

Genomics & Bioinformatics

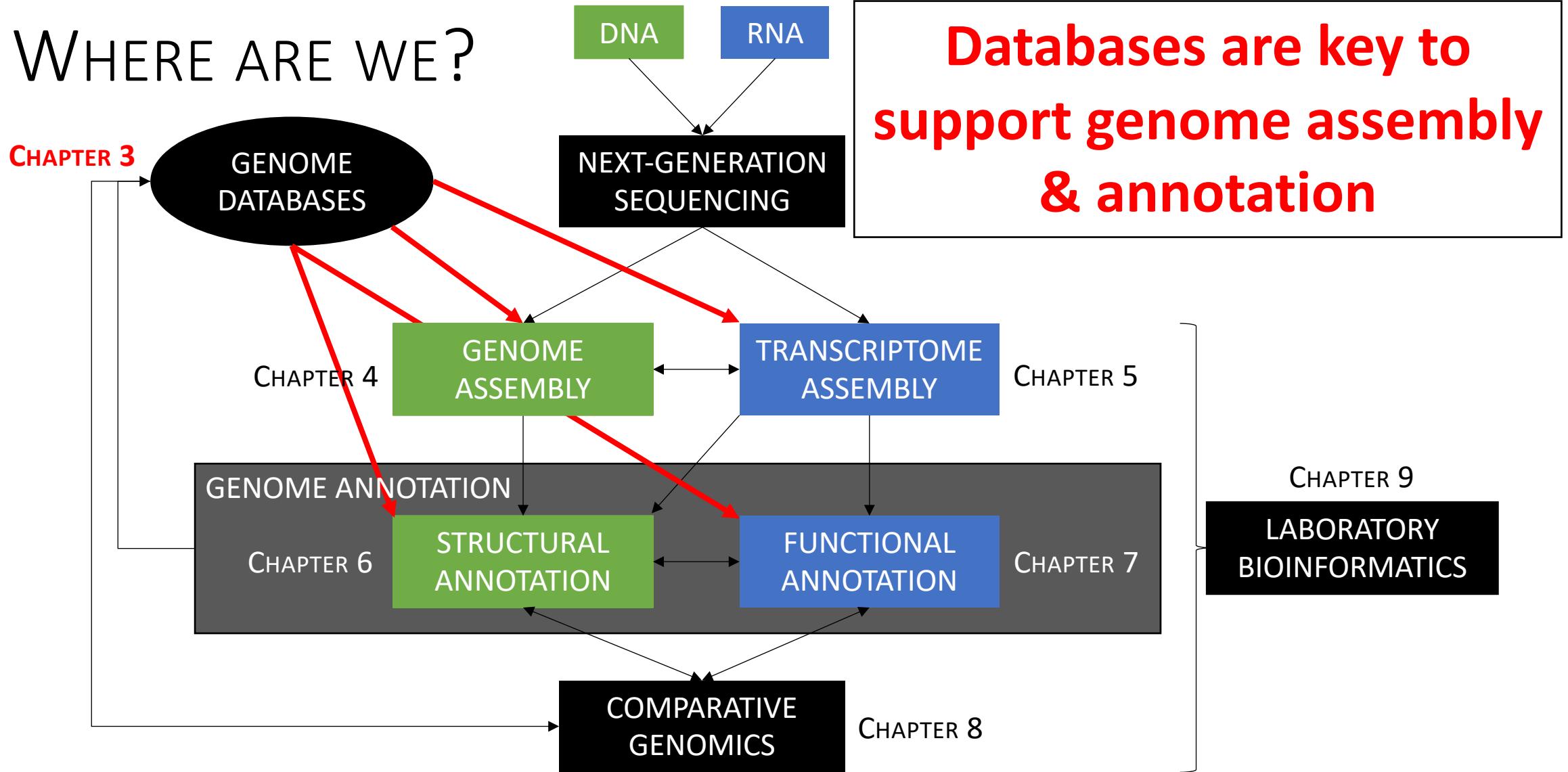


BIOL 497, 597

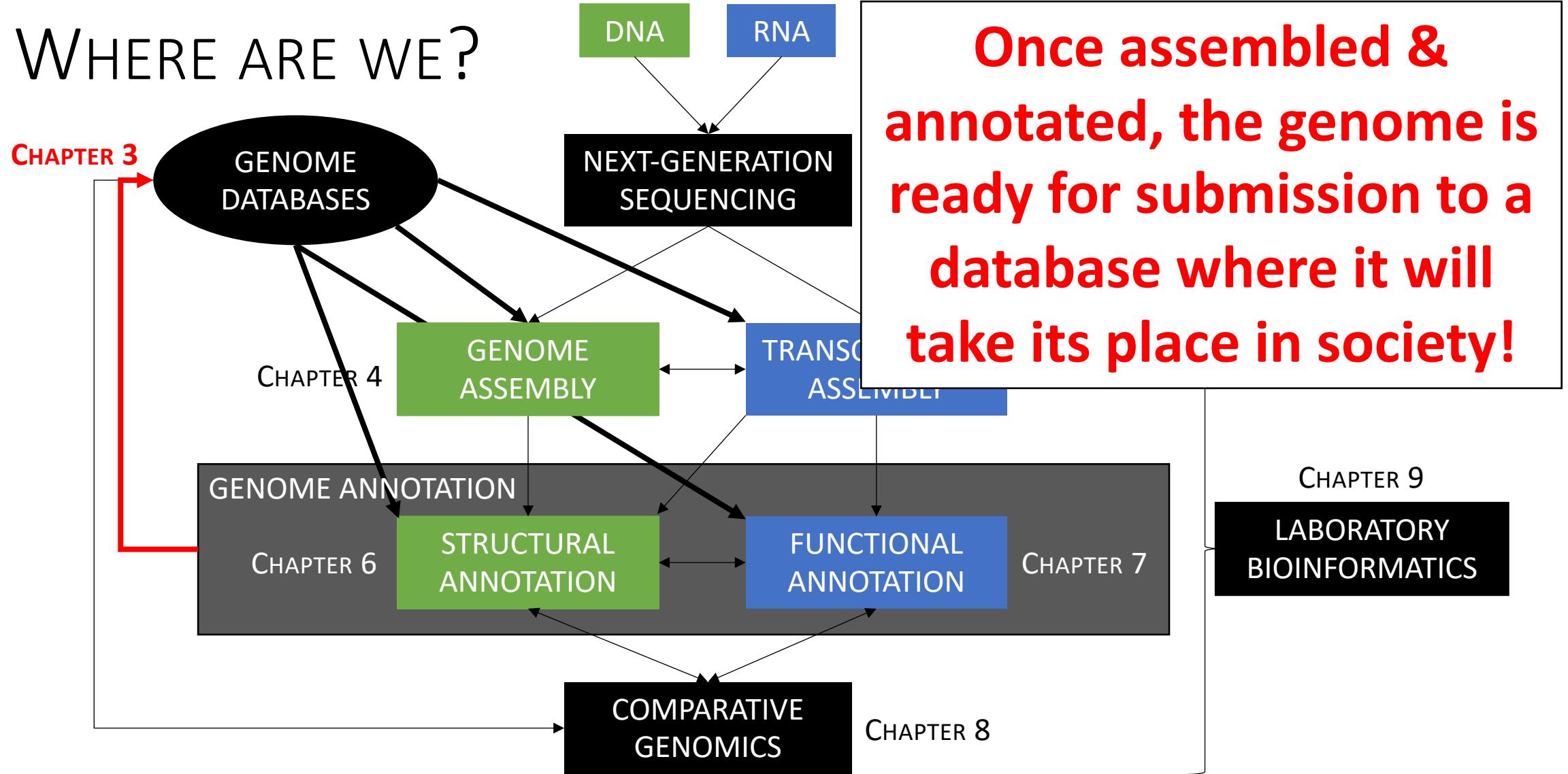
