# How to use NGS data



# GnCBIO

# Introduction

GNC Bio is a biotechnology company established in 2007. We mainly perform research services related to genomics (genome sequencing, DEG analysis, pathogen analysis, mitochondrial and chloroplast genome analysis, evolutionary research).

GNC Bio's Bioinformation Analysis Team performed various analyzes using NGS data. If you want to make various types of tables and figures using NGS data, please contact GNC Bio.

We will guide you on how to utilize NGS data along with friendly consultation.

#### Please wake up the sleeping data now!







sales@gncbio.kr

4th floor, 36, Banseok-ro, Yuseong-gu, Daejeon, Republic of Korea



## **Genome comparative analysis**

- 병원성 인자(virulence factor, VF)
- 항생제 내성(antimicrobial resistant, AMR) 유전자



■ 참조균주와의 비교 유전체 분석 등 기타

#### Visualization of core gene & gene cluster



\*Reference : Molecular characterization of megaplasmids encoding the type VI secretion system in *Campylobacter jejuni* isolated from chicken livers and gizzards

Heat map for nucleotide identity by orthology (OrthoANI)



\*Reference : Genome Sequence of a Potentially New Buttiauxella Species, Strain B2, Isolated from Rhizosphere of Olivillo Trees (Aextoxicon punctatum)



\*Reference : Whole genome-based characterisation of antimicrobial resistance and genetic diversity in *Campylobacter jejuni* and *Campylobacter coli* from ruminants



4 🗸 GnCBio

\*Reference : Whole genome-based characterisation of antimicrobial resistance and genetic diversity in *Campylobacter jejuni* and *Campylobacter coli* from ruminants

### Phylogenetic distance using virulence factor(VF)



\*Reference : Genomic Analysis and Antimicrobial Resistance of Campylobacter jejuni and Campylobacter coli in Peru



#### Visualization for Point mutation of drug-resistance genes

\*Reference : Prevalence and extent of heteroresistance by next generation sequencing of multidrugresistant tuberculosis

#### Genome structure comparison of phages



\*Reference : The Shiga toxin 2 production level in enterohemorrhagic Escherichia coli O157:H7 is correlated with the subtypes of toxinencoding phage



## Pairwise of chromosomal synteny between strain

\*Reference : Pangenome analyses of the wheat pathogen *Zymoseptoria tritici* reveal the structural basis of a highly plastic eukaryotic genome

9

### Visualization for pangenome analysis results



- a. Venn diagram of the singleton, accessory, and core genes of the pangenome.
- b. Categorization of the pangenome into singleton, accessory, and core genes.
- c. Number of accessory gene clusters.
- d. Estimation of the sizes of the core genome and the pangenome.

\*Reference : Pangenome analyses of the wheat pathogen Zymoseptoria tritici reveal the structural basis of a highly plastic eukaryotic genome

# **Phylogenetic** analysis

- VNTR (variable number tandem repeat)
- SNP (Single nucleotide polymorphism)
- InDeL (Insertions / Deletions)
- SSR (Simple sequence repeat)
- STR (Short Tandem Repeat)

### Phylogenetic tree of VNTR or SNP



**Profiling & Visualization** 

\*Reference : Whole genome sequencing analysis of multidrug-resistant tuberculosis in Singapore, 2006–2018



\*Reference : Serotype Diversity and Antimicrobial Resistance among Salmonella enterica Isolates from Patients at an Equine Referral Hospital

# **Amplicon sequence analysis**

- Metabarcoding sequencing
- Metagenome sequencing



### Distribution & statistical analysis

Target region sequencing

### Visualization of charts and statistics

# Proportions(%) of sequence sources of variety samples

\*Reference : Diet composition of the Korean wild boar Sus scrofa coreanus (Suidae) at Mt. Jeombongsan, Korea









confidence cutoff = 0.70

\*Reference : GNCBIO's report

# **NGS read analysis**

- Amplicon sequence
- In vivo transcription library sequence
- Polymerase cycling assembly(PCA) library sequence



#### Read classification by BLASTN searches

**Read classification &** 

clustering

\*Reference : A metagenomics study for the identification of respiratory viruses in mixed clinical specimens: an application of the iterative mapping approach

#### Read distributions of the viral genomes revealed by read mapping



\*Reference : A metagenomics study for the identification of respiratory viruses in mixed clinical specimens: an application of the iterative mapping approach



Analysis of random-sequence genetic oligomer pools

\*Reference : Random-sequence genetic oligomer pools display an innate potential for ligation and recombination

# Virome analysis





# **Building a Pipeline**



Workflow of prediction of avian influenza virus subtype (PAIVS).

\*Reference : PAIVS: prediction of avian influenza virus subtype



# Workflow of variant analysis (SNVs, single nucleotide variants. CNVs, copy number variation)

\*Reference : Building a Robust Tumor Profiling Program: Synergy between Next-Generation Sequencing and Targeted Single-Gene Testing

# **Correlation analysis**





Nonmetric multidimensional scaling (NMDS) ordination of a presence-absence matrix of families detected in each sample.

\*Reference : Testing multiple substrates for terrestrial biodiversity monitoring using environmental DNA metabarcoding

# Statistical significance analysis



Analysis of gene expression level change by inoculation cycle

#### **(A)** 'Peason r' values

	TNF-α	TGF-β	IL-17	IL-12	IL-10	L-6	IL-2	IL-1β	IFN-y	CXCL10
TNF-α	1.00									
TGF-β	0.98	1.00								
IL-17	0.98	0.92	1.00							
IL-12	0.93	0.83	0.98	1.00						
IL-10	0.41	0.27	0.60	0.69	1.00					
IL-6	0.99	0.96	0.99	0.95	0.50	1.00				
IL-2	0.93	0.83	0.98	1.00	0.69	0.95	1.00			
IL-1β	0.84	0.93	0.79	0.64	0.19	0.84	0.64	1.00		
IFN-γ	0.75	0.59	0.84	0.94	0.78	0.79	0.94	0.33	1.00	
CXCL10	0.38	0.14	0.48	0.64	0.63	0.39	0.64	-0. 18	0.86	1.00

#### **B** 'Peason P' values

	TNF-α	TGF-β	IL-17	IL-12	IL-10	L-6	IL-2	IL-1β	IFN-γ	CXCL10
TNF-α	0.00									
TGF-β	0.02	0.00								
IL-17	0.02	0.08	0.00							
IL-12	0.07	0.17	0.02	0.00						
IL-10	0.59	0.73	0.40	0.31	0.00					
IL-6	0.01	0.04	0.01	0.05	0.50	0.00				
IL-2	0.07	0.17	0.02	0.00	0.31	0.05	0.00			
IL-1β	0.16	0.07	0.21	0.38	0.81	0.16	0.36	0.00		
IFN-γ	0.25	0.41	0.16	0.08	0.22	0.21	0.06	0.67	0.00	
CXCL10	0.64	0.86	0.54	0.38	0.37	0.61	0.36	0.82	0.14	0.00





# Change is chance

www.gncbio.com



34069대전광역시 유성구 반석로 36 예건프라자 4층 TEL. 042-824-8258 FAX. 042-824-8268 E-MAIL. sales@gncbio.kr