

DNAMAN

Sequence Analysis Software

*Sequence
Database*

BIOINFORMATICS PLATFORM

Sequence Database

- ▶ DNAMAN Database
- ▶ Database Engines
- ▶ Database Setup
- ▶ Database Functions
- ▶ Shared Database

DNAMAN Database

- ▶ Designed to manage customer sequence data
- ▶ Not to replace large public databases
- ▶ Common database management system to provide performance, scalability and reliability
- ▶ Database can be administrated by dedicated users
- ▶ Built-in user-level security
- ▶ Compatible with third party database tools

Database Engines

▶ SQLite3

- ▶ Available to all DNAMAN user licenses
- ▶ Self-contained, high-reliability, full-featured, and in public-domain
- ▶ Not Server-Client based

▶ MySQL

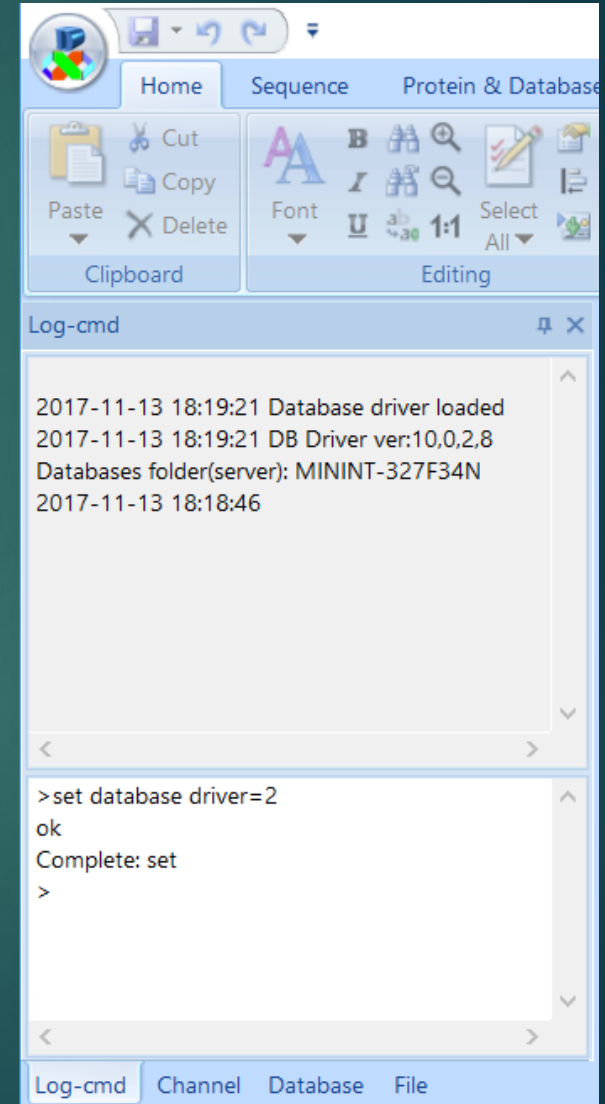
- ▶ Available to Institutional user licenses
- ▶ Open-Source, high performance and reliability

▶ Microsoft SQL Server

- ▶ Available to Institutional user licenses
- ▶ Microsoft proprietary

Database Setup

- ▶ SQLite3 is the default database engine in DNAMAN
- ▶ Institutional users may opt to MySQL or MS SQL using command line operations
 - ▶ MySQL: set database driver=2
 - ▶ MS SQL: set database driver=3
- ▶ Turn off sequence compression
 - ▶ set database compression=0
- ▶ Turn on database logs
 - ▶ set database enablelog=1