Boise State University

Spring

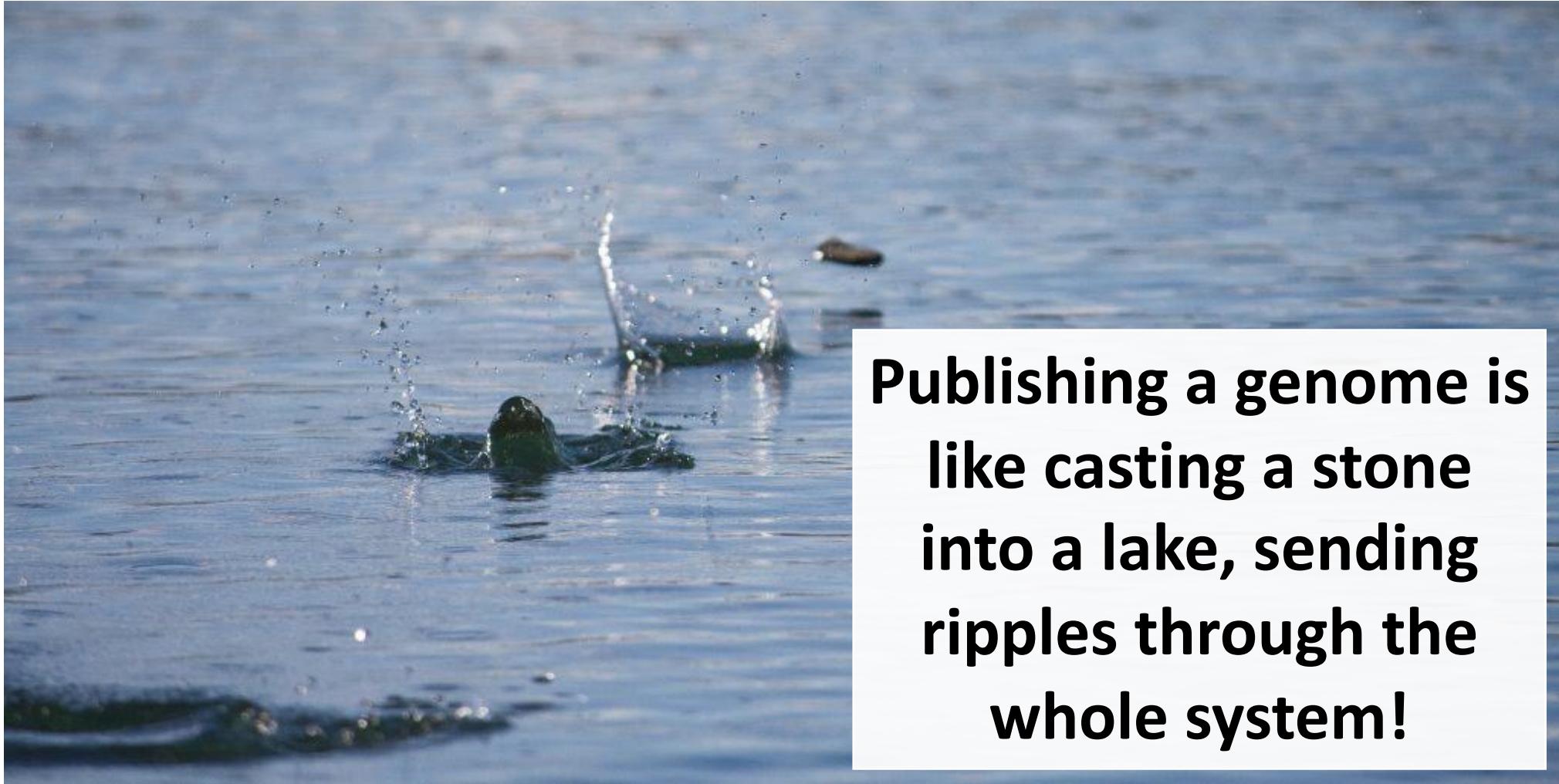
WHERE ARE WE?



WHERE ARE WE?

