Choose Example.xlsx
  - Link associated ID to fastq files



Copyright©2018 GGA Corp. All rights reserved.

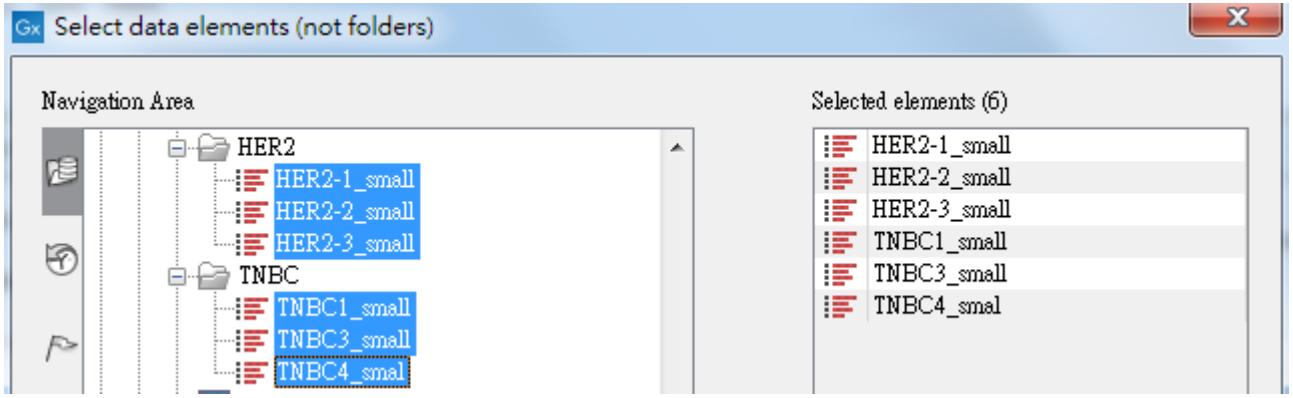
RNA-Seq Analysis – Breast Cancer
Import Metadata

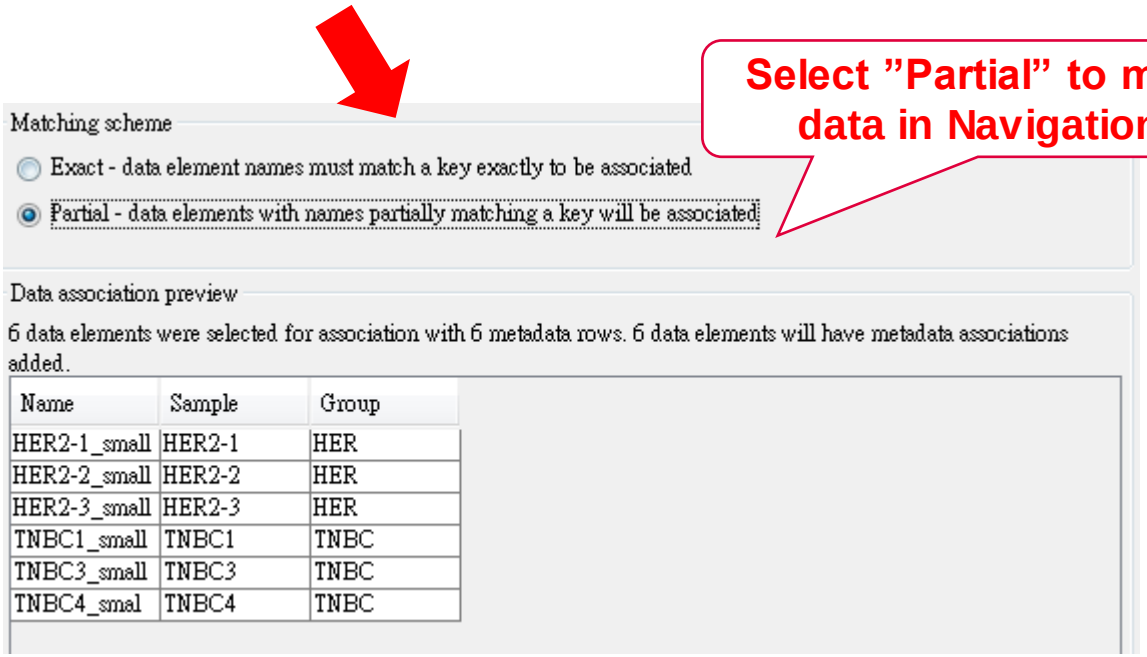




Copyright©2018 GGA Corp. All rights reserved.

RNA-Seq Analysis – Breast Cancer Import Metadata

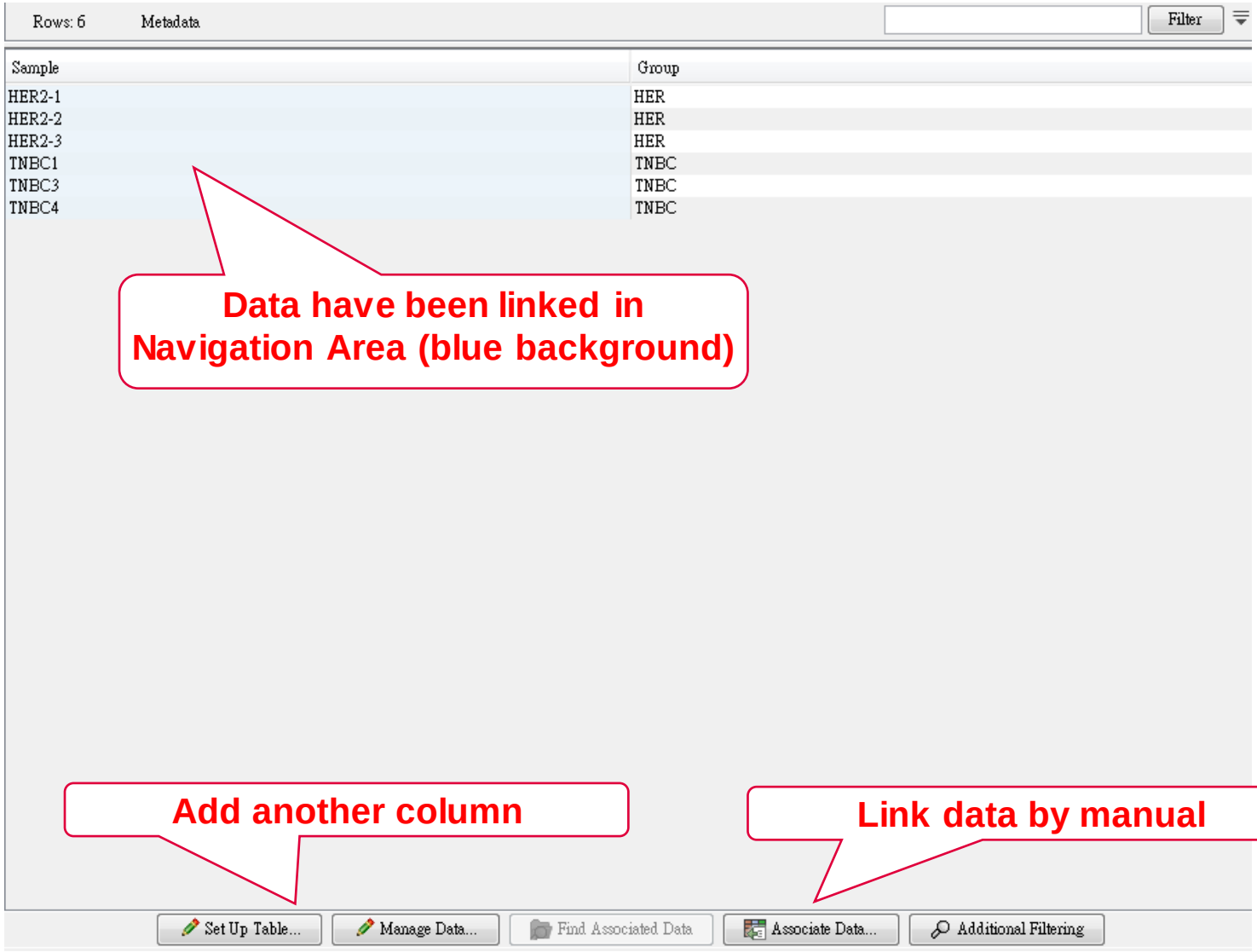




Select "Partial" to match the data in Navigation Area

Copyright©2018 GGA Corp. All rights reserved.

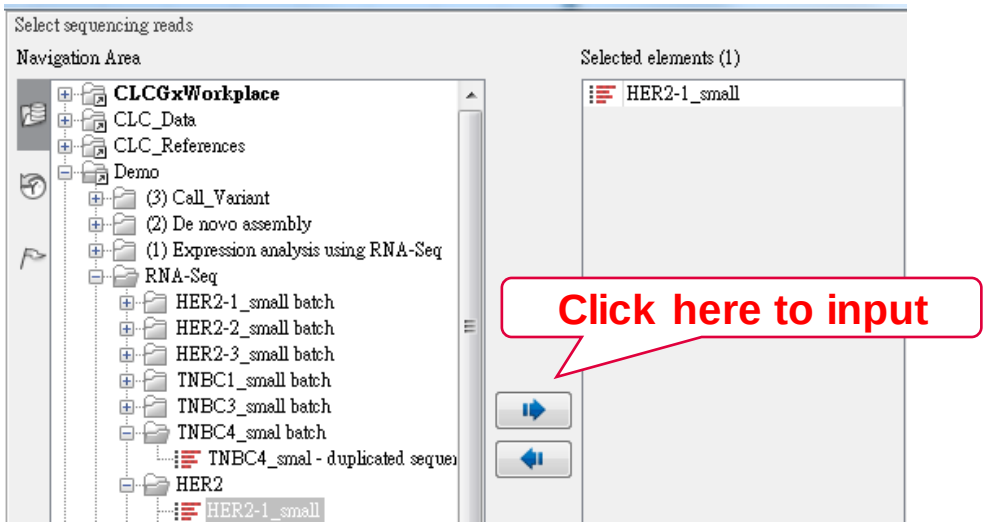
RNA-Seq Analysis – Breast Cancer Metadata

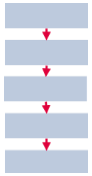


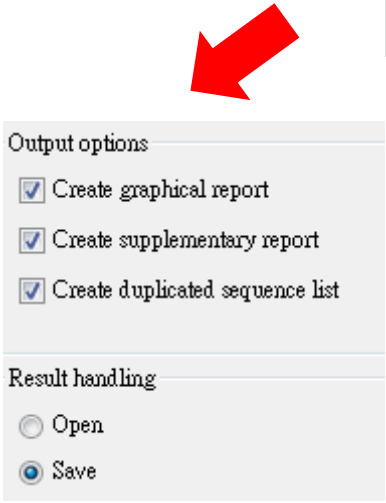
Copyright©2018 GGA Corp. All rights reserved.

