# The StreptomeDB 2.0

## Knowledge database of secondary metabolites produced by streptomycetes

**Klementz D.; Döring K.; Lucas X.; Telukunta K.; Thomas O.; Günther S.**

*Pharmaceutical Bioinformatics, Institute of Pharmaceutical Sciences, University of Freiburg - Germany*

dennis.klementz@pharmazie.uni-freiburg.de; kersten.doering@pharmazie.uni-freiburg.de

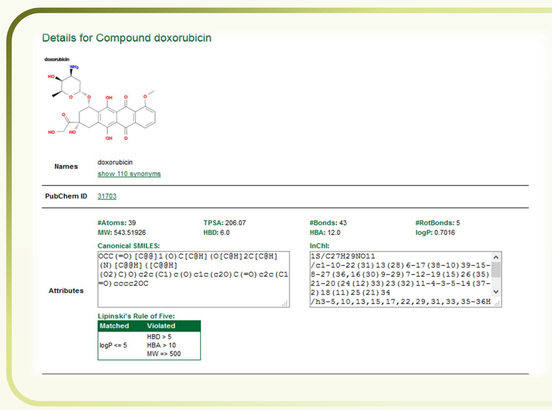
## Overview

**Streptomycetes** are one of the most relevant *genera* for pharmaceutical research since 1943, when Albert Schatz isolated streptomycin for the first time[1]. Nowadays, over 60% of all known antibiotics are produced by these bacteria[2]. An **overwhelming amount of data** concerning streptomycetes has been produced in the past 70 years, but it has never been offered in a **centralized**, **freely accessible** database until the publication of the first version of **StreptomeDB** in January 2013[2]. Since that date, we could increase the initial amount of molecules in our database from 2,400 to 4,000 compounds. In addition, we have included several **new features** such as the integration of **genomic and phylogenetic data**, an advanced **scaffold-based navigation system**, and a comprehensive **literature collection** with specialized search options.(url: http://www.pharmaceutical-bioinformatics.de/streptomedb/)

*Albert Schatz [The Guardian magazin]*

## Compounds

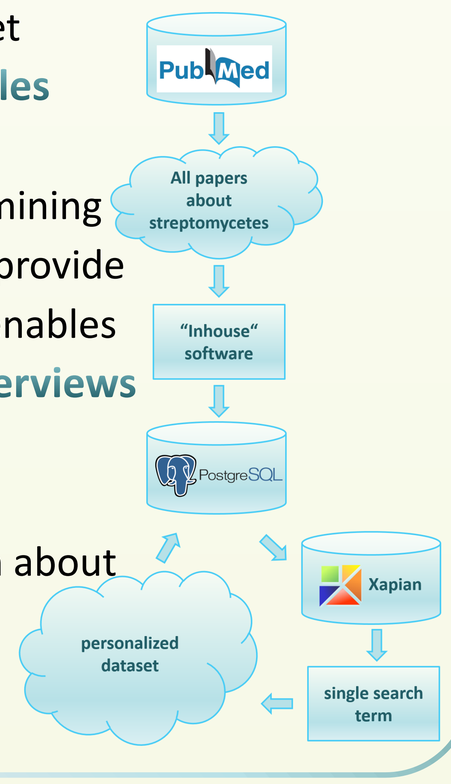
StreptomeDB is currently the largest collection of **natural compounds** produced by streptomycetes, containing a plethora of recently discovered structures, of which over 1,000 cannot yet be found in commonly used databases such as PubChem. Additionally to the vast amount of **curated data**, it offers extensive **background information** for all included compounds such as chemical structures or predicted chemical properties.