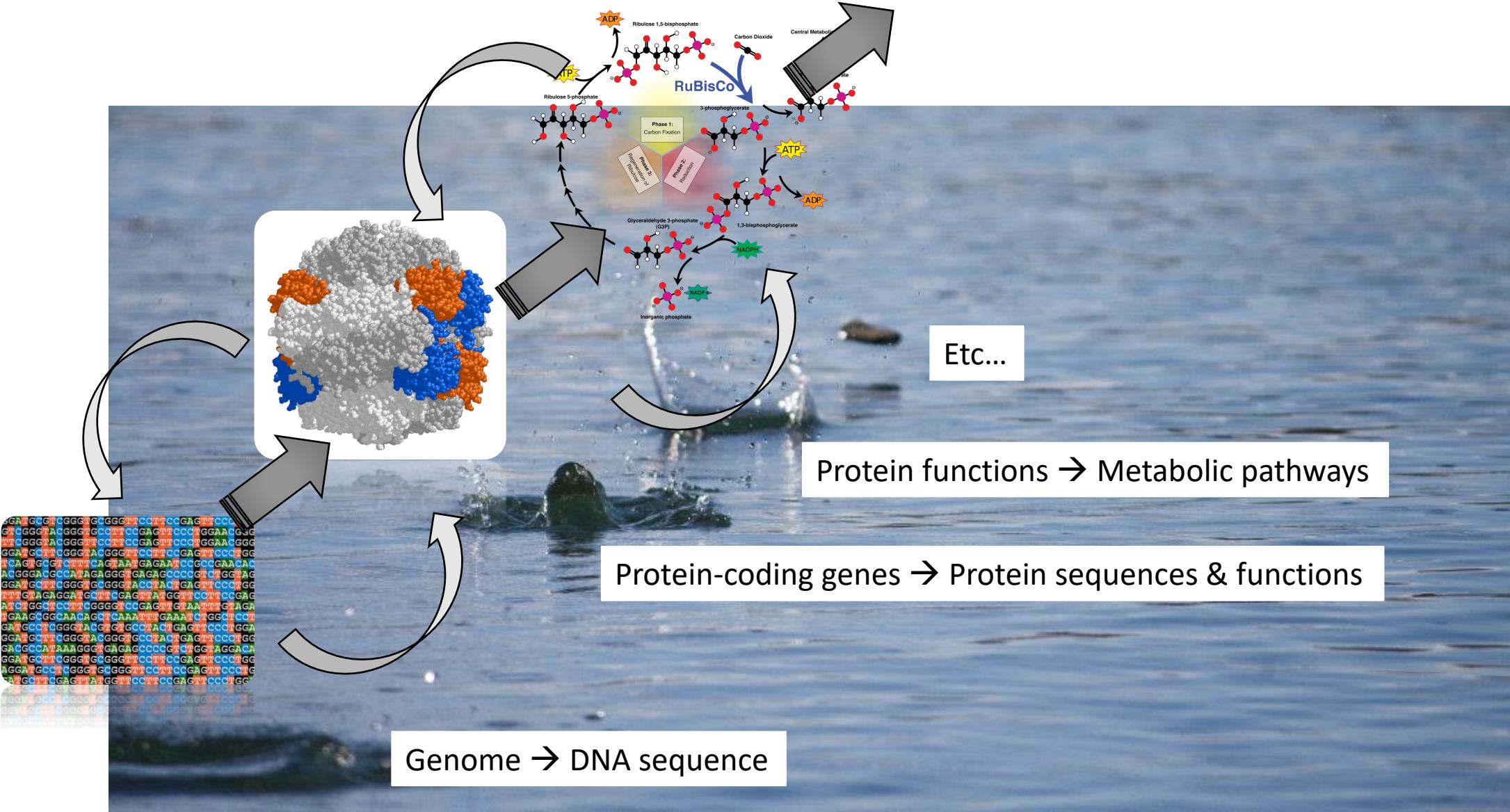


DATABASES FORM AN INTERLOCKING NETWORK

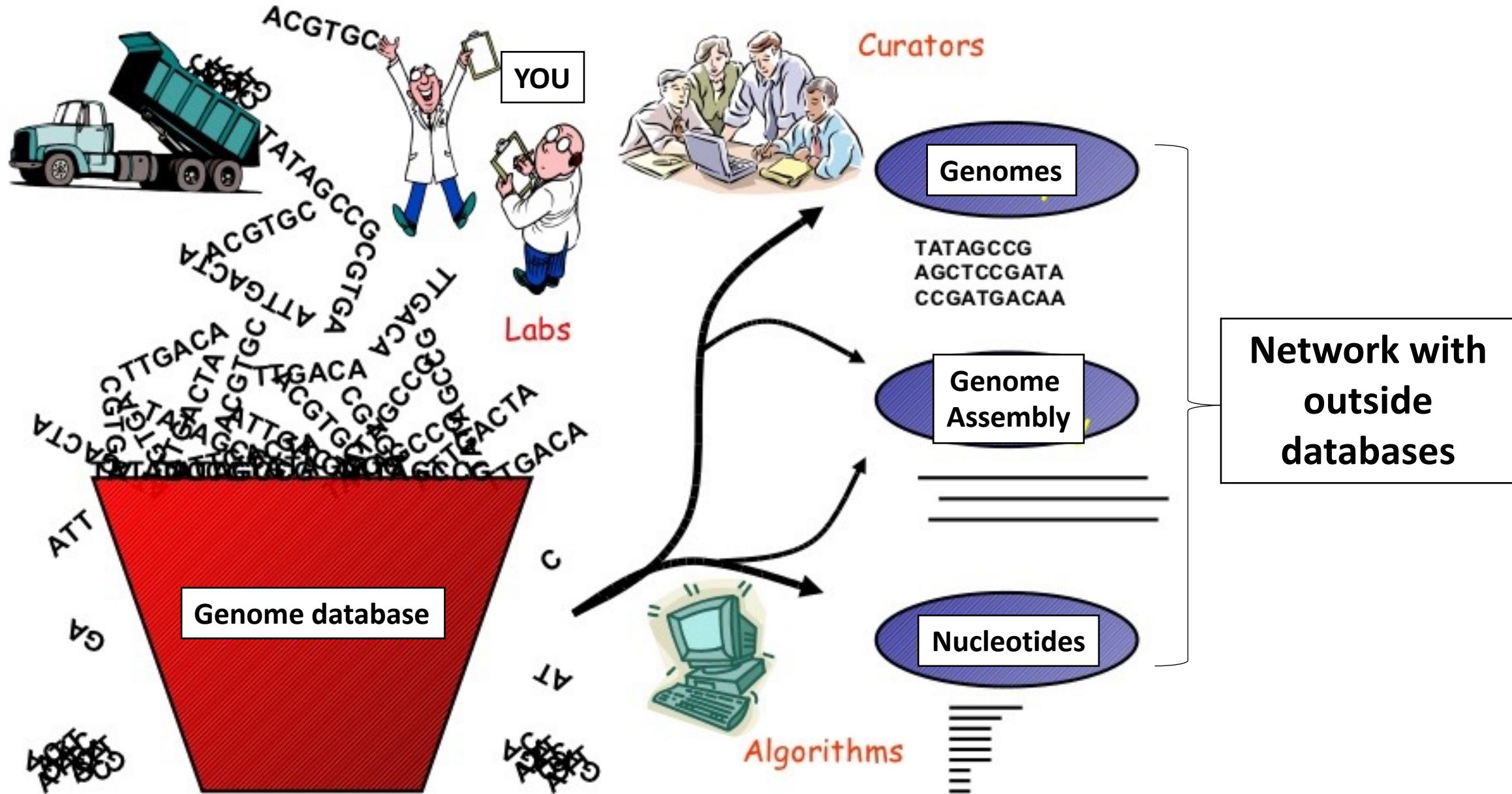


**Publishing a genome is
like casting a stone
into a lake, sending
ripples through the
whole system!**

DATABASES FORM AN INTERLOCKING NETWORK



GENOME DATABASES – OVERVIEW OF WORKFLOW



REQUIREMENTS OF GENOME DATABASES

Institutions in charge of molecular databases develop tools to:

- **Harvest and curate data** (plus annotations) – that is, check both for accuracy and format – and distribute them.
- **Track and archive data** so that they do not get lost.
- **Record provenance and other information on samples** (e.g. location, vouchers, tissue type, taxonomy).
- **Provide links from the data to relevant items** in other databases, including bibliographical libraries (e.g. PubMed).
- **Provide information retrieval and analysis software** to support research: recovery of selected data and calculations with them (e.g. SRA toolkit, BLAST).
- **Provide documentation and tutorials.**
- Keep up with scientific advances in both biology and informatics.
- Be responsive to users' needs.

GENOME DATABASES – MAJOR TYPES

- Nucleic acid sequences databases
- Protein sequences databases
- Gene ontology databases
- Metabolic pathways databases
- Specialized annotated genomes portals

GENOME DATABASES – MAJOR TYPES

- Nucleic acid sequences databases (In class)
- Protein sequences databases
- Gene ontology databases
- Metabolic pathways databases
- Specialized annotated genomes portals.

Mini-Report 2



NUCLEIC ACID SEQUENCES DATABASES

The International Nucleotide Sequence Database Collaboration (INSDC) is a partnership between 3 DNA seq. databases:

- ✓ **DDBJ:** DNA Data Bank of Japan.
- ✓ **EMBL-EBI:** European Bioinformatics Institute.
- ✓ **NCBI:** National Center for Biotechnology Information (USA).



NUCLEIC ACID SEQUENCES DATABASES

Data type	DDBJ	EMBL-EBI	NCBI
Next generation reads	Sequence Read Archive		Sequence Read Archive