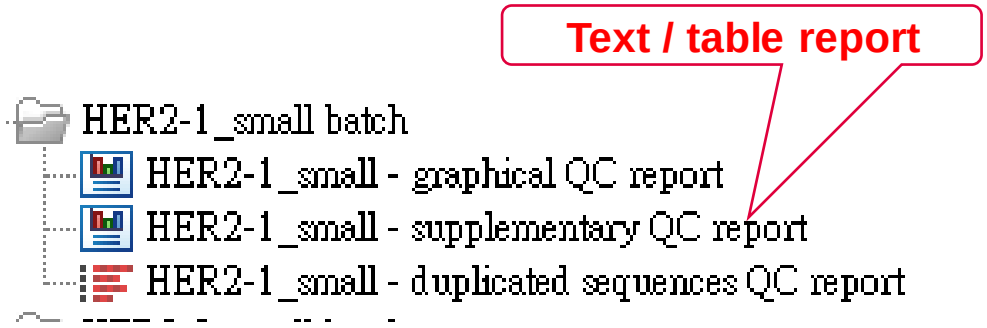
RNA-Seq Analysis - Breast Cancer Quality Control

- **Toolbox > NGS Core Tools > Create Sequencing QC Report**





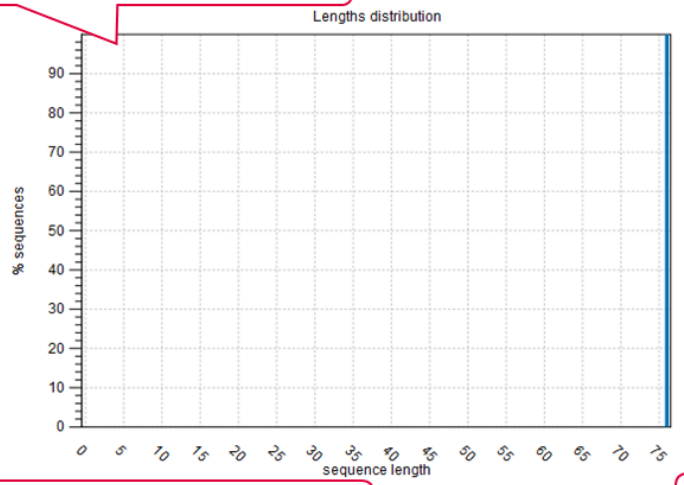




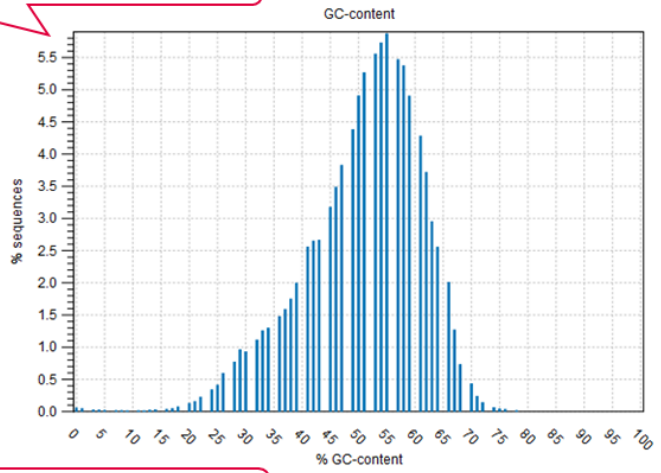
Copyright©2018 GGA Corp. All rights reserved.

RNA-Seq Analysis - Breast Cancer Quality Control

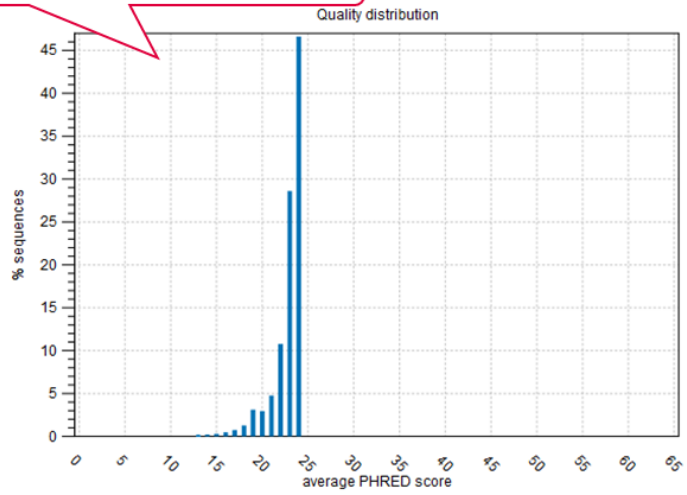
Lengths distribution



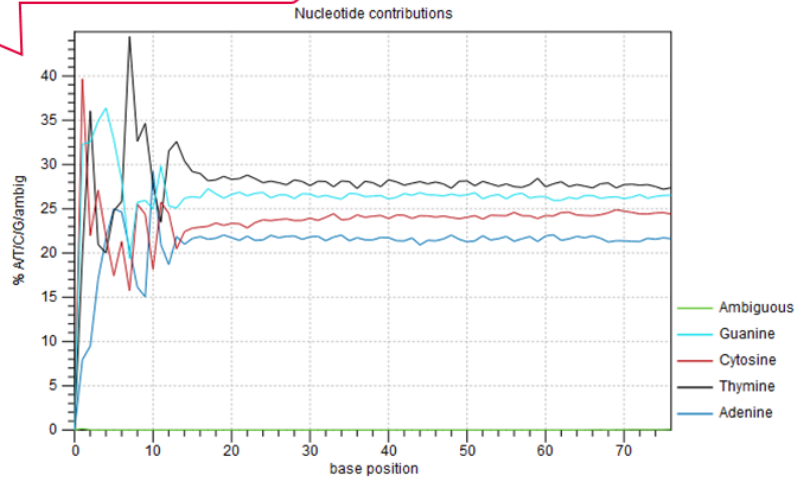
GC-Content



Quality distribution



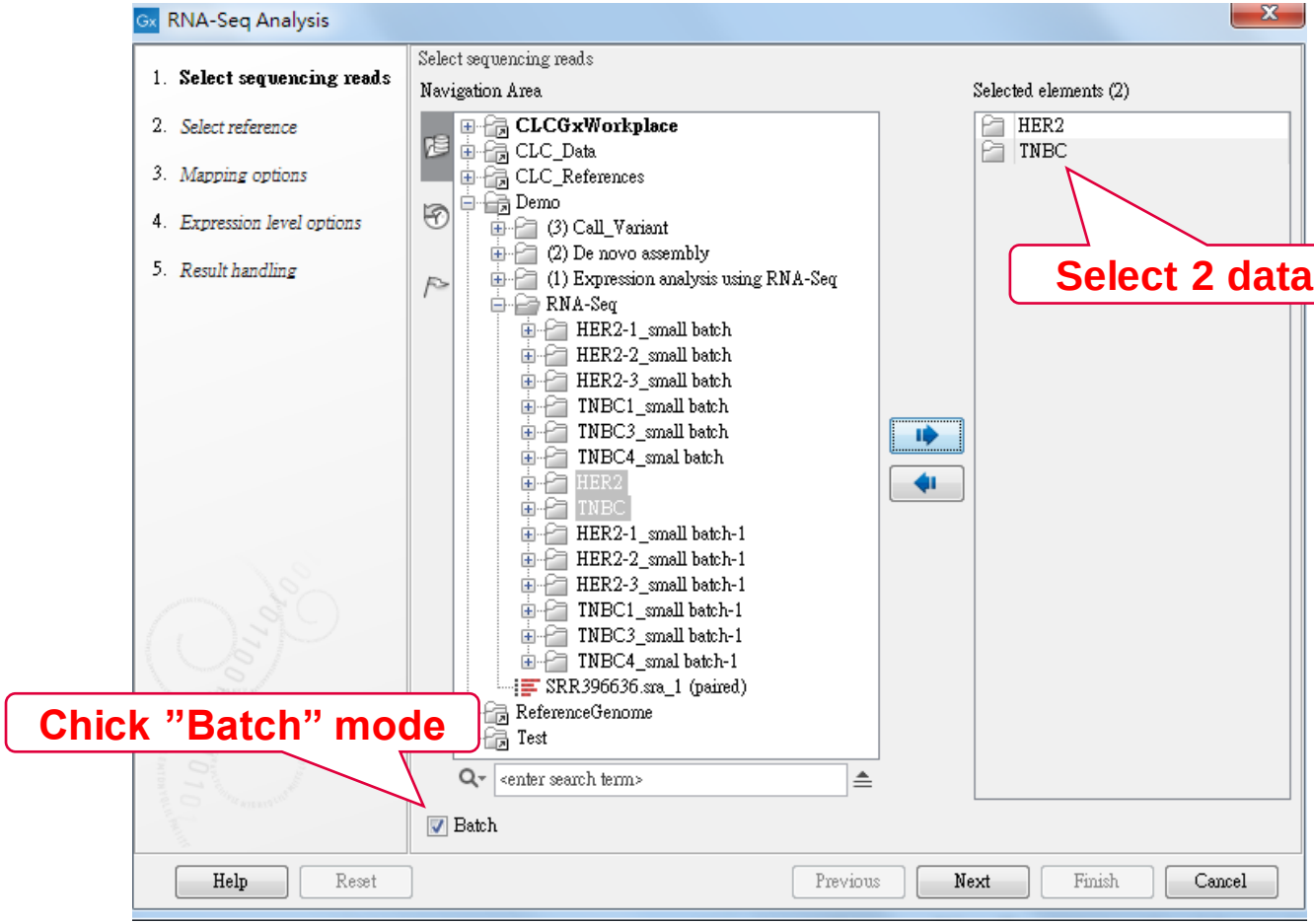
Nucleotide Contribution



Copyright©2018 GGA Corp. All rights reserved.