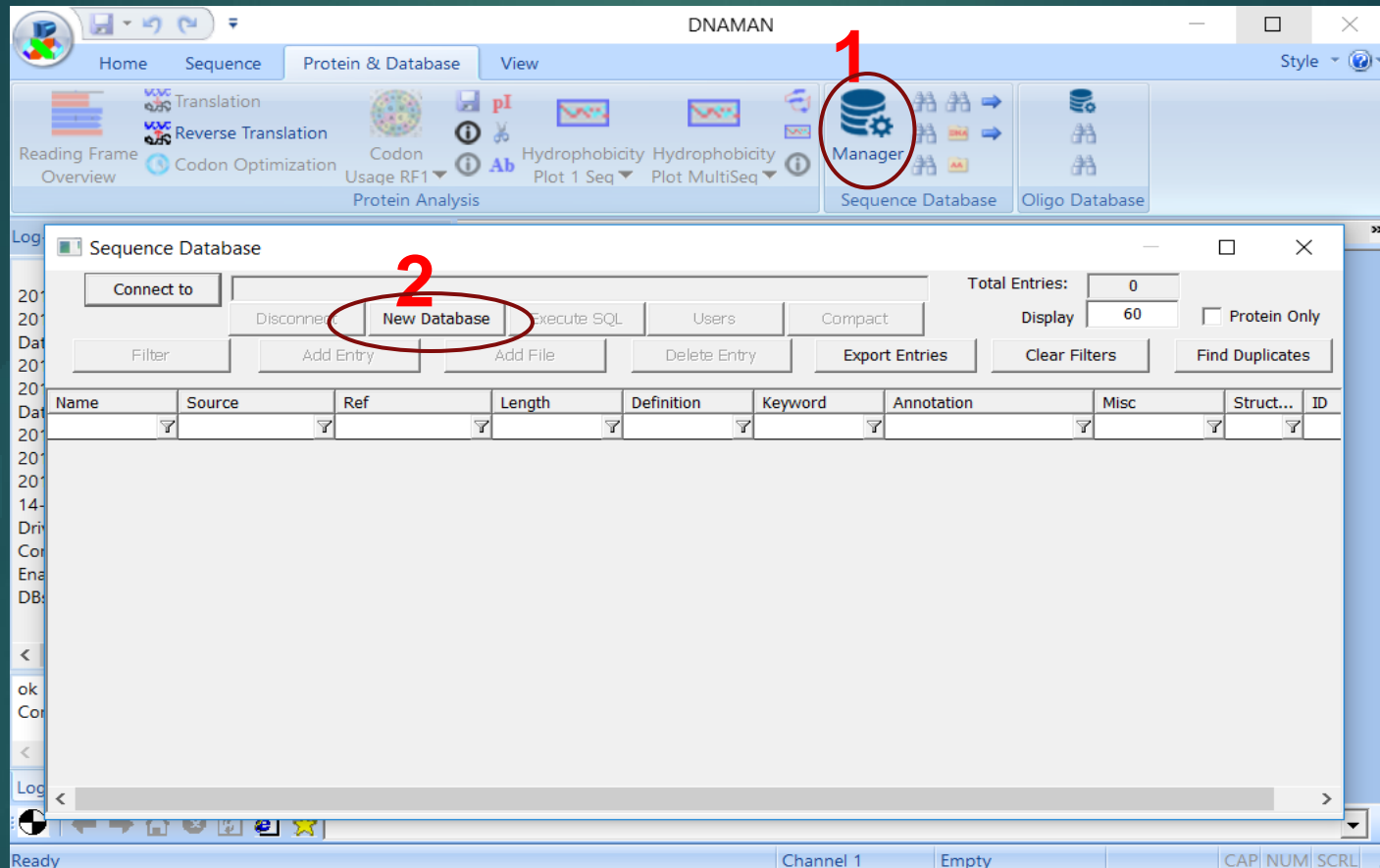
Sequence Database

- SQLite Database Structure

- ▶ One data table: DNAMAN Seq DB
- ▶ 8 TEXT fields including "Name", "Source", "Ref", "Misc", "Definition", "Keyword", "Annotation", "cds"
- ▶ 3 INT fields including "Type", "Length" and "NumAnnotation"
- ▶ 3 LongBinary fields including "Pwd", "Structure" and "Sequence"