Capillary reads	Trace Archive	European Nucleotide Archive (ENA)	Trace Archive
Annotated sequences	DDBJ		GenBank
Samples	BioSample		BioSample
Studies	BioProject		BioProject

NEXT GENERATION READS - SRA

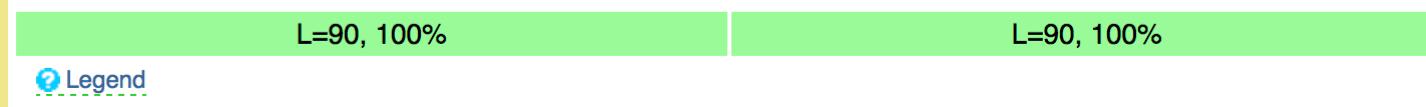
- Sequence Read Archive ([SRA](#)) makes biological sequence data available to the research community to enhance reproducibility and allow for new discoveries.
- The SRA stores raw sequencing data (WGS, RNA-Seq) and alignment information from high-throughput sequencing platforms.
- [NCBI SRA Toolkit](#) allows to remotely download SRA files.
- We will learn protocols to download SRA files in class (Chapter 4).

WGS of *Apostasia shenzhenica*: 180 insert size (SRR5759389)

Metadata Analysis (alpha) Reads Download

Run	Spots	Bases	Size	GC content	Published	Access Type
SRR5759389	84.1M	15.1Gbp	11.3G	35.5%	2017-06-27	public

This run has 2 reads per spot:



Experiment	Library Name	Platform	Strategy	Source	Selection	Layout
SRX2959224	Apostasia180	Illumina	WGS	GENOMIC	PCR	PAIRED
to BLAST						

Design:

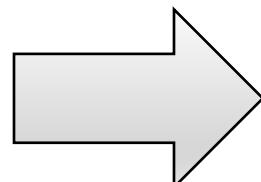
180 insert size library on Illumina

Biosample	Sample Description	Organism	Links
SAMN04453324 (SRS2316248)		Apostasia shenzhenica	<ul style="list-style-type: none"> PRJNA310678 [Apostasia shenzhenica isolate:ASH160606] The Apostasia genome and the evolution of orchids.

Bioproject	SRA Study	Title
PRJNA310678	SRP109877	Apostasia shenzhenica isolate:ASH160606 Genome sequencing and assembly
Show abstract		



NUCLEIC ACID SEQUENCES DATABASES



Data type	DDBJ	EMBL-EBI	NCBI
Next generation reads	Sequence Read Archive		Sequence Read Archive
Capillary reads	Trace Archive	European Nucleotide Archive (ENA)	Trace Archive
Annotated sequences	DDBJ		GenBank
Samples	BioSample		BioSample
Studies	BioProject		BioProject

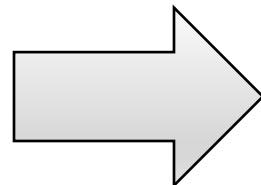
CAPILLARY READS – TRACE ARCHIVES

[Trace Archive](#) serves as repository of sequencing data from capillary platforms e.g. Applied Biosystems ABI 3730 (= Sanger sequencing).