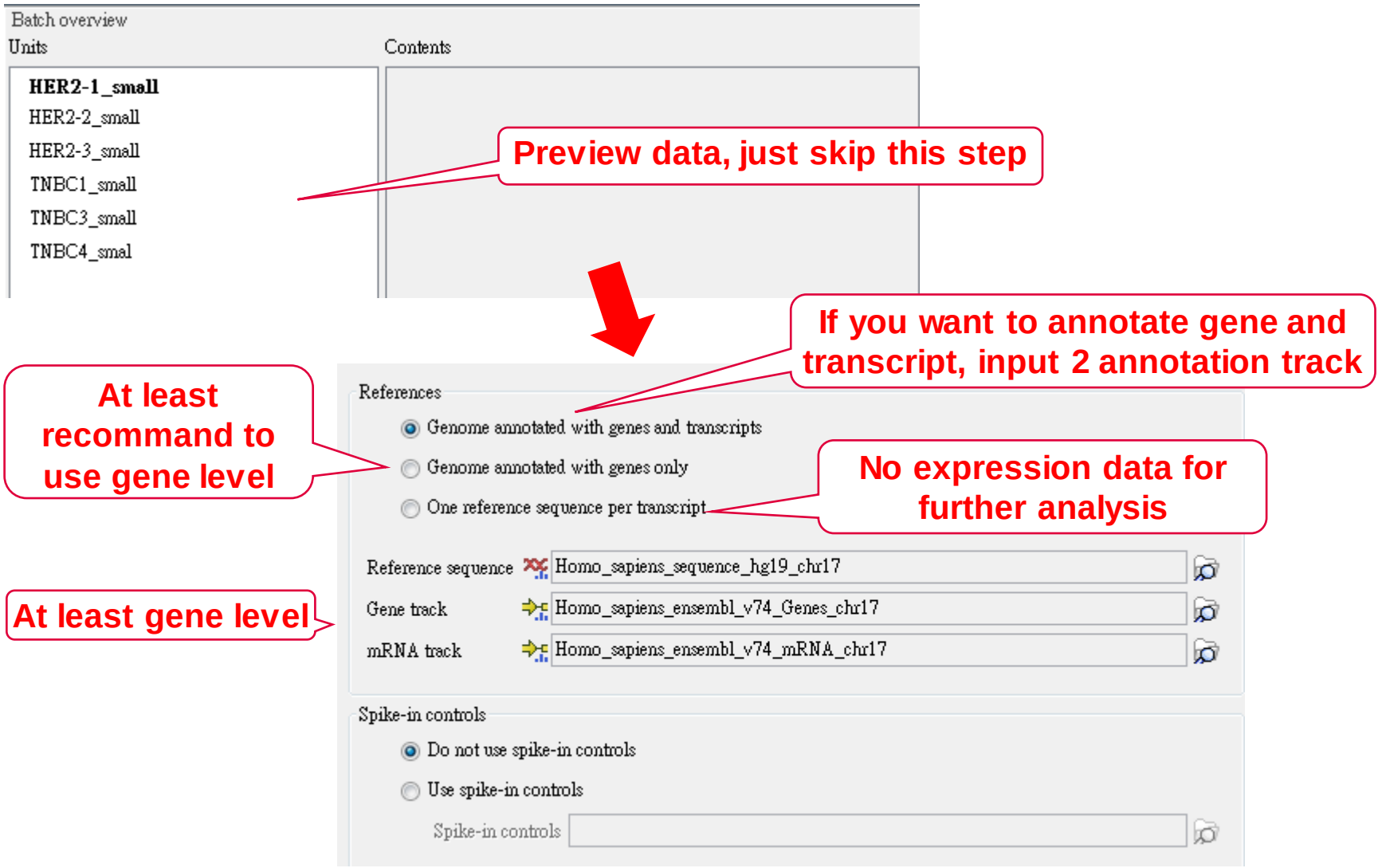
RNA-Seq Analysis – Breast Cancer RNA-Seq Quantification

- **Toolbox > RNA-Seq Analysis > RNA-Seq Analysis**



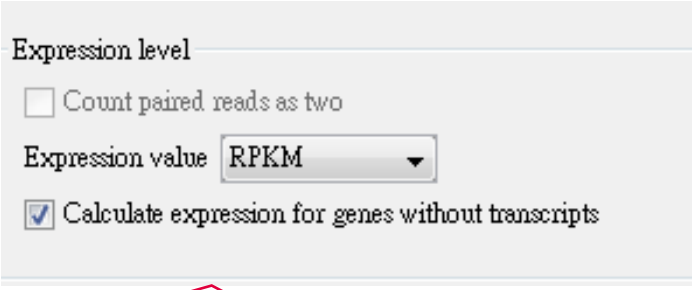
Copyright©2018 GGA Corp. All rights reserved.

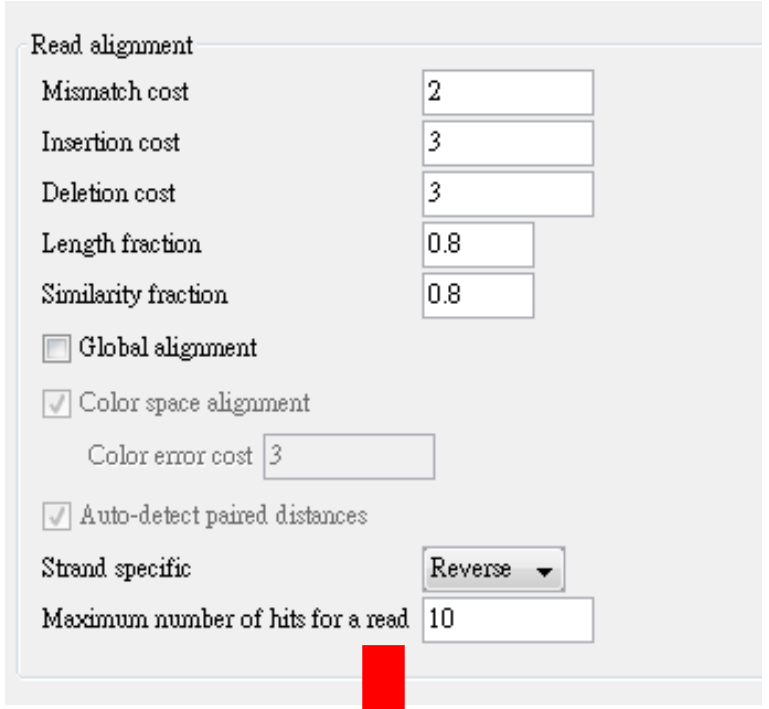
RNA-Seq Analysis – Breast Cancer RNA-Seq Quantification



Copyright©2018 GGA Corp. All rights reserved.

RNA-Seq Analysis – Breast Cancer RNA-Seq Quantification



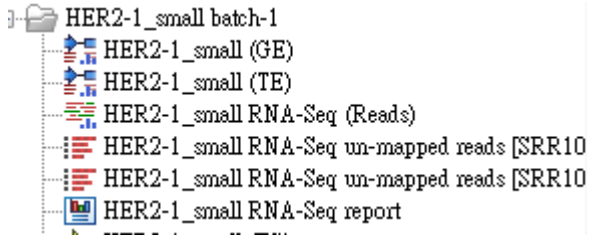


TPM:
Transcript **P**er **M**illion

RPKM:
Reads **P**er **K**ilobase per **M**illion

Unique counts: only reads that are **uniquely mapped** are counted

Total reads: total number of reads mapped to the gene.



Copyright©2018 GGA Corp. All rights reserved.

RNA-Seq Analysis – Breast Cancer
RNA-Seq Quantification

GE data (Gene level)

Name	TPM	Exons	Gene ID	ENSEMBL	Transcripts ...	Detected tra...	Exon length	Unique exo...	Total exon r...
SNORA21_2	467,791.48	1	ENSG00000199293	ENSG00000199293	1	0	132	0	147
RPL19	114,181.04	18	ENSG00000108298	ENSG00000108298	7	4	2031	5547	554
RPL23	78,054.92	24	ENSG00000125691	ENSG00000125691	10	8	5656	10503	1056
PSMB3	48,792.20	14	ENSG00000108294	ENSG00000108294	6	4	1179	1376	137
ERBB2	41,174.79	76	ENSG00000141736	ENSG00000141736	24	15	10321	10165	1016
LASP1	21,587.19	22	ENSG00000002834	ENSG00000002834	9	5	7061	3646	364
MIEN1	18,429.78	10	ENSG00000141741	ENSG00000141741	5	1	1200	529	52
MED1	18,199.88	25	ENSG00000125686	ENSG00000125686	5	4	7330	3191	319

TE data (Transcript level)

Name	TPM	Gene name	Transcrip...	Exons	Gene ID	Transcript ID	ENSEMBL	Unique tr...	Total tran...
SNORA21_2_1	227,651.85	SNORA21_2	132	1	ENSG00000199293		ENSG00000199293	0	1477
RPL23_2	183,523.13	RPL23	588	5	ENSG00000125691	ENST00000245857	ENSG00000125691,	0	5304
RPL23_8	130,730.92	RPL23	625	6	ENSG00000125691	ENST00000394332	ENSG00000125691,	3	4016
RPL19_6	111,676.93	RPL19	732	6	ENSG00000108298	ENST00000225430	ENSG00000108298,	28	4018
RPL19_3	42,695.70	RPL19	690	6	ENSG00000108298	ENST00000579374	ENSG00000108298,	0	1448
PSMB3_5	35,031.86	PSMB3	773	6	ENSG00000108294	ENST00000225426	ENSG00000108294,	107	1331
ERBB2_18	25,210.21	ERBB2	4341	27	ENSG00000141736	ENST00000541774	ENSG00000141736,	0	5379
LASP1_6	17,666.61	LASP1	4109	7	ENSG00000002834	ENST00000318008	ENSG00000002834,	488	3568

Copyright©2018 GGA Corp. All rights reserved.

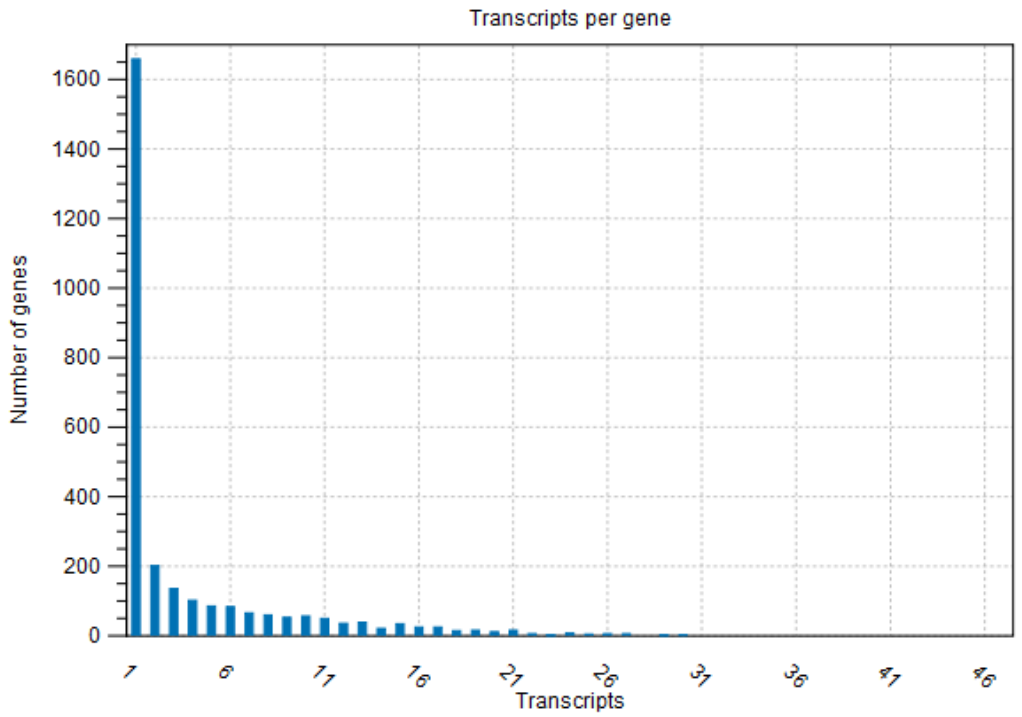
RNA-Seq Analysis – Breast Cancer RNA-Seq Quantification