Database Functions

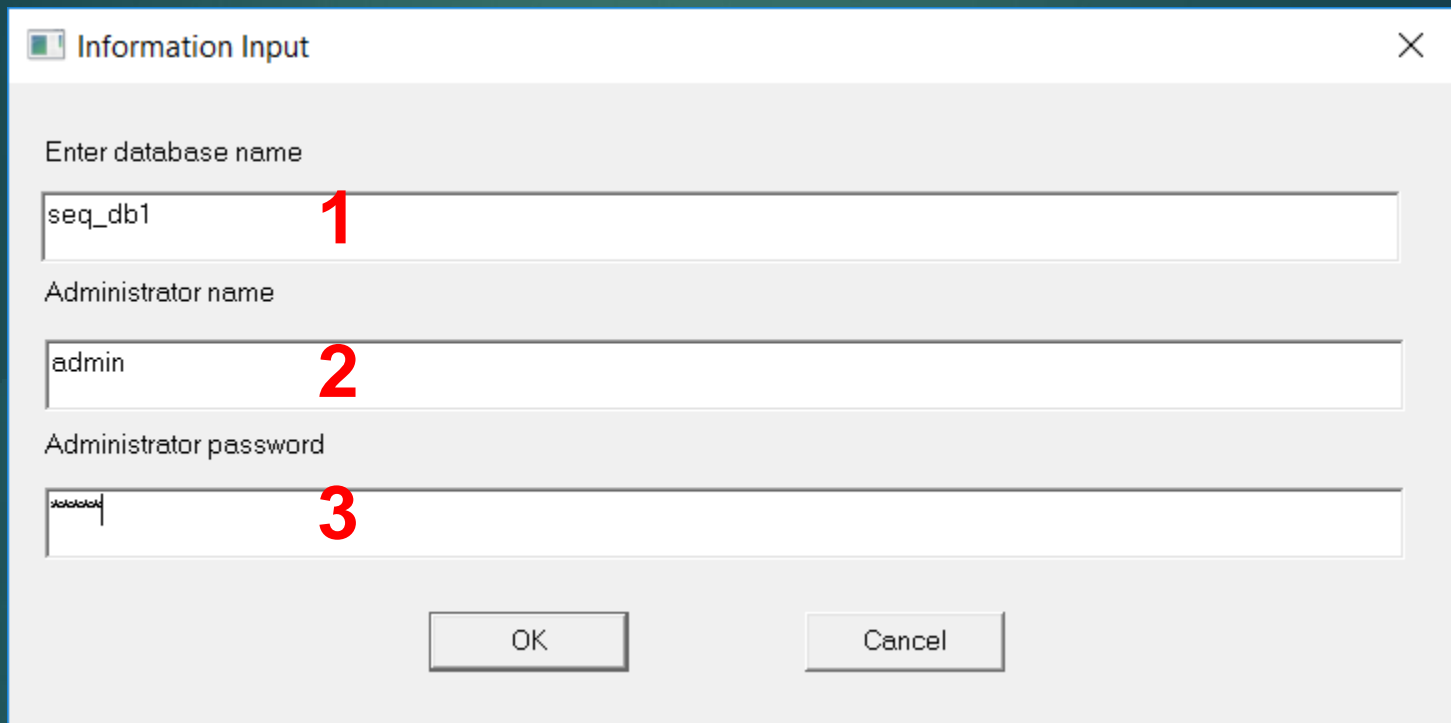
- Create New Database



1. Click Database Manager in Sequence Database tab
2. Click New Database button

Create New Database

- SQLite Database



Information Input

Enter database name

seq_db1 1

Administrator name

admin 2

Administrator password

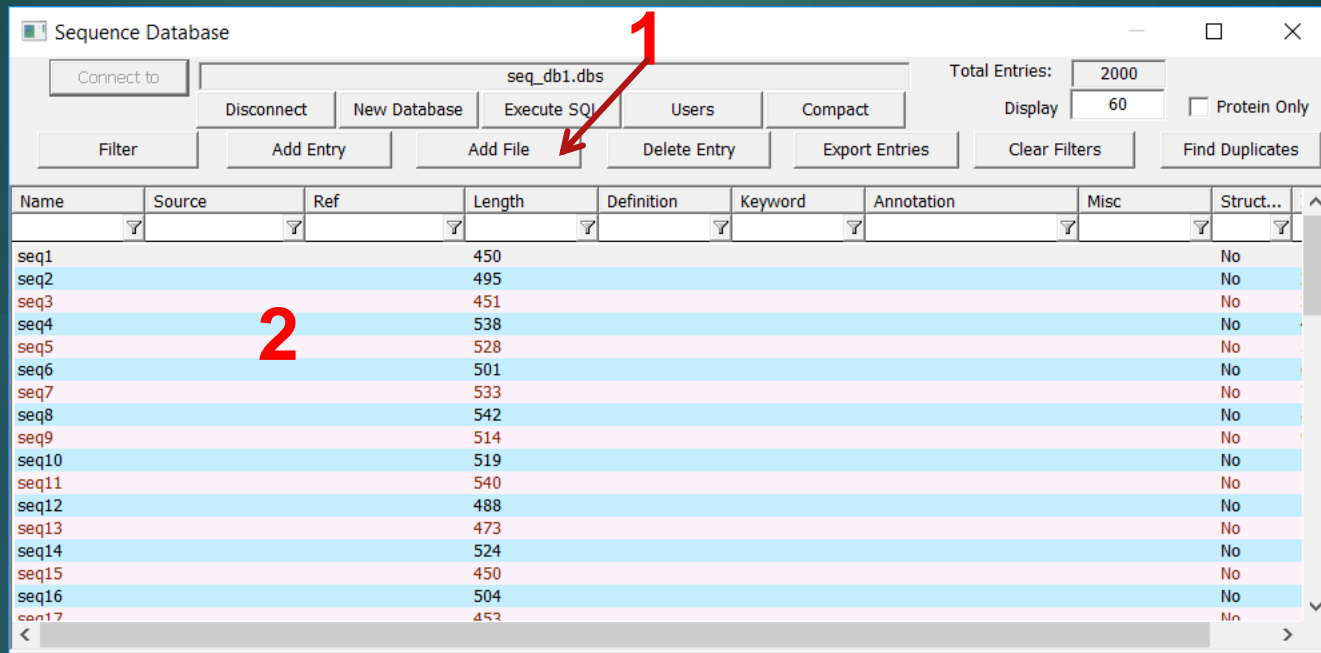
XXXXXXXXXX 3

OK Cancel

1. Enter database name
2. Enter Administrator name, normally the user name of current computer
3. Enter password. Leave it blank if no user control is needed.

Add Sequence Records

- SQLite Database



1. Press Add File button
 2. Select "AssemblySamples.seq" from DNAMAN folder and wait for all 2000 sequences added to database
- Multiple sequences can be retrieved from Genbank and added to database. Genbank is preferred format and annotations/references can be kept in database

Edit Sequence Records

- SQLite Database

1. Double click a record in sequence list
2. Edit record information

Only users with administrative privileges have the right to edit records if password has been set when creating database

The screenshot shows the 'Sequence Database' application interface. At the top, there is a toolbar with buttons for 'Connect to', 'Disconnect', 'New Database', 'Execute SQL', 'Users', 'Compact', 'Filter', 'Add Entry', 'Add File', 'Delete Entry', 'Export Entries', 'Clear Filters', and 'Find Duplicates'. The 'vectors_puc18.dbs' database is selected, showing 30 total entries and a display of 60. A table lists sequence records with columns for Name, Source, Ref, Length, Definition, Keyword, Annotation, Misc, and Struct... The record 'AB593377' is selected. A red arrow points to the 'Add Entry' button. A 'DNAMAN Database Entry' dialog box is open, showing the details for 'AB593379'. The dialog has fields for Name, Definition, Keywords, Source, Reference, Ref Title, Comment, and Annotations. The 'Circular DNA' radio button is selected. The 'Sequence' field shows a DNA sequence starting with '00001 TCGCGCGTTT CGGTGATGAC GGTGAAAACC TCTGACACAT GCAGCTCCCG GAGACGGTCA CAGCTTGCTC...'. A red number '2' is placed over the 'Keywords' field.