NUCLEIC ACID SEQUENCES DATABASES



Data type	DDBJ	EMBL-EBI	NCBI
Next generation reads	Sequence Read Archive		Sequence Read Archive
Capillary reads	Trace Archive	European Nucleotide Archive (ENA)	Trace Archive
Annotated sequences	DDBJ		GenBank
Samples	BioSample		BioSample
Studies	BioProject		BioProject

ANNOTATED SEQUENCES – GENBANK

NCBI Resources How To

[Sign in to NCBI](#)

GenBank

Nucleotide



Search

GenBank

Submit

Genomes

WGS

Metagenomes

TPA

TSA

INSDC

Other

GenBank Overview

What is GenBank?

GenBank® is the NIH genetic sequence database, an annotated collection of all publicly available DNA sequences ([Nucleic Acids Research, 2013 Jan;41\(D1\):D36-42](#)). GenBank is part of the [International Nucleotide Sequence Database Collaboration](#), which comprises the DNA DataBank of Japan (DDBJ), the European Nucleotide Archive (ENA), and GenBank at NCBI. These three organizations exchange data on a daily basis.

A GenBank release occurs every two months and is available from the [ftp site](#). The [release notes](#) for the current version of GenBank provide detailed information about the release and notifications of upcoming changes to GenBank. Release notes for [previous GenBank releases](#) are also available. GenBank growth statistics for both the traditional GenBank divisions and the WGS division are available from each release. GenBank growth [statistics](#) for both the traditional GenBank divisions and the WGS division are available from each release.

GenBank Resources

[GenBank Home](#)

[Submission Types](#)

[Submission Tools](#)

[Search GenBank](#)

[Update GenBank Records](#)

ANNOTATED SEQUENCES – GENBANK

There are several ways to search and retrieve data from GenBank:

- a. Search GenBank for sequence identifiers and annotations with [Entrez Nucleotide](#).
 - Entrez is divided into three divisions: [CoreNucleotide](#) (the main collection), [dbEST](#) (Expressed Sequence Tags), and [dbGSS](#) (Genome Survey Sequences).
- b. Search and align GenBank sequences to a query sequence using [BLAST](#) (Basic Local Alignment Search Tool).

ANNOTATED SEQUENCES – GENBANK

There are several ways to search and retrieve data from GenBank:

- c. Search, link, and download sequences programmatically using [NCBI e-utilities](#).
- d. The ASN.1 and flatfile formats are available at **NCBI's anonymous FTP server**: <ftp://ftp.ncbi.nlm.nih.gov/ncbi-asn1> and <ftp://ftp.ncbi.nlm.nih.gov/genbank>.

ANNOTATED SEQUENCES – GENBANK

GenBank Flat File format

<u>LOCUS</u>	SCU49845	<u>5028 bp</u>	<u>DNA</u>	<u>PLN</u>	21-JUN-1999
<u>DEFINITION</u>	Saccharomyces cerevisiae TCP1-beta gene, partial cds, and Axl2p (AXL2) and Rev7p (REV7) genes, complete cds.				
<u>ACCESSION</u>	U49845				
<u>VERSION</u>	U49845.1 GI:1293613				
<u>KEYWORDS</u>	.				
<u>SOURCE</u>	Saccharomyces cerevisiae (baker's yeast)				
<u>ORGANISM</u>	Saccharomyces cerevisiae Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Saccharomyces.				
<u>REFERENCE</u>	1 (bases 1 to 5028)				
<u>AUTHORS</u>	Torpey,L.E., Gibbs,P.E., Nelson,J. and Lawrence,C.W.				
<u>TITLE</u>	Cloning and sequence of REV7, a gene whose function is required for DNA damage-induced mutagenesis in <i>Saccharomyces cerevisiae</i>				
<u>JOURNAL</u>	Yeast 10 (11), 1503-1509 (1994)				
<u>PUBMED</u>	7871890				
<u>REFERENCE</u>	2 (bases 1 to 5028)				
<u>AUTHORS</u>	Roemer,T., Madden,K., Chang,J. and Snyder,M.				
<u>TITLE</u>	Selection of axial growth sites in yeast requires Axl2p, a novel plasma membrane glycoprotein				
<u>JOURNAL</u>	Genes Dev. 10 (7), 777-793 (1996)				
<u>PUBMED</u>	8846915				
<u>REFERENCE</u>	3 (bases 1 to 5028)				
<u>AUTHORS</u>	Roemer,T.				
<u>TITLE</u>	Direct Submission				
<u>JOURNAL</u>	Submitted (22-FEB-1996) Terry Roemer, Biology, Yale University, New Haven, CT, USA				

