StreptomeDB

Update 2.0

Klementz D.; Döring K.; Lucas X.; Günther S.

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

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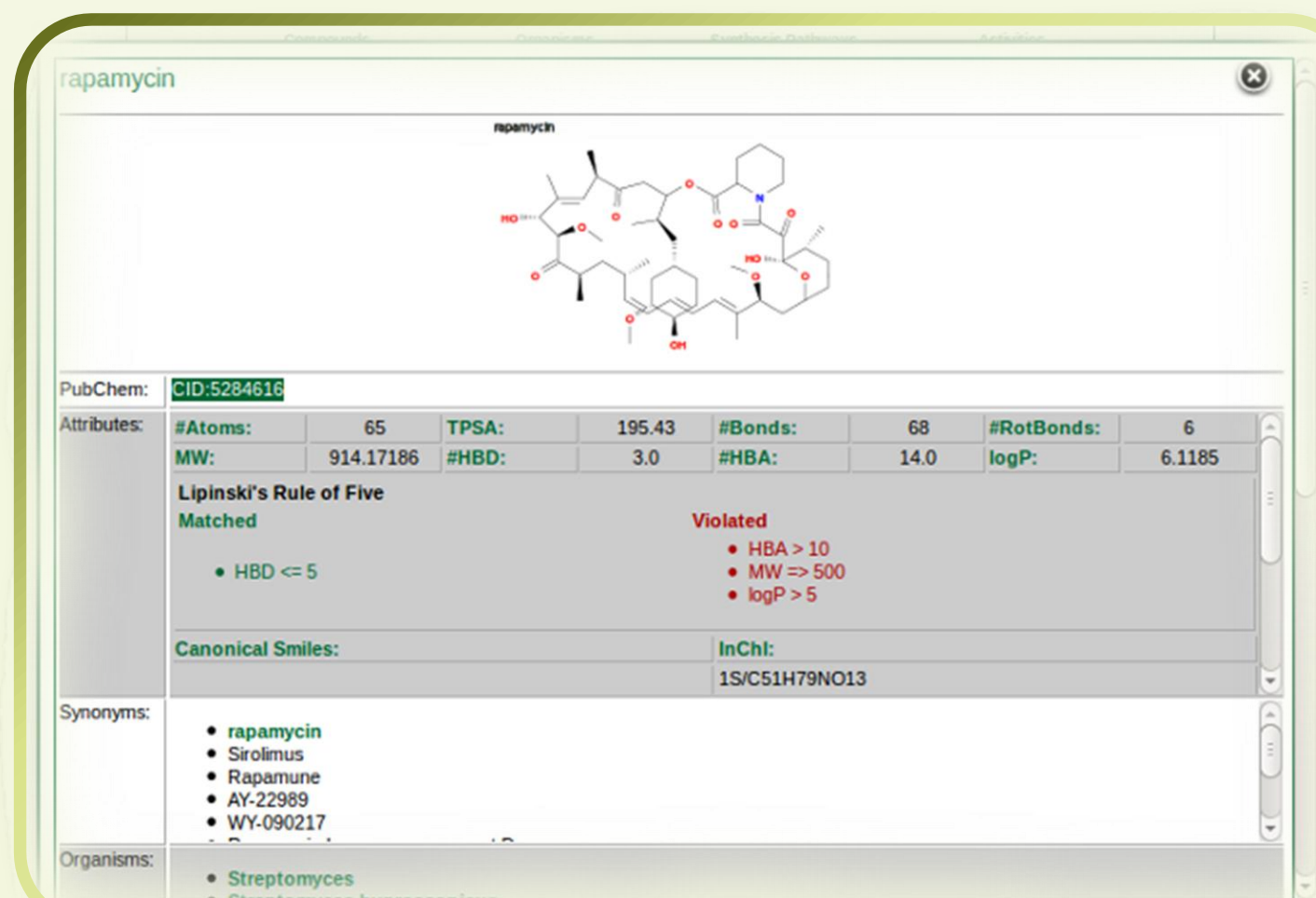
Overview

Streptomycetes became 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** and **freely-accessible** database until the publication of the first version of **StreptomeDB** in January 2013[2]. Since that date, we could increase the initial size of our database by several hundred compounds. In addition, we have included several **new features** such as the integration of **genomic and phylogenetic data**, an advanced **scaffold-based search system**, and a comprehensive **literature collection** with specialized search options.