1.2 Reference Sequences

References	Length	Genes	Transcripts
1	81,195,210	2,903	12,358

2 References

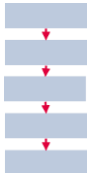
2.1 Transcripts per gene



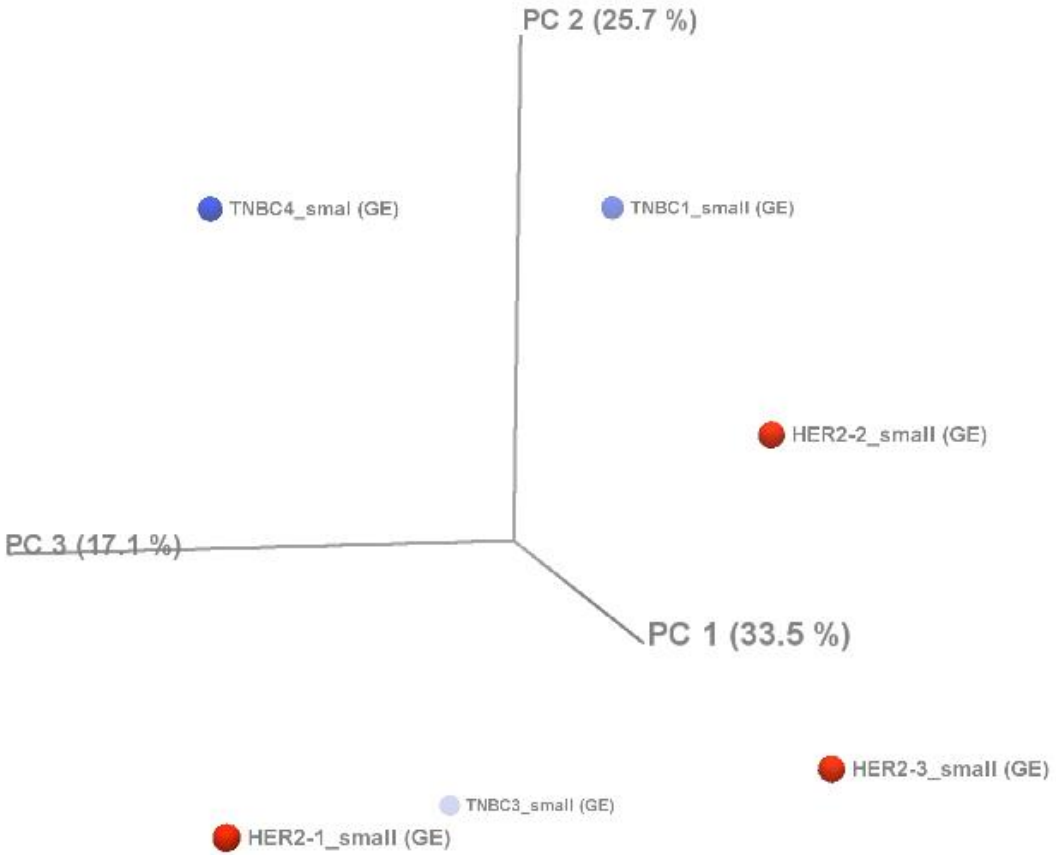
Copyright©2018 GGA Corp. All rights reserved.

RNA-Seq Analysis – Breast Cancer PCA Plot

- **Toolbox** > **RNA-Seq Analysis** > **PCA for RNA-Seq**
- Input : All GE data from RNA-Seq Analysis



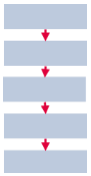
PCA



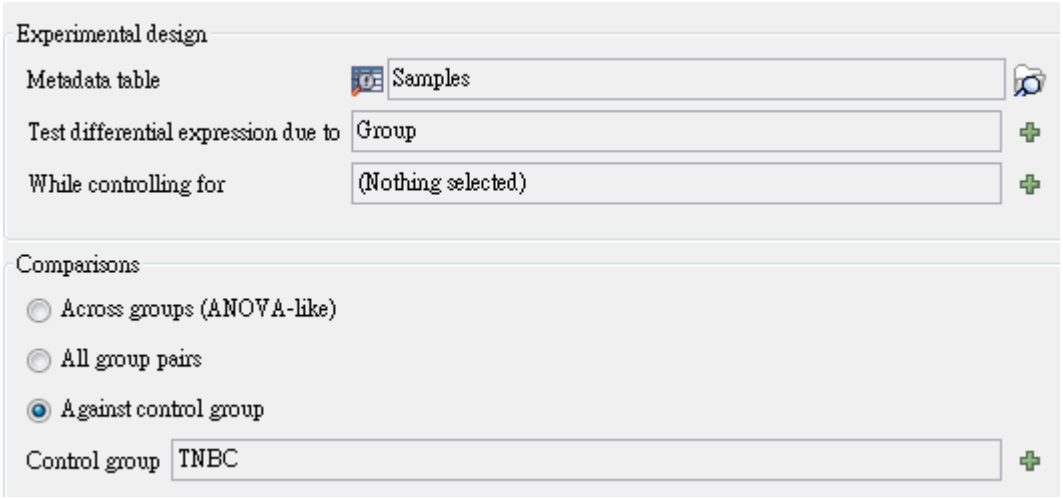
Copyright©2018 GGA Corp. All rights reserved.

RNA-Seq Analysis – Breast Cancer Differential Expression Analysis

- **Toolbox > RNA-Seq Analysis > Differential Expression for RNA-Seq**
  - Input / Parameters:



Differential expression



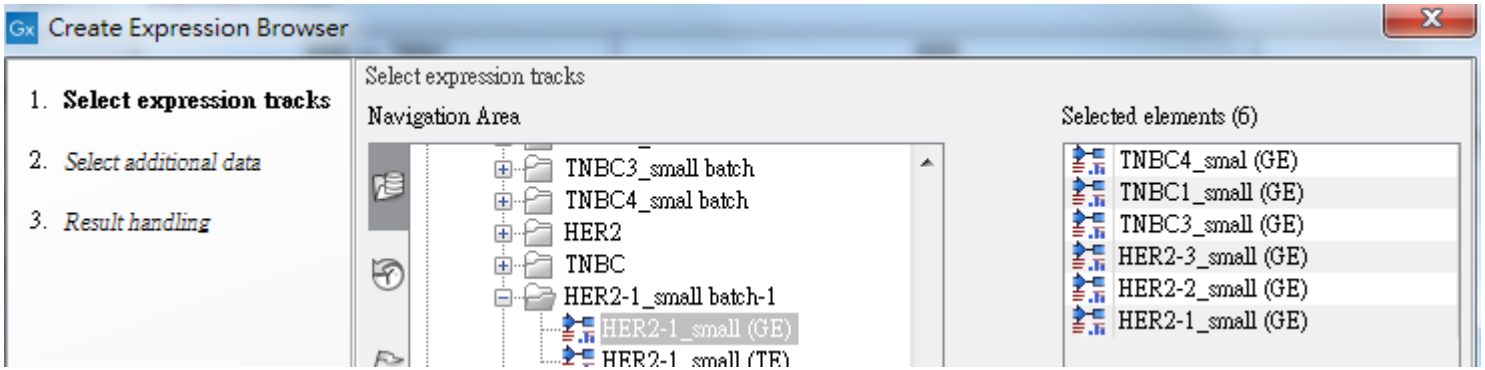
P-value for differential expressed gene

Name	Chromosome	Region	Max group m...	Log2 fold ch...	Fold change	P-value	FDR p-value	Bonferroni
ERBB2	17	37844167..37886679	24,086.35	4.31	19.77	2.43E-9	7.04E-6	7.04E-6
SNORA21_2	17	complement(37009116..37009247)	56,904.43	6.16	71.33	9.86E-7	1.43E-3	2.86E-3
STARD3	17	37793318..37819737	3,452.13	2.82	7.08	8.60E-6	8.32E-3	0.02
CTB-131K11.1	17	37558046..37562486	3,584.05	2.04	4.10	1.68E-4	0.12	0.49
FBXL20	17	complement(37415384..37558776)	779.28	1.61	3.05	5.33E-4	0.27	1.00
PGAP3	17	complement(37827375..37853050)	3,046.74	2.36	5.14	5.87E-4	0.27	1.00
GRB7	17	37894180..37903544	4,824.23	3.79	13.80	6.47E-4	0.27	1.00

Copyright©2018 GGA Corp. All rights reserved.

RNA-Seq Analysis – Breast Cancer Create Expression Table

- **Toolbox > RNA-Seq Analysis > Create Expression Browser**





Copyright©2018 GGA Corp. All rights reserved.