Name	Source	Ref	Length	Definition	Keyword	Annotation	Misc	Struct...
AB593379	piggyBac donor ...	<a="Mitsutake,H...	8127	PiggyBac don...	CIRCULAR .	<t="source" id="" pos=...		No
AB593378								No
AB593377								No
EU234497								No
EU234496								No
KX357893								No
AY737006								No
AY737005								No
AY737004								No
AY619005								No
AY619004								No
AY599233								No
AY599231								No
AY599228								No
AY599227								No
AY599226								No
F11106640								No

DNAMAN Database Entry

Name: AB593379 Linear DNA Circular DNA Protein

Definition: PiggyBac donor vector pPIGA3Fluc DNA, complete sequence.

Keywords: CIRCULAR .

Source: piggyBac donor vector pPIGA3Fluc

Reference: <a="Mitsutake,H. and Kobayashi,J." j="Unpublished"><a="Mitsutake,H. a

Ref Title: piggyBac-mediated stable transformation of cultured Bombyx mori

Comment:

Annotations: <t="source" id="" pos="0 8126" src="piggyBac_donor"><t="misc_feature" id="" pos="0 403" att="cloning_vector;"><t="repeat_region" id="" pos="555 590" att="piggyBac_5'_inv;"><t="gene" id="hsp70" pos="1372 1764"><t="regulatory" id="hsp70" pos="1372 1764" att="Hsp70Bb;"><t="gene" id="pac" pos="1863 2462"><t="cds" id="puromycin_N-ace" pos="1863 2462"><t="gene" id="PGK" pos="2516 2821"><t="misc_feature" id="PGK" pos="2516 2821" att="PGK_polyadenyla;"><t="gene" id="A3" pos="2839 3494"><t="regulatory" id="A3" pos="2839 2965" att="actin_A3;"><t="exon" id="A3" pos="2966 3034" att="actin_A3;"><t="intron" id="A3" pos="3035 3494" att="actin_A3;"><t="gene" id="luc" pos="3524 5176"><t="cds" id="luciferase" pos="3524 5176"