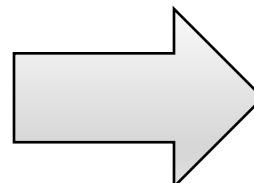
<u>FEATURES</u>	<u>Location/Qualifiers</u>
<u>source</u>	1..5028 /organism="Saccharomyces cerevisiae" /db_xref="taxon:4932" /chromosome="IX" /map="9" <u><1..206</u> /codon_start=3 /product="TCP1-beta" <u>/protein_id=AAA98665.1</u> /db_xref="GI:1293614" <u>/translation=Ssiyngistsgldlnngtiadmrqlgivesyklkravvssasea</u> AEVLLRVDNIIRARPTANRQHM"
<u>CDS</u>	687..3158 /gene="AXL2"
<u>gene</u>	687..3158 /gene="AXL2" /note="plasma membrane glycoprotein" /codon_start=1 /function="required for axial budding pattern of S. cerevisiae" /product="Ax12p" <u>/protein_id=AAA98666.1</u> /db_xref="GI:1293615" <u>/translation=Mtqlqislltatislhvvatpyeaypigkqyppvarvnesf</u> TFQISNDTYKSSVDKTAQITYNCFDLPSWLSFDSSRTFSGEPSSDLLSDANTTLYFN VILEGTDSDADSTSLNNTYQFVVTNRPSISLSSDFNLLALLKNYGYTNGKNALKLDPNE VFNVTFDRSMFTNEESIVSYYGRSQLYNAPLPNWLFDFSGELKFTGTAPVINSIAPE TSYSFVIIATDIEGFSAVEVEFELVIGAHQLTTSIQNSLIINVDTGNVSYDLPLNYV YLDDDPPISSDKLGSINLLDAPDWALDNATISGSVPDELLGKNSNPANFSVSIYDTYG DVIYFNFEVVSTTDLFAISSLPNINATRGGEWFSSYYFLPSQFTDYVNTNVSLFTNSSQ DHDWVKFQSSNLTLAGEVPKNFDKLGLKANQGSQSQELYFNIIGMDSKITHSNHSA NATSTRSSHSTSTSSYTSSTYTAKISSTSAAATSSAPAALPAANKTSSHNKAVAIA CGVAIPLGVILVALICFLIFWRRRRENPDDENLPHAIISGPDLNNPANKPNQENATPLN
<u>CDS</u>	687..3158 /gene="AXL2" /note="plasma membrane glycoprotein" /codon_start=1 /function="required for axial budding pattern of S. cerevisiae" /product="Ax12p" <u>/protein_id=AAA98666.1</u> /db_xref="GI:1293615" <u>/translation=Mtqlqislltatislhvvatpyeaypigkqyppvarvnesf</u> TFQISNDTYKSSVDKTAQITYNCFDLPSWLSFDSSRTFSGEPSSDLLSDANTTLYFN VILEGTDSDADSTSLNNTYQFVVTNRPSISLSSDFNLLALLKNYGYTNGKNALKLDPNE VFNVTFDRSMFTNEESIVSYYGRSQLYNAPLPNWLFDFSGELKFTGTAPVINSIAPE TSYSFVIIATDIEGFSAVEVEFELVIGAHQLTTSIQNSLIINVDTGNVSYDLPLNYV YLDDDPPISSDKLGSINLLDAPDWALDNATISGSVPDELLGKNSNPANFSVSIYDTYG DVIYFNFEVVSTTDLFAISSLPNINATRGGEWFSSYYFLPSQFTDYVNTNVSLFTNSSQ DHDWVKFQSSNLTLAGEVPKNFDKLGLKANQGSQSQELYFNIIGMDSKITHSNHSA NATSTRSSHSTSTSSYTSSTYTAKISSTSAAATSSAPAALPAANKTSSHNKAVAIA CGVAIPLGVILVALICFLIFWRRRRENPDDENLPHAIISGPDLNNPANKPNQENATPLN

ORIGIN

1 gatcctccat atacaacggc atctccacct caggtttaga tctcaacaac ggaaccattg
61 ccgacatgag acagtttagt atcgctgaga gttacaagct aaaacgagca gtagtcagct
121 ctgcatactga agccgctgaa gttctactaa gggtggataa catcatccgt gcaagaccaa
181 gaaccgcca tagacaacat atgtaacata tttaggatat acctcgaaaa taataaaccg
241 ccacactgtc attattataa ttagaaacag aacgcaaaaa ttatccacta tataattcaa
301 agacgcgaaa aaaaaagaac aacgcgtcat agaactttg gcaattcgcg tcacaaataa
361 attttggcaa cttatgtttc ctcttcgagc agtactcgag ccctgtctca agaatgtaat
421 aatacccatc gtaggtatgg ttaaagatag catctccaca acctcaaagc tccttgccga
481 gagtcgcct cctttgtcga gtaatttca ctttcatat gagaacttat tttcttattc
541 tttactctca catcctgttag tgattgacac tgcaacagcc accatcacta gaagaacacaga
601 acaattactt aatagaaaaa ttatatcttc ctcgaaacga tttcctgctt ccaacatcta
661 cgtatatcaa gaagcattca cttaccatga cacagctca gatttcatta ttgctgacag
721 ctactatatc actactccat ctagtagtgg ccacgcctta tgaggcatat cctatcgaa
781 aacaataccc cccagtggca agagtcaatg aatcgttac atttcaaatt tccaatgata
841 cctataaaatc gtctgttagac aagacagctc aaataacata caattgcttc gacttaccga
901 gctggcttc gtttgactct agttctagaa cgttctcagg tgaaccttct tctgacttac
961 tatctgatgc gaacaccacg ttgttattca atgtaatact cgagggtacg gactctgccg
021 acagcacgtc tttgaacaat acataccaaat ttgttgttac aaaccgtcca tccatctcgcc
081 tatcgtcaga tttcaatcta ttggcgttgt taaaaaaacta tggttatact aacggaaaaa
141 acgctctgaa actagatcct aatgaagtct tcaacgtgac ttttgaccgt tcaatgttca
201 ctaacgaaga atccattgtg tcgttattacg gacgttctca gttgtataat gcgcgcgttac
261 ccaattggct gttcttcgat tctggcgagt tgaagttac tgggacggca ccggtgataa
321 actcggcgat tgctccagaa acaagctaca gtttgtcat catcgctaca gacattgaag
381 gattttctgc cgttgaggta gaattcgaat tagtcatcg ggctcaccag ttaactacct
441 ctattcaaaa tagttgata atcaacgtta ctgacacagg taacgttca tatgacttac
501 ctctaaacta tgtttatctc gatgacgatc ctatttcttc tgataaaattt ggttctataa
561 acttatttggta tgctccagac tgggtggcat tagataatgc taccatttcc gggctgtcc
621 cagatgaatt actcggtaag aactccaatc ctgccaattt ttctgtgtcc atttatgata
681 cttatggta tgtgatttat ttcaacttcg aagttgtctc cacaacggat ttgtttgcca
741 tttagttctt tcccaatatt aacgctacaa ggggtgaatg gttctcctac tatttttgc
801 cttctcagtt tacagactac gtgaatacaa acgtttcatt agagtttact aattcaagcc
861 aagaccatga ctgggtgaaa ttccaaatcat ctaatttaac attagctgga gaagtgc
921 agaatttcga caagcttca ttaggttga aagcgaacca aggttcacaa tctcaagagc
981



NUCLEIC ACID SEQUENCES DATABASES



Data type	DDBJ	EMBL-EBI	NCBI
Next generation reads	Sequence Read Archive		Sequence Read Archive
Capillary reads	Trace Archive	European Nucleotide Archive (ENA)	Trace Archive
Annotated sequences	DDBJ		GenBank
Samples	BioSample		BioSample
Studies	BioProject		BioProject

SAMPLES – BIOSAMPLES

The [BioSample](#) database contains descriptions of biological source materials used in experimental assays.