RNA-Seq Analysis – Breast Cancer Create Expression Table

Differential Expression

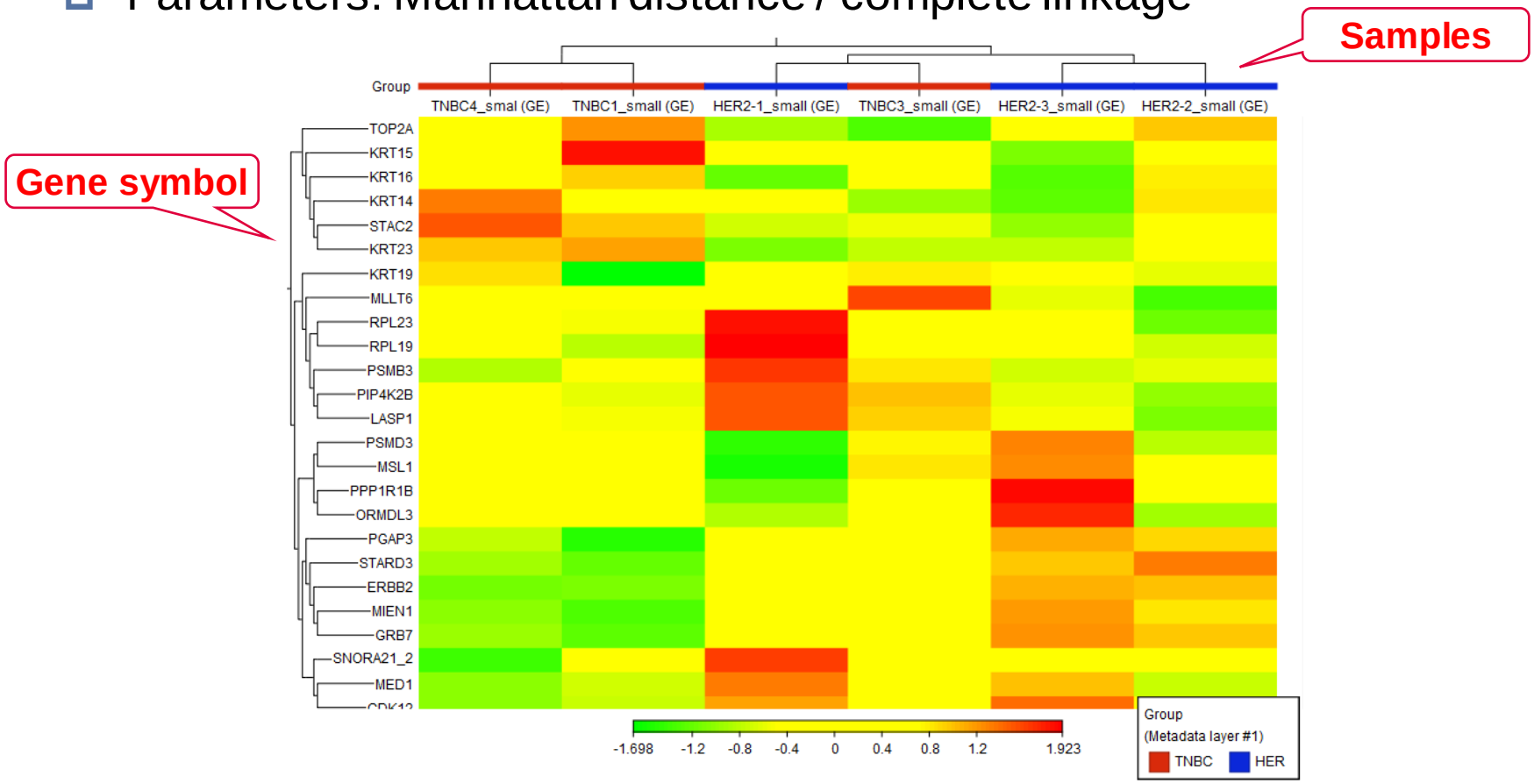
Raw coverage

Name	HER vs. TNBC			HER				TNBC			
	Log fold change	P-val...	FDR p-value	HER2-1_small... RPKM	HER2-2_sma... RPKM	HER2-3_sma... RPKM	Mean	TNBC1_smal... RPKM	TNBC3_smal... RPKM	TNBC4_smal... RPKM	Mean
ERBB2	4.31	2.43E-9	7.04E-6	13,849.97	32,241.57	26,167.50	24,086.35	1,542.72	4,101.30	1,395.79	2,346.60
SNORA21_2	6.16	9.86E-7	1.43E-3	157,351.10	3,347.40	10,014.79	56,904.43	1,997.09	2,342.62	0.00	1,446.57
STARD3	2.82	8.60E-6	8.32E-3	1,340.85	6,040.05	2,975.49	3,452.13	638.60	1,295.10	811.59	915.10
CTB-131K11.1	2.04	1.68E-4	0.12	5,412.44	1,321.17	4,018.56	3,584.05	1,348.73	2,097.46	1,363.46	1,603.22
FBXL20	1.61	5.33E-4	0.27	1,134.14	387.13	816.56	779.28	388.03	663.78	359.58	470.46
PGAP3	2.36	5.87E-4	0.27	1,199.66	4,141.92	3,798.62	3,046.74	359.25	2,336.44	787.56	1,161.08
GRB7	3.79	6.47E-4	0.27	2,150.23	6,163.95	6,158.50	4,824.23	373.77	1,093.70	523.91	663.79
KRT15	-4.37	8.83E-4	0.32	246.68	265.65	40.07	184.13	18,563.86	199.28	989.33	6,584.16
STAC2	-4.01	2.15E-3	0.69	46.01	154.90	22.47	74.46	1,772.07	76.65	4,284.94	2,044.55
THRA	-1.73	2.66E-3	0.73	101.04	108.85	100.41	103.43	473.52	834.40	589.74	632.55
KRT42P	-4.65	2.78E-3	0.73	0.00	0.00	4.47	1.49	74.94	118.03	39.69	77.55
KRT9	-2.74	4.44E-3	0.97	0.00	20.70	0.00	6.90	69.16	127.70	91.56	96.14
MIEN1	2.84	6.00E-3	0.97	6,199.23	12,229.95	13,062.14	10,497.11	1,515.79	4,981.96	1,919.40	2,805.72
KRT13	-2.37	7.39E-3	0.97	4.83	16.25	0.00	7.03	63.35	112.05	28.75	68.05
AC003958.2	-5.81	0.01	0.97	0.00	0.00	0.00	0.00	25.93	114.05	51.48	63.82
GSDMA	-2.13	0.02	0.97	18.92	7.08	8.47	11.49	130.03	130.96	18.78	93.26
C17orf96	1.57	0.03	0.97	714.67	542.81	468.18	575.22	490.14	207.62	350.39	349.38
DUSP14	-1.35	0.03	0.97	191.33	479.92	249.38	306.88	1,360.80	1,876.72	1,122.77	1,453.43
CSF3	-4.81	0.03	0.97	0.00	0.00	0.00	0.00	23.75	100.60	37.73	54.03
KRT17	-2.61	0.03	0.97	2,078.61	2,002.90	342.42	1,474.64	6,712.49	832.89	37,173.36	14,906.24
LINC00974	-4.70	0.03	0.97	0.00	0.00	0.00	0.00	49.25	80.24	19.56	49.68
TMEM99	1.32	0.03	0.97	270.55	269.10	156.89	232.18	195.95	157.74	201.42	185.04

Copyright©2018 GGA Corp. All rights reserved.

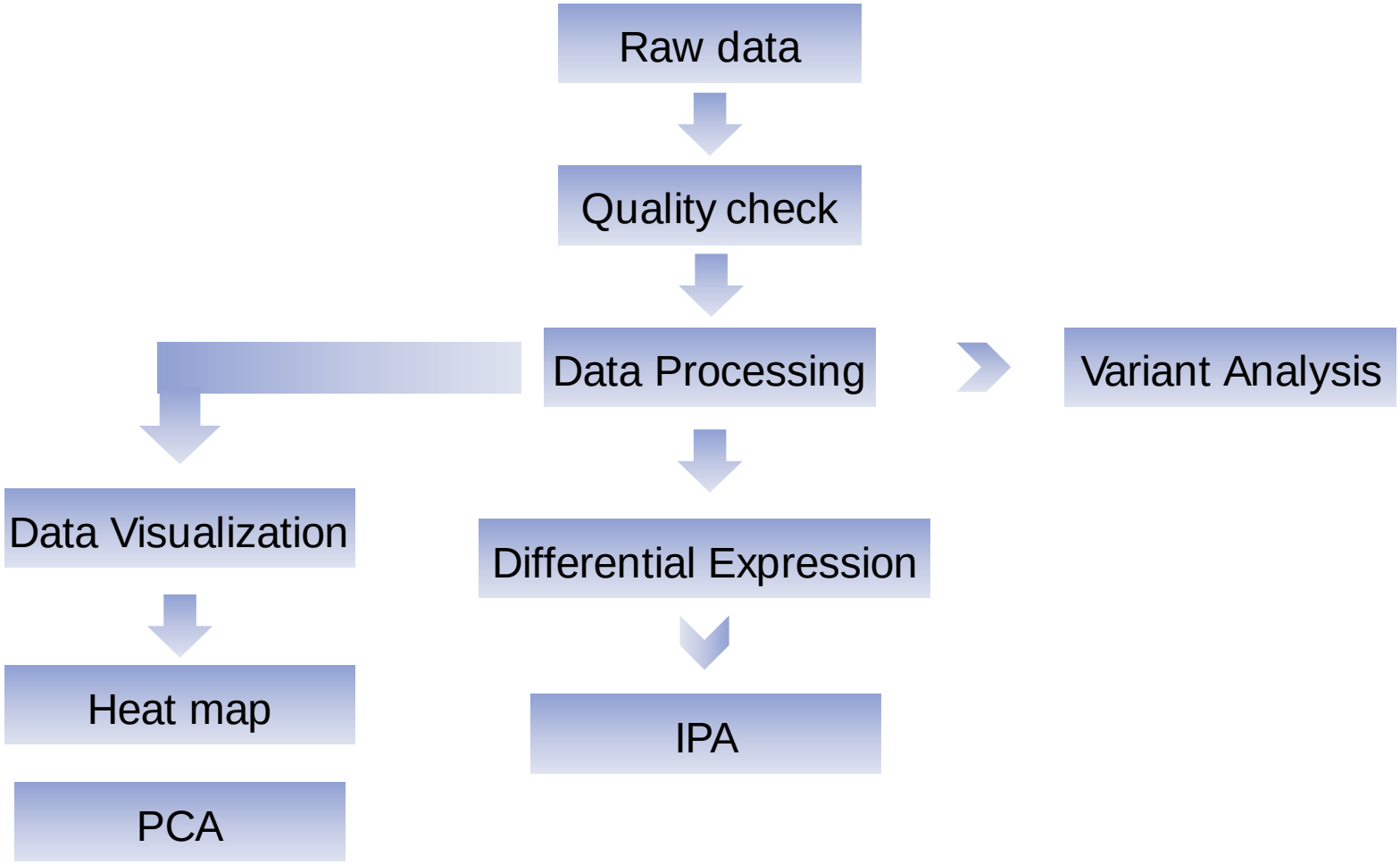
RNA-Seq Analysis – Breast Cancer Heat Map

- **Launch > Create Heat Map for RNA-Seq**
  - Input : GE data
  - Parameters: Manhattan distance / complete linkage



Copyright©2018 GGA Corp. All rights reserved.

Summary of Analysis by RNA-Seq



Copyright©2018 GGA Corp. All rights reserved.

- Analysis of NGS Data
  - Call Variant
- Introduction Workflow in CLC-Gx

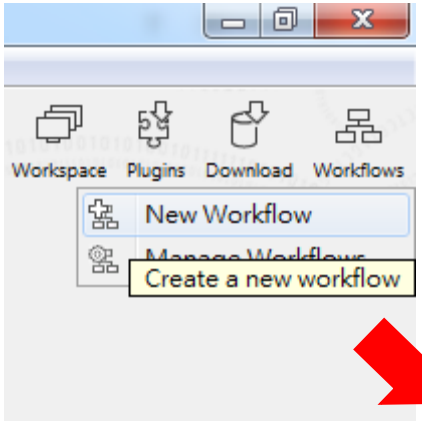
Copyright©2018 GGA Corp. All rights reserved.