Sequence:

```
00001 TCGCGCGTTT CGGTGATGAC GGTGAAAACC TCTGACACAT GCAGCTCCCG GAGACGGTCA CAGCTTGCTC
00071 GTAAGCGGAT GCCGGGAGCA GACAAGCCCG TCAGGGCGCG TCAGCGGGTG TTGGCGGGTG TCGGGGCTGG
00141 CTTAACTATG CGGCATCAGA GCAGATTGTA CTGAGAGTGC ACCATATGCG GTGTGAAATA CCGCACAGAT
00211 GCGTAAGGAG AAAATACCGC ATCAGGCGCC ATTCCGCCATT CAGGCTGGCC AACTGTTGGG AAGGGCGATC
00281 GGTGCGGGCC TCTTCGCTAT TACGCCAGTC GGCGAAAAGG GAGATGTGCTG CAAGGCGATT AAGTTGGGTA
00351 ACGCCAGGGT TTTCOCAGTC ACGACGTTGT AAAACGACGG CCAAGTGCCAA GCTTGCATCG CTCACGGTCC
00421 ACGCTCGCGC GACTTGGTIT GCCATTCTTT AGCGCGCGTC CGGTACACCA GCTTGGCCAC AATGTGGTIT
00491 TTGTCAAACG AAGATTCTAT GACGTGTTTA AAGTTTAGGT CGAGTAAAGC GCAAAATCTTT TTTAACCCTA
00561 GAAAGATAGT CTGCGTAAAA TTGACGCATG CATTCTTGAA ATATTGCTCT CTCTTCTAA ATAGCGCGAA
00631 TCCGTCGCTG TGCATTTAGG ACATCTCAGT CGCGCGTTGG AGCTCCCGTG AGGCGTGCTT GTCAATGCGG
```

OK Exit

Edit Database User List

- SQLite Database

The screenshot shows the 'Sequence Database' application window. The 'Users' button is highlighted with a red arrow and the number '1'. A dialog box titled 'Database Users' is open, showing a table with columns 'User' and 'Write Permission'. The dialog box is highlighted with a red '2'. The main window displays a list of sequence entries with columns: Name, Source, Ref, Length, Definition, Keyword, Annotation, Misc, and Struct... The 'Total Entries' is 30 and 'Display' is 60. The 'Protein Only' checkbox is unchecked.

Name	Source	Ref	Length	Definition	Keyword	Annotation	Misc	Struct...
AB593379	piggyBac donor ...	<a=				source" id="" pos=...	No	
AB593378	piggyBac donor ...	<a=				source" id="" pos=...	No	
AB593377	piggyBac donor ...	<a=				source" id="" pos=...	No	
EU234497	Cosmid vector p...	<a=				source" id="" pos=...	No	
EU234496	Cosmid vector p...	<a=				source" id="" pos=...	No	
KX357893	Cloning vector p...	<a=				source" id="" pos=... #Assembly...	No	
AY737006	Cloning vector p...	<a=				source" id="" pos=...	No	
AY737005	Cloning vector p...	<a=				source" id="" pos=...	No	
AY737004	Cloning vector p...	<a=				source" id="" pos=...	No	
AY619005	Cloning vector p...	<a=				source" id="" pos=...	No	
AY619004	Cloning vector p...	<a=				source" id="" pos=...	No	
AY599233	Cloning vector p...	<a=				source" id="" pos=...	No	
AY599231	Cloning vector p...	<a=				source" id="" pos=...	No	
AY599228	Cloning vector p...	<a=				source" id="" pos=...	No	
AY599227	Cloning vector p...	<a=				source" id="" pos=...	No	
AY599226	Cloning vector p...	<a=				source" id="" pos=...	No	
EU1196040	Transformation	<a=				source" id="" pos=...	No	

1. Click Users button

2. Edit user list

Only users with write permission have the rights to edit user list if password has been set when creating database

Use Database Sequence

1. Double click a record in Database control
2. Sequence opened and loaded into channel for analysis

The screenshot shows the DNAMAN software interface. The top menu bar includes Home, Sequence, Protein & Database, and View. The toolbar contains various icons for sequence analysis, including Blastn, Blastn RF1, Blastn Pro, SIRNA, Seq Fold, Align Multiple, Seq Assembly, Enzyme Cut, Restriction Analysis, Oligo Seq, and PCR Primer.