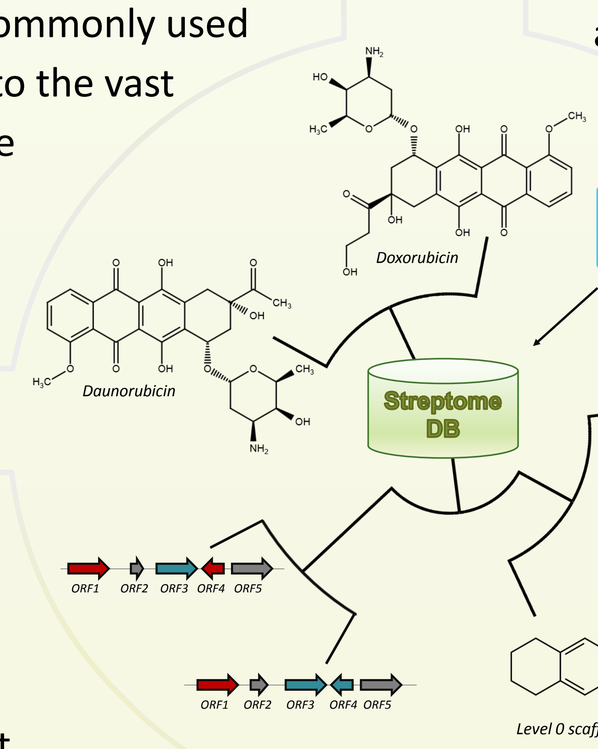



*Example of the current StreptomeDB*

## Literature

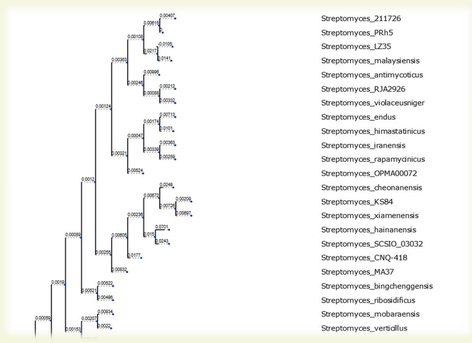
To provide further information, a set of more than **18,000 full text articles** concerning streptomycetes will be integrated in the database by text mining methods. Additionally, we want to provide a **multilayer search system**, that enables the generation of **personalized overviews** and **datasets,** including advanced features such as relevance prediction, information about authors, publication year, or number of citations.




## Genomics

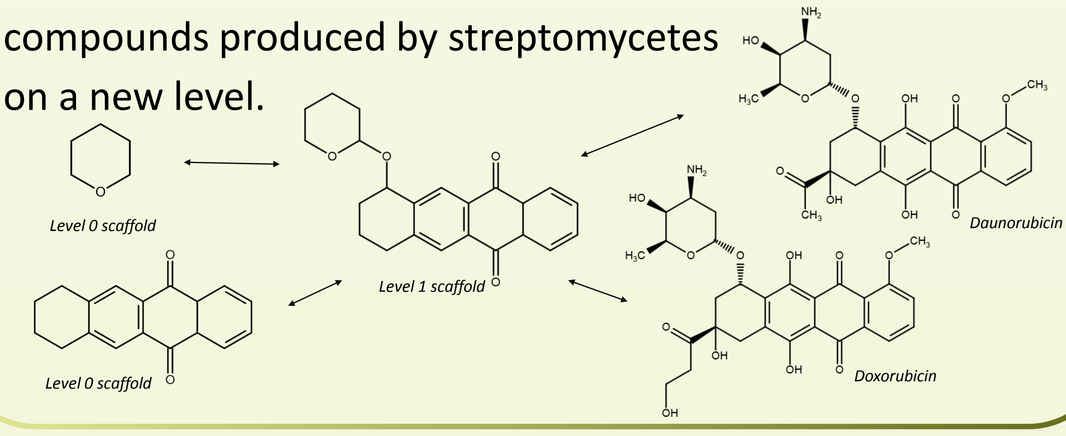
We have integrated a comprehensive collection of **genomic data** for the strains contained in StreptomeDB. It includes 489 gene clusters and all available full and draft genomes. The StreptomeDB offers an **interactive phylogenetic tree,** which is representative for about 1,500 strains. For example, it allows for the visualization of the frequency of a given scaffold in an evolutionary context.




*Phylogenetic tree of Streptomyces 16S rRNA*

## Scaffolds

Chemical scaffolds play a major role in computational high-throughout **screening methods**[3], *and* drug discovery[4]. With an advanced **scaffold search system**, we will offer the possibility to explore the **chemical diversity** of natural compounds produced by streptomycetes on a new level.




## Conclusion

StreptomeDB comprises an **extensive collection** of data related to streptomycetes, with a focus on **natural compounds**. The presented upgrades will allow for an advanced exploration of the **chemical diversity** of compounds produced by streptomycetes and a deeper understanding of their **evolutionary background**.

**Pharmazeutische Bioinformatik**

www.pharmaceutical-bioinformatics.de

JUNIOR RESEARCH GROUP PROGRAMME

DFG LIS 45

-factors RTG1976

References:

[1] Schatz, A.; Bugle, E.;Waksman, S. A. Exp. Biol. Med. **1944**, 55, 66-69.

[2] Lucas, X.; Senger, C.; Erxleben, A.; Grüning, B. A.; Döring, K.; Mosch, J.; Flemming, S.; Günther, S. Nucleic Acids Res. **2013**, D1130-D1136.

[3] Shelat, A.; Guy K. Nat. Chem. Biol. **2007**, 3, 442-446

[4] Lucas, X.; Grüning BA.; Bleher S.; Günther S.; J Chem In f Model. **2015**, 55(5), 915-24

UNI FREIBURG