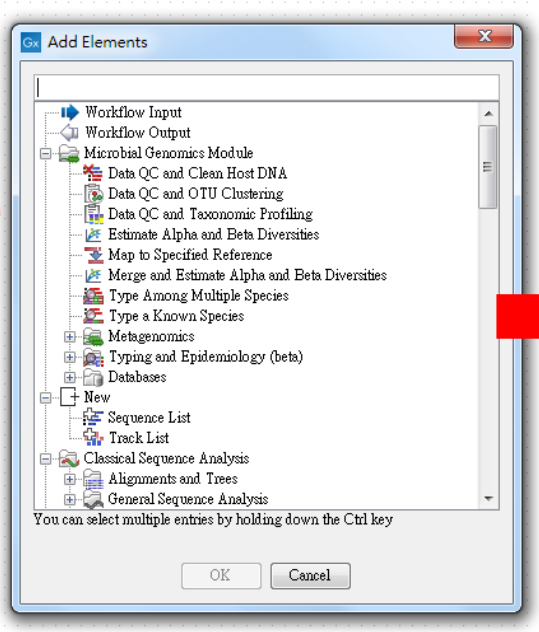
Workflow

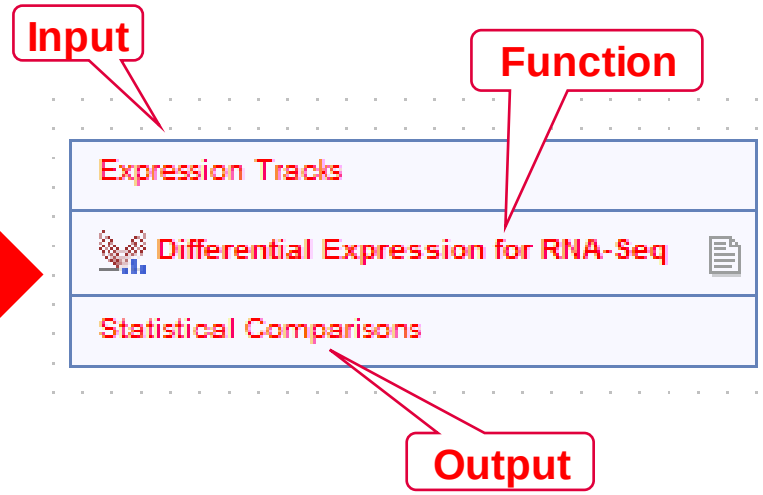
- Advantage:
  - Workflow includes multiple functions
  - Users can build their own pipelines
  - Users can automatically produce data
  - Users can share workflow data with each other
  - Users need to insert input just one time
  - Save time

Copyright©2018 GGA Corp. All rights reserved.

Workflow in CLC-Gx









Copyright©2018 GGA Corp. All rights reserved.

New Workflow Made by Users

- Common Icons:
  - Input data



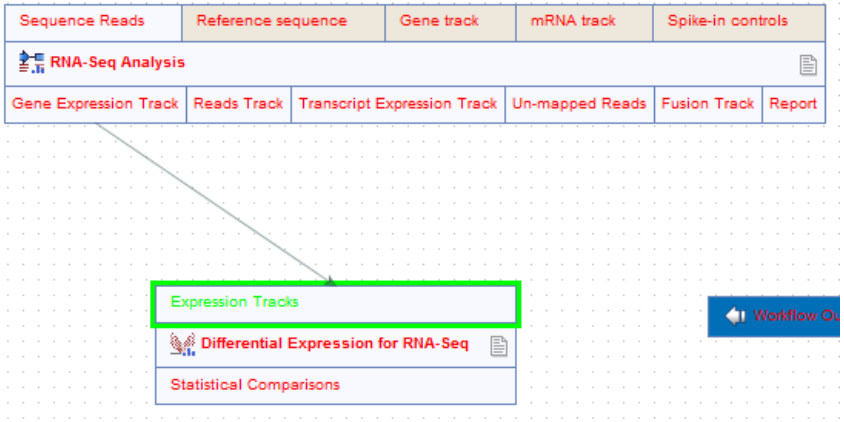
    - Based on input for function
  - Function



  - Export data



    - Base on output from function

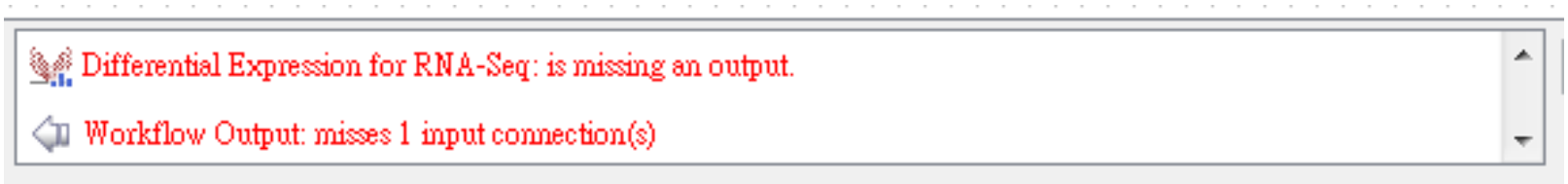


- If the output is the input of another function, we can link the function to build the workflow.

Copyright©2018 GGA Corp. All rights reserved.

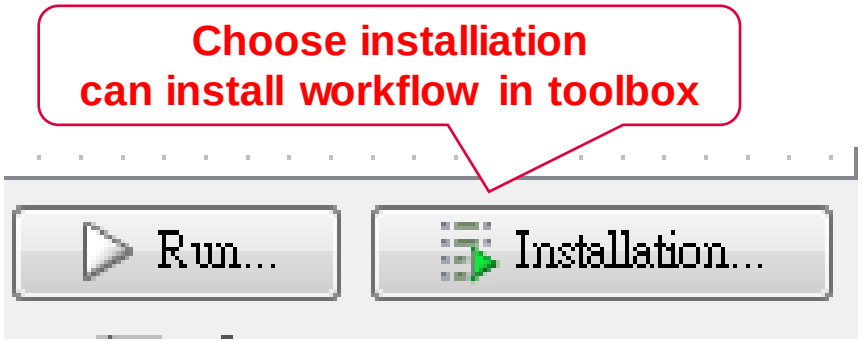
Verify the Workflow

- Lack of connection



- Workflow verified
  - Workflow must be saved

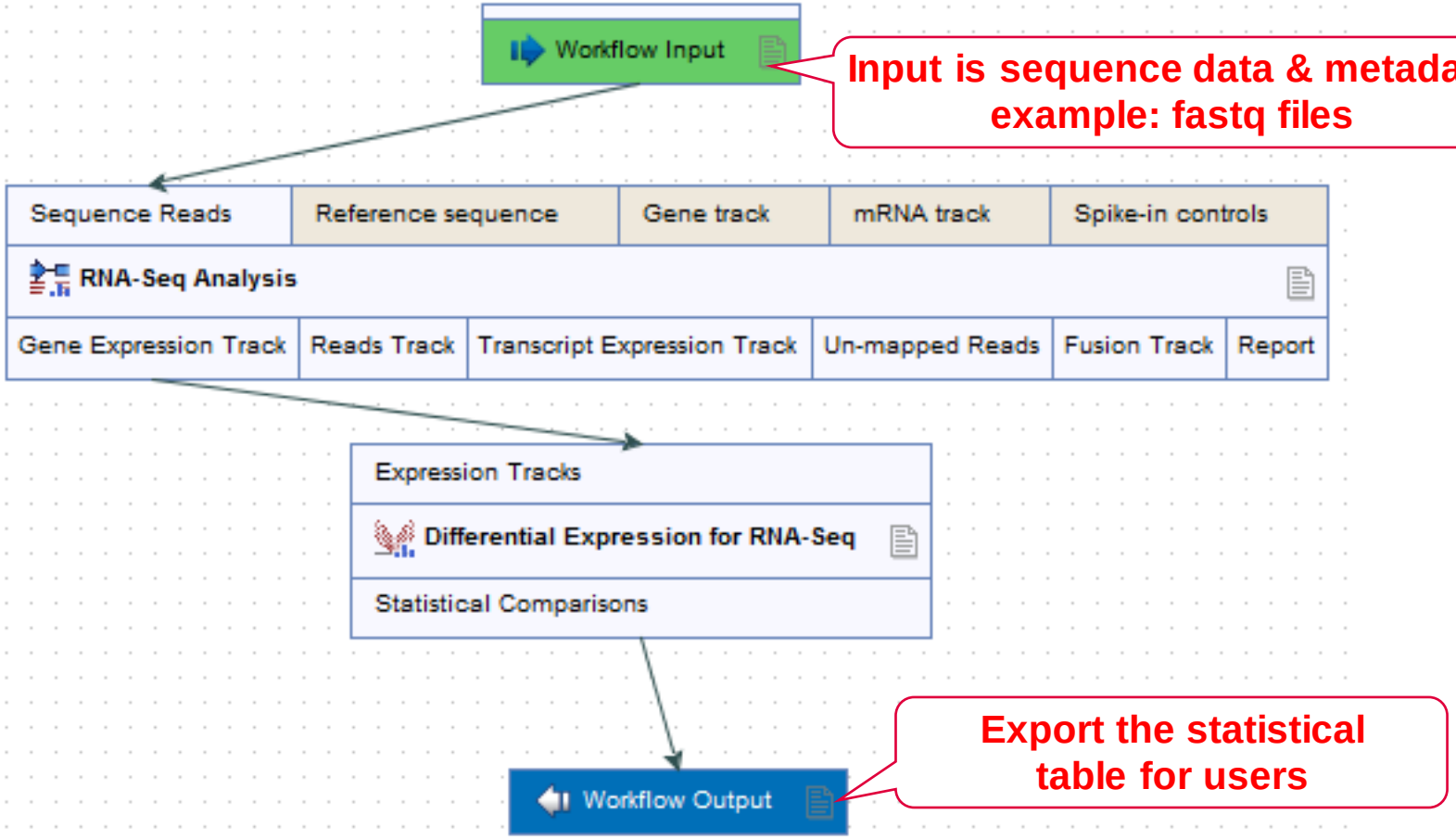




- Run workflow by **Run** icon

Copyright©2018 GGA Corp. All rights reserved.