Plant sample from *Apostasia shenzhenica*

Identifiers	BioSample: SAMN07191733; Sample name: <i>Apostasia shenzhenica</i> tuber; SRA: SRS2300266
Organism	<i>Apostasia shenzhenica</i> cellular organisms; Eukaryota; Viridiplantae; Streptophyta; Streptophytina; Embryophyta; Tracheophyta; Euphylophyta; Spermatophyta; Magnoliophyta; Mesangiospermae; Liliopsida; Petrosaviidae; Asparagales; Orchidaceae; Apostasioideae; Apostasia
Package	Plant; version 1.0
Attributes	isolate wild <i>Apostasia shenzhenica</i> development stage reproductive growth geographic location China: Shenzhen tissue tuber
BioProject	PRJNA310678 <i>Apostasia shenzhenica</i> isolate:ASH160606 Retrieve all samples from this project
Submission	The National Orchid Conservation Center of China, Zhongjian Liu; 2017-06-04

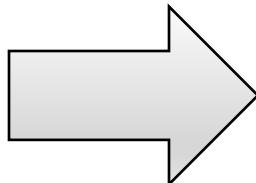
Accession: SAMN07191733 ID: 7191733

[BioProject](#) [SRA](#)



NUCLEIC ACID SEQUENCES DATABASES

Data type	DDBJ	EMBL-EBI	NCBI
Next generation reads	Sequence Read Archive		Sequence Read Archive
Capillary reads	Trace Archive	European Nucleotide Archive (ENA)	Trace Archive
Annotated sequences	DDBJ		GenBank
Samples	BioSample		BioSample
Studies	BioProject		BioProject



STUDIES – BIOPROJECT

- A [BioProject](#) is a collection of biological data related to a single initiative, originating from a single organization or from a consortium.
- A BioProject record provides users a single place to find links to the diverse data types generated for that project.

Apostasia shenzhenica isolate:ASH160606 Genome sequencing

Apostasia shenzhenica has a karyotype of 2N=2X=68 with uniform small-size chromosomes . [More...](#)

Accession	PRJNA310678
Data Type	Genome sequencing
Scope	Monoisolate
Organism	Apostasia shenzhenica [Taxonomy ID: 1088818] Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Orchidaceae; Apostasioideae; Apostasia; Apostasia shenzhenica
Publications	Zhang GQ <i>et al.</i> , "The Apostasia genome and the evolution of orchids.", <i>Nature</i> , 2017 Sep 13;549(7672):379-383
Submission	Registration date: 2-Feb-2016 Shenzhen Key Laboratory for Orchid Conservation and Utilization The National Orchid Conservation Center of China
Relevance	Evolution

Project Data:

Resource Name	Number of Links
SEQUENCE DATA	
Nucleotide (total)	2986
WGS master	1
SRA Experiments	45
Protein Sequences	21743
PUBLICATIONS	
PubMed	1
OTHER DATASETS	
BioSample	38

▼ SRA Data Details

Parameter	Value
Data volume, Gbases	467
Data volume, Tbytes	0.24

ADDITIONAL USEFUL DATABASES – NCBI

NCBI (via Entrez platform) allows accessing 39 databases:

NCBI Resources How To Sign in to NCBI

Search NCBI databases

Literature

- Books books and reports
- MeSH ontology used for PubMed indexing
- NLM Catalog books, journals and more in the NLM Collections
- PubMed scientific & medical abstracts/citations
- PubMed Central full-text journal articles

Health

- ClinVar human variations of clinical significance
- dbGaP genotype/phenotype interaction studies
- GTR genetic testing registry
- MedGen medical genetics literature and links
- OMIM online mendelian inheritance in man
- PubMed Health clinical effectiveness, disease and drug reports

Genomes

- Assembly genome assembly information
- BioCollections museum, herbaria, and other biorepository collections
- BioProject biological projects providing data to NCBI
- BioSample descriptions of biological source materials
- Clone genomic and cDNA clones
- dbVar genome structural variation studies
- Genome genome sequencing projects by organism
- GSS genome survey sequences
- Nucleotide DNA and RNA sequences
- Probe sequence-based probes and primers
- SNP short genetic variations
- SRA high-throughput DNA and RNA sequence read archive
- Taxonomy taxonomic classification and nomenclature catalog