The Database control panel on the left displays a table with the following data:

Name	Def	Length
DQ457004	Conjugative vector p...	4664
HE584556	Cloning vector pAJM-...	6045
AF311601	Reporter vector pJDL...	5847
AJ537514	Synthetic construct pl...	10828
Y07862	Cloning vector pEAV0...	15528
SYCCMA	Synechocystis sp. PC...	2393
AB593379	PiggyBac donor vecto...	8127
AB593378	PiggyBac donor vecto...	7172
AB593377	PiggyBac donor vecto...	9375
EU234497	Cosmid vector pPIBT...	7384
EU234496	Cosmid vector pPIBT...	7436
KX357893	Cloning vector pUC1...	12082
AY737006	Cloning vector pUC1...	5874
AY737005	Cloning vector pUC1...	4136
AY737004	Cloning vector pUC1...	6078
AY619005	Clonina vector pUC1...	8347

A red arrow labeled '1' points to the record HE584556. The right pane shows the sequence details for HE584556, including source information, misc_features, gene, and cds. A red number '2' is placed next to the sequence details. The sequence is displayed in the ORIGIN section:

```
ORIGIN
1      TGGTGTAAC AAATGACGC TTAGACAACT TAATAACACA TTGCGGACGT TTTTCGGATC
61     TGCCCGATCT AGTAACATAG ATGACACCGC GCGCGATAAT TTATCCTAGT TTGCGCGCTA
121    TATTTTGTTC TCTATCGCGT ATTAATATGTA TAATTTGCGGG ACTCTAATCA TAAAAACCCA
181    TCTCATAAAT AACGTCATGC ATTACATGTT AATTATTACA TGCTTAACGT AATTCAACAG
241    AAATTATATG ATAATCATCG CAAGACCGGC AACAGGATTC AATCTTAAGA AACTTTAATG
301    CCAAAATGTT GAACGATCGG GGATCCCGGA CGAGTGCTGG GCGTCGGTTC TCCACTATCG
```

Sequence Database

- MYSQL Database Structure

- ▶ One data table: DNAMAN_Seq_DB
- ▶ 8 TEXT fields including "Name", "Source", "Ref", "Misc", "Definition", "Keyword", "Annotation", "cds"
- ▶ 4 INT fields including "ROWID", "Type", "Length" and "NumAnnotation"
- ▶ 1 BLOB "Structure" and 1 LONGBLOB "Sequence"
- ▶ One log table if enabled: DNAMAN_Log_Msg

Sequence Database

- MySQL Database Management

- ▶ MySQL databases are managed by local administrator. DNAMAN is not responsible for server side operations.
- ▶ MySQL database administrator is responsible to create users and assign read-write rights to users
- ▶ Database logs reside on server. DNAMAN may not track all logs

Sequence Database

- Microsoft SQL Database Structure

- ▶ One data table: DNAMAN_Seq_DB
- ▶ 8 fields of varchar(max) including "Name", "Source", "Ref", "Misc", "Definition", "Keyword", "Annotation", "cds"
- ▶ 4 INT fields including "ROWID", "Type", "Length" and "NumAnnotation"
- ▶ 2 fields of varbinary(max) including "Structure" and "Sequence"
- ▶ One log table if enabled: DNAMAN_Log_Msg

Sequence Database

- MS SQL Database Management

- ▶ Microsoft SQL databases are managed by local administrator. DNAMAN is not responsible for server side operations.
- ▶ Microsoft SQL database administrator is responsible to create users and assign read-write rights to users
- ▶ Database logs reside on server. DNAMAN may not track all logs

Shared Database

- ▶ MySQL and Microsoft SQL databases are Server-Client based, so databases can be shared natively
- ▶ SQLite is not Server-Client based, so database sharing may be restricted and performance may decrease with simultaneous access by a large number of users

Shared Database

- SQLite database

- ▶ Create a folder on shared drive: eg T:\shared_database
- ▶ Point DNAMAN database folder to the shared folder

1. Press Settings in Information tab
2. Enter database folder on shared drive
3. Click OK to save the change

Make sure shared drive is accessible by users