Example Workflow: Expression by RNA-Seq Analysis

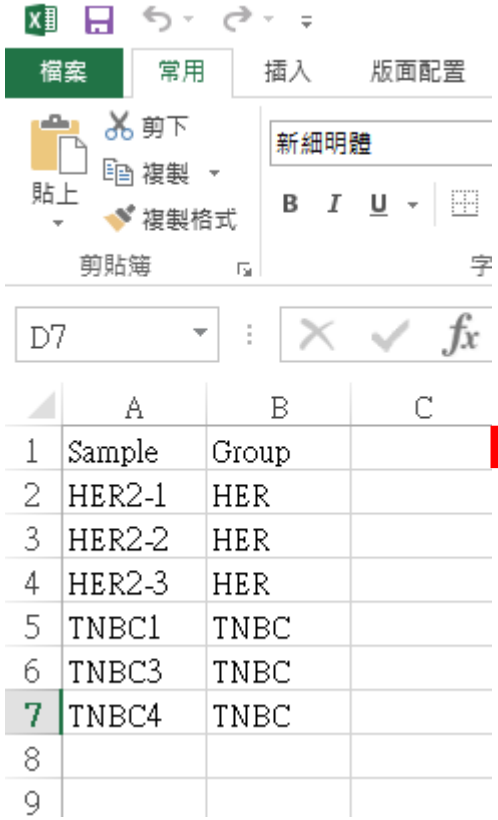


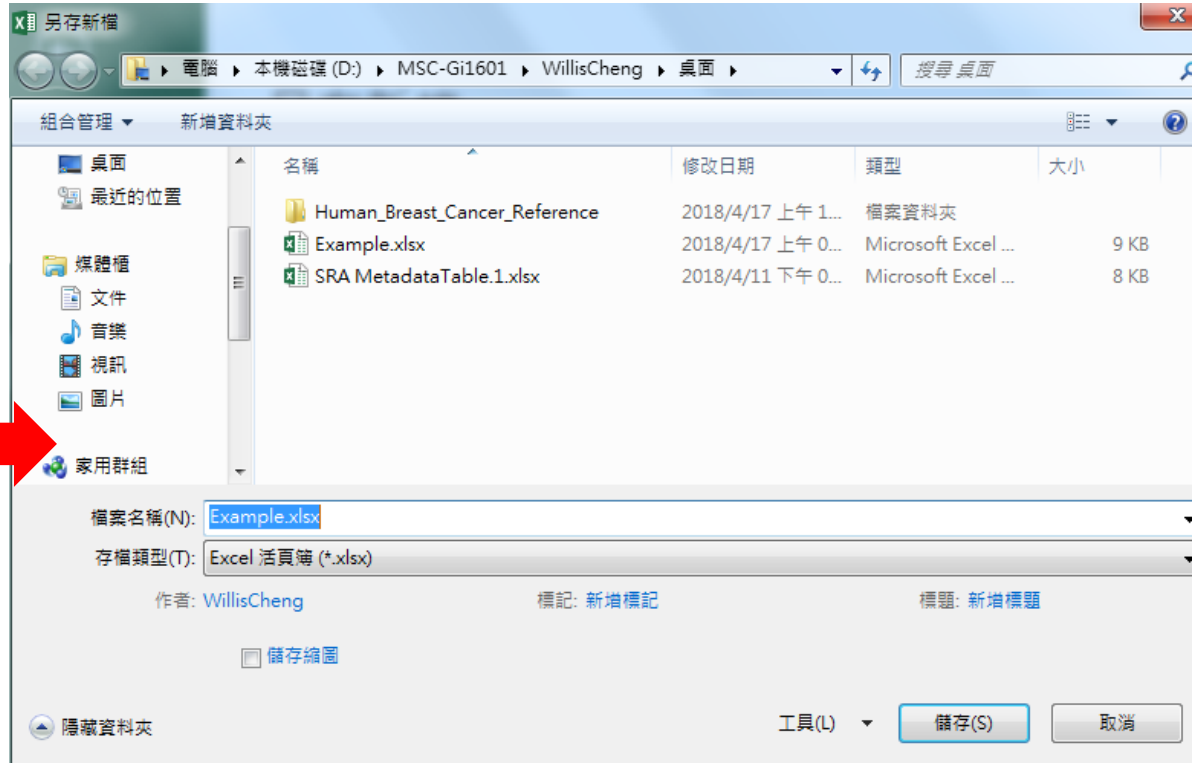
→ Save the workflow and share with other users!

Copyright©2018 GGA Corp. All rights reserved.

Build Metadata

- Use EXCEL, and save as xlsx file

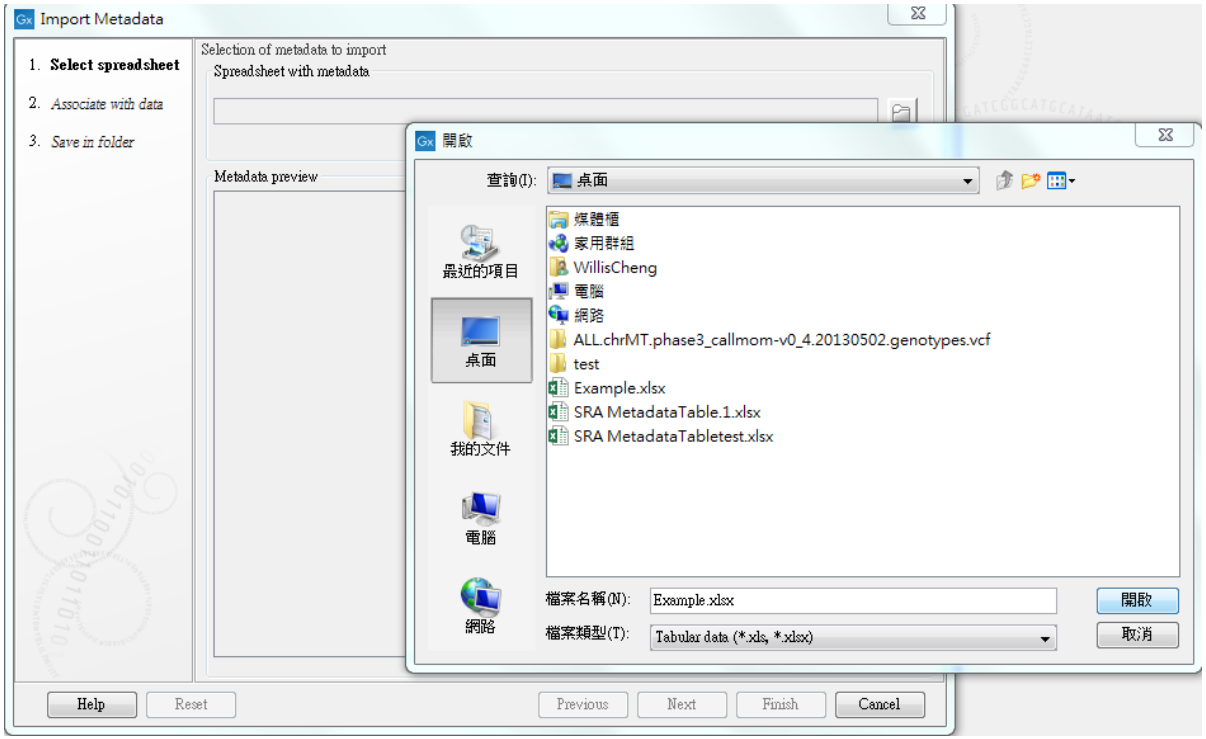




Copyright©2018 GGA Corp. All rights reserved.

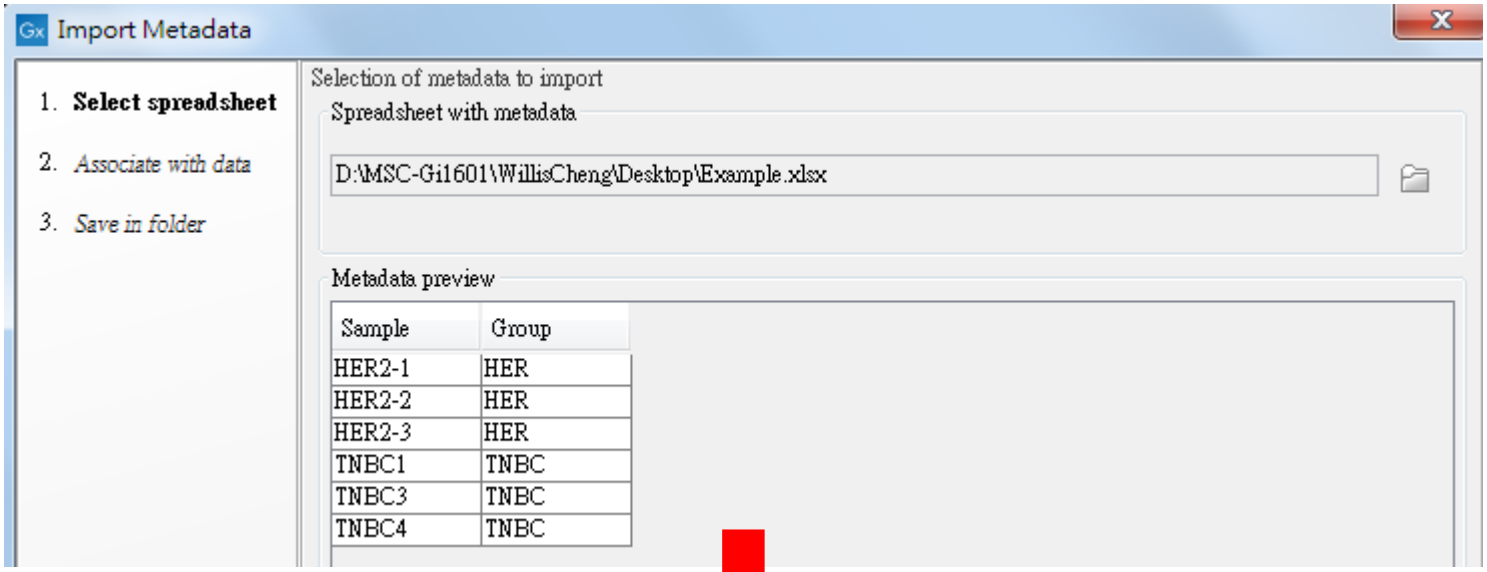
Import Metadata

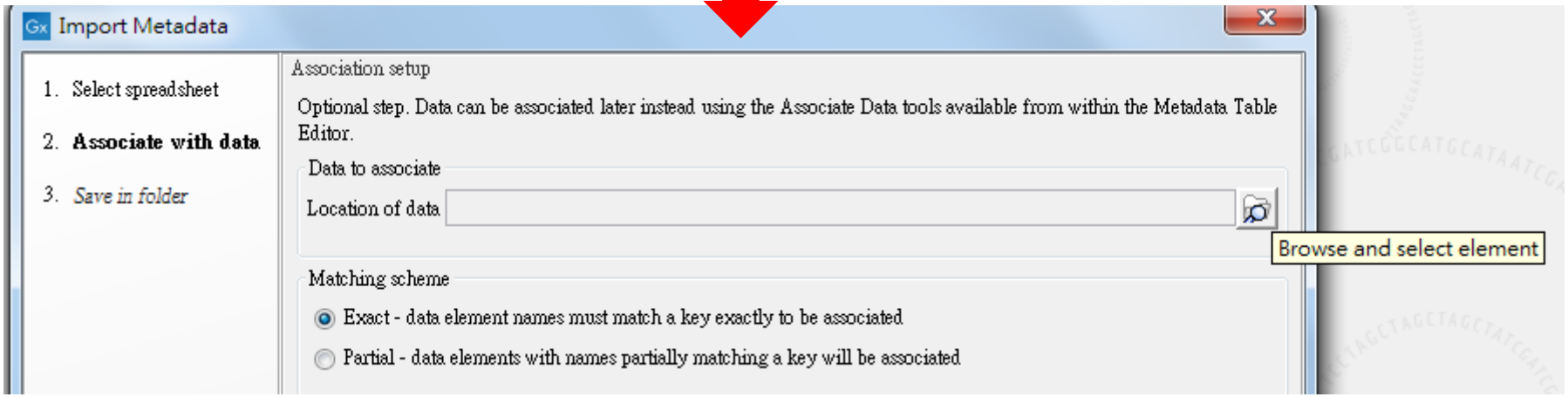
- **Import > Import Metadata**
  - Choose Example.xlsx
  - Link associated ID to fastq files



Copyright©2018 GGA Corp. All rights reserved.