Genes

- EST
- Gene
- GEO DataSets
- GEO Profiles
- HomoloGene

PopSet

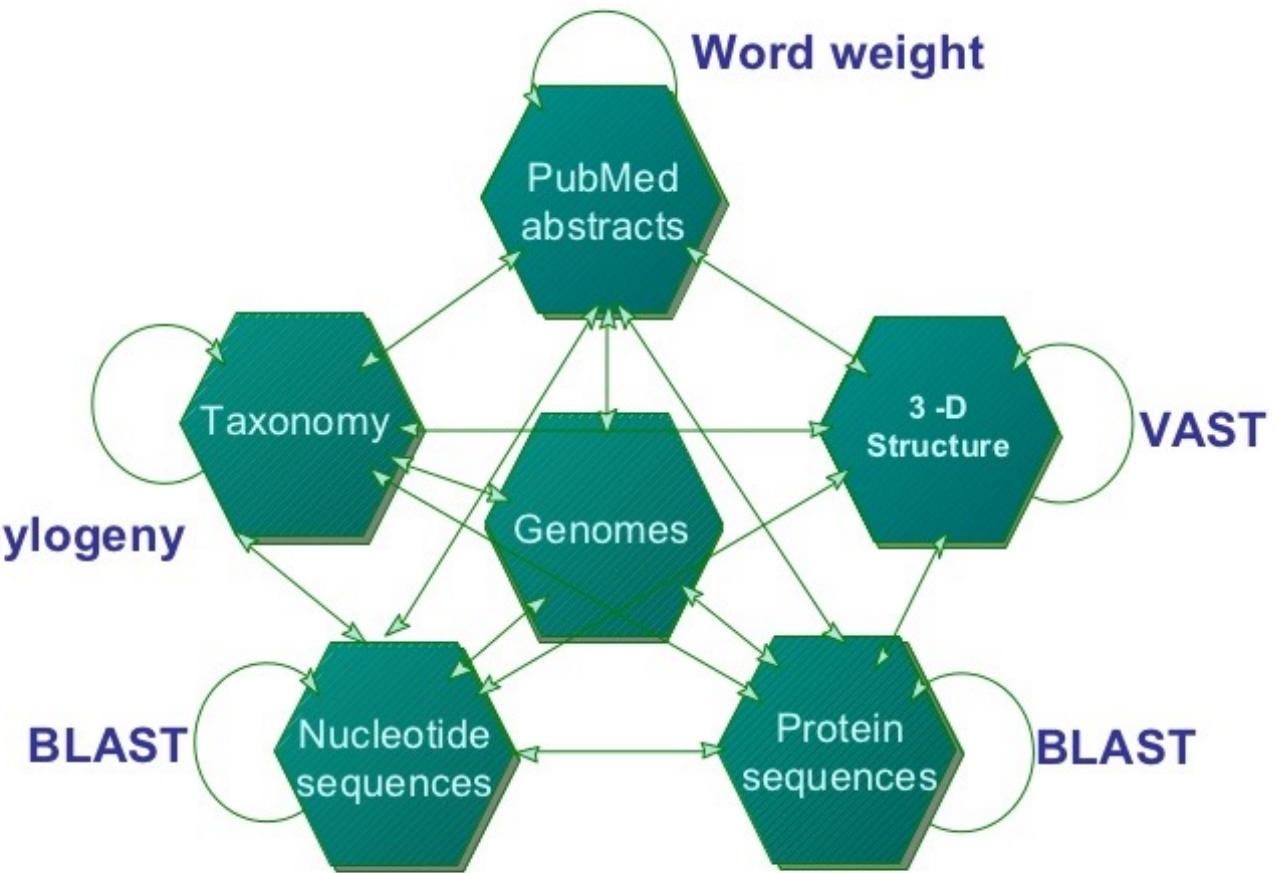
UniGene

Proteins

- Conserved Domains
- Protein
- Protein Clusters
- Structure

Chemicals

- BioSystems
- PubChem BioAssay
- PubChem Compound
- PubChem Substance



ASSEMBLY

- The new [Assembly Archive](#) at NCBI is a **repository of fully and partially complete genomic assemblies** that exists in association with sequence submissions in GenBank and trace submissions in the NCBI Trace Archive.
- The repository provides users with the **ability to access and evaluate the assemblies** from which finished genomic nucleotide sequence has been derived.
- Many benefits accrue to users of this data, including for example the **ability to determine that a spurious frame shift has occurred, or that a putative SNP is not well supported by adequate coverage**.

ASSEMBLY

NCBI Resources How To Sign in to NCBI

Sequence Set Browser Show help

Project: PEFY01 Search List of all Projects

PEFY00000000.1 Apostasia shenzhenica

Master Contigs Download History

# of Contigs:	12,380	The Apostasia shenzhenica whole genome shotgun (WGS) project has the project accession PEFY00000000. This version of the project (01) has the accession number PEFY01000000, and consists of sequences PEFY01000001-PEFY01012380.
# of Proteins:	21,743	##Genome-Assembly-Data-START##
# of Scaffolds/Chrs:	2,985	Assembly Method : AllPaths v. 49292; Pbjelly v. 14.1; fragscaff v. 140324
Total length:	322,899,837 bp	Genome Representation : Full
BioProject:	PRJNA310678	Expected Final Version: Yes
BioSample:	SAMN04453324	Genome Coverage : 455.49x
Keywords:	WGS	Sequencing Technology : Illumina; PacBio
Annotation:	Scaffolds	##Genome-Assembly-Data-END##
Organism:	Apostasia shenzhenica – show lineage	
Biosource:	/country = China: Shenzhen /ecotype = Shenzhen /isolate = ASH160606 /mol_type = genomic /tissue_type = stem; leaf	
WGS:	PEFY01000001:PEFY01012380	
Scaffolds:	KZ451883:KZ454867 2,985 scaffolds, 21,743 proteins, total length is 348,733,136 bases	
Reference:	The Apostasia genome and the evolution of orchids : Nature 549 (7672), 379-383 (2017) – show 35 authors	
Submission:	Submitted (25-OCT-2017) Shenzhen Key Laboratory for Orchid Conservation and Utilization, The National Orchid Conservation Center of China, Wangtong Road, Shenzhen 518114, China – Liu,Z.-J.	

GENOME

- This resource organizes information on genomes including sequences, maps, chromosomes, assemblies, and annotations.
- You can download genomes directly using this [ftp](#) or browse by [organisms](#).

Acanthisitta chloris (rifleman)

Acanthisitta chloris overview

Lineage: Eukaryota[2661]; Metazoa[876]; Chordata[376]; Craniata[368]; Vertebrata[368]; Euteleostomi[362]; Archelosauria[104]; Archosauria[99]; Dinosauria[95]; Saurischia[95]; Theropoda[95]; Coelurosauria[95]; Aves[95]; Neognathae[92]; Passeriformes[23]; Acanthisittidae[1]; Acanthisitta[1]; Acanthisitta chloris[1]

The rifleman (*Acanthisitta chloris*) is a small suboscine passerine bird native to New Zealand. Perching birds (i.e., passerines) are usually divided into the songbirds (oscines) and non-songbirds (suboscines) based, in part, on differences in the anatomy of the tracheal structures used to produce sound. The rifleman is thought to belong to an [More...](#)

Summary

Submitter:	BGI
Assembly level:	Scaffold
Assembly:	GCA_000695815.1 ASM69581v1 scaffolds: 53,875 contigs: 120,312 N50: 20,602 L50: 14,656
BioProjects:	PRJNA253841, PRJNA212877
Whole Genome Shotgun (WGS):	INSDC: JJRS00000000.1
Statistics:	total length (Mb): 1035.88 protein count: 16077 GC%: 41.6
NCBI Annotation Release:	100

Publications

- Comparative genomics reveals insights into avian genome evolution and adaptation. Zhang G, et al. Science 2014 Dec 12
- tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Lowe TM, et al. Nucleic Acids Res 1997 Mar 1

Genome Assembly Annotation

Loc	Type	Name	RefSeq	INSDC	Size (Mb)	GC%	Protein	tRNA	Other RNA	Gene	Pseudogene
	master WGS		-	JJRS00000000.1	1,031.1	41.6	2,369	-	-	2,713	146

Genome Region

[Acanthisitta chloris isolate BGI_N310 unplaced genomic scaffold, ASM69581v1 scaffold1876, whole genome shotgun sequence](#)

Go to nucleotide: [Graphics](#) [FASTA](#) [GenBank](#)