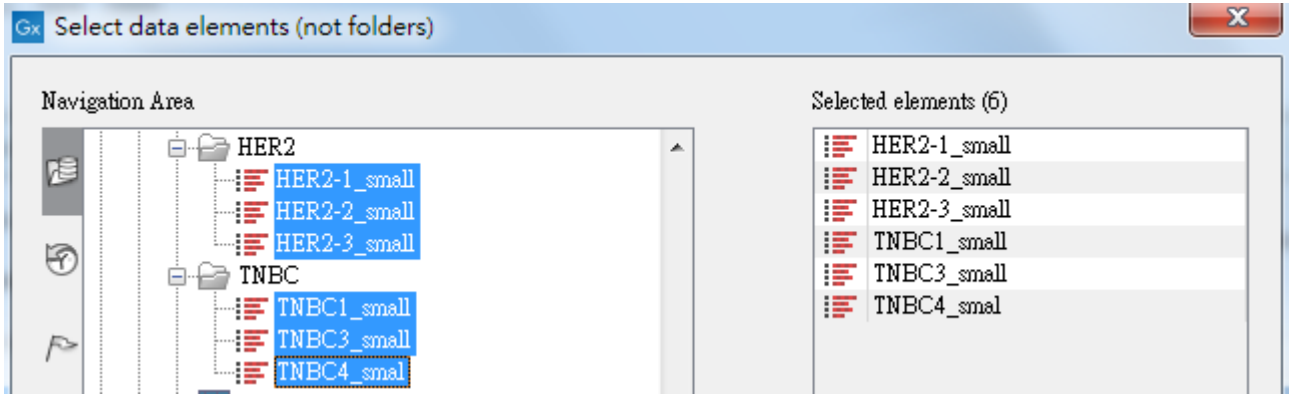
Import Metadata

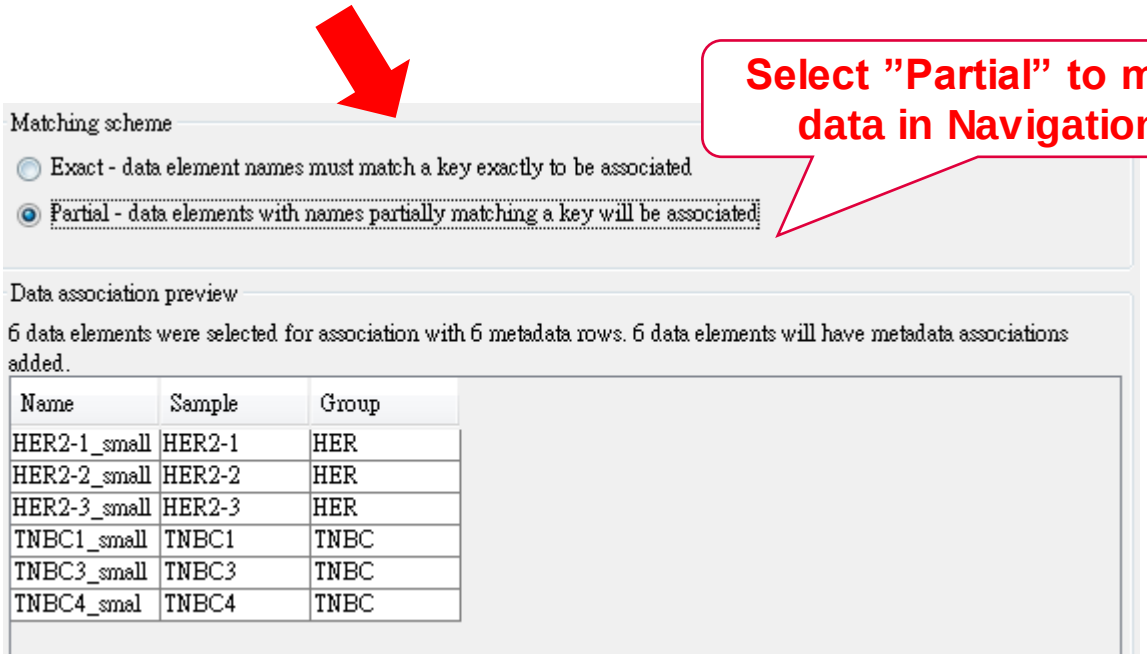




Copyright©2018 GGA Corp. All rights reserved.

Import Metadata

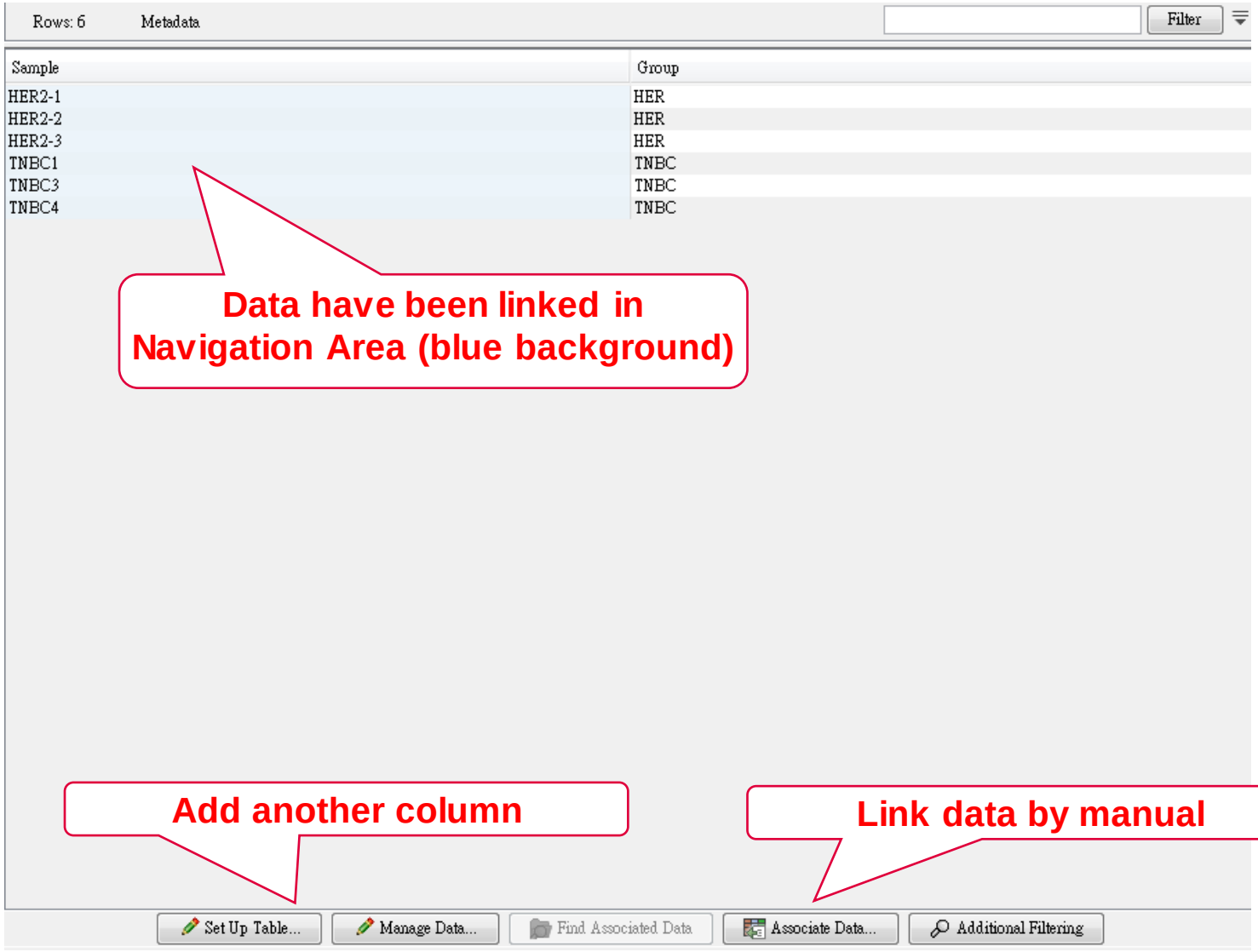




Select "Partial" to match the data in Navigation Area

Copyright©2018 GGA Corp. All rights reserved.

Metadata



Copyright©2018 GGA Corp. All rights reserved.

Reference

- CLC bio GWB / BxWB Download Link
  - https://www.qiagenbioinformatics.com/products/clc-genomics-workbench-direct-download/
  - https://www.qiagenbioinformatics.com/products/biomedical-genomics-workbench-direct-download/
- IPA / IVA Manual
  - https://www.qiagenbioinformatics.com/products/ingenuity-pathway-analysis/
  - https://www.qiagenbioinformatics.com/products/ingenuity-variant-analysis/
- Other Links
  - http://www.qiagenbioinformatics.com/
  - https://www.qiagenbioinformatics.com/product-login/

Copyright©2018 GGA Corp. All rights reserved.



www.QIAGEN.com
www.QIAGENBIOINFORMATICS.com

GGA Corp. : Willis Cheng, 02-2795-1777#3019
WillisCheng@GGA.ASIA

Copyright©2018 GGA Corp. All rights reserved.

Q&A

Copyright©2018 GGA Corp. All rights reserved.