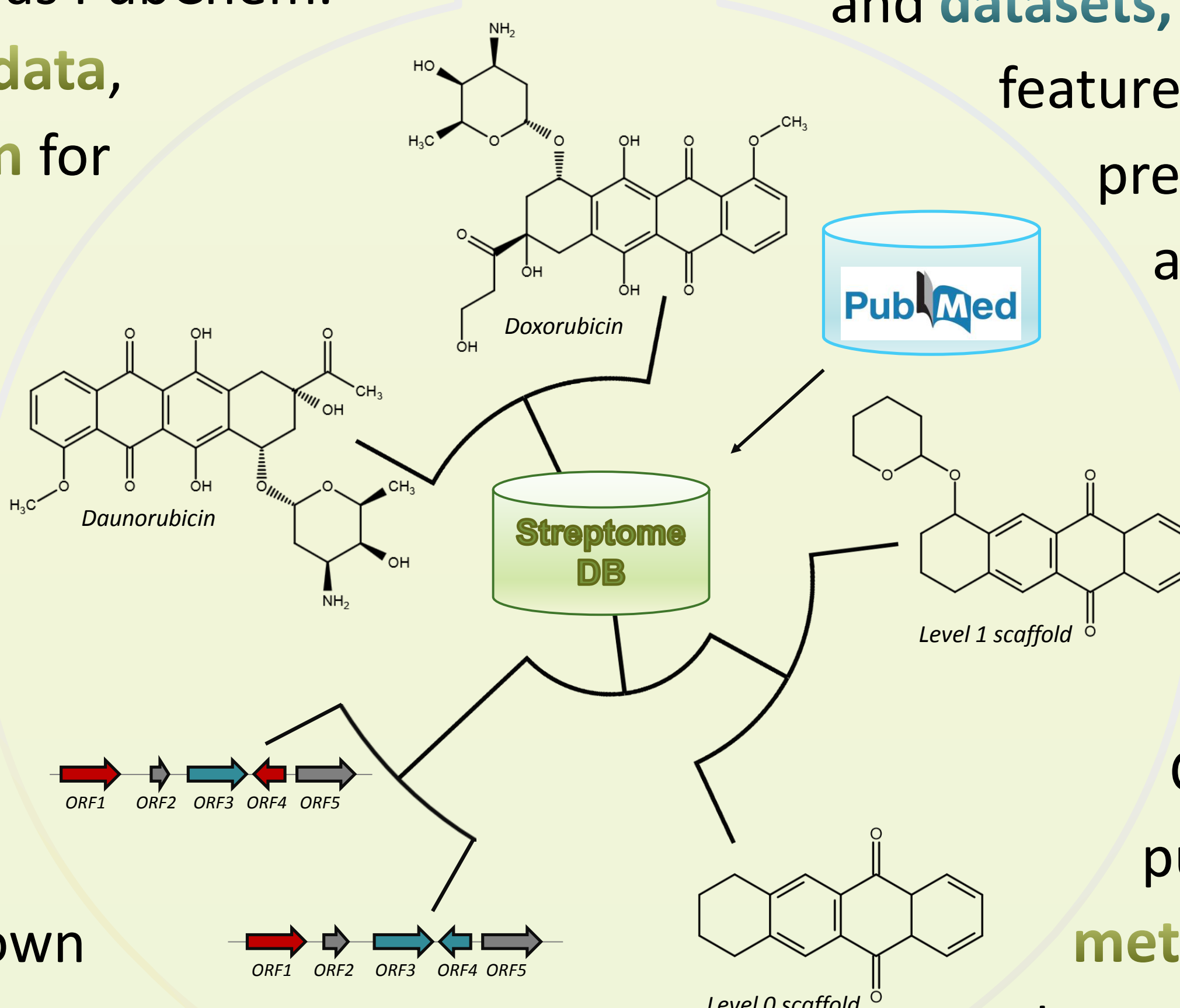*Albert Schatz [The Guardian magazin]*

Compounds

StreptomeDB is the currently largest collection of **natural compounds** produced by streptomycetes, containing plenty of recently discovered structures, which cannot be found in commonly used databases such as PubChem. Additionally to the vast amount of **curated data**, we offer extensive **background information** for all included compounds such as chemical structures, predicted chemical properties, source organisms, and many more.

Example of the current StreptomeDB

Literature

We have collected information from a set of more than **18,000 full text articles** concerning streptomycetes by text mining methods, which were integrated in our database. Additionally, we provide a **multi-layer search system**, that enables an easy 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. We included known gene clusters, all available full and draft genomes, and **preprocessed and reviewed** rRNA sequences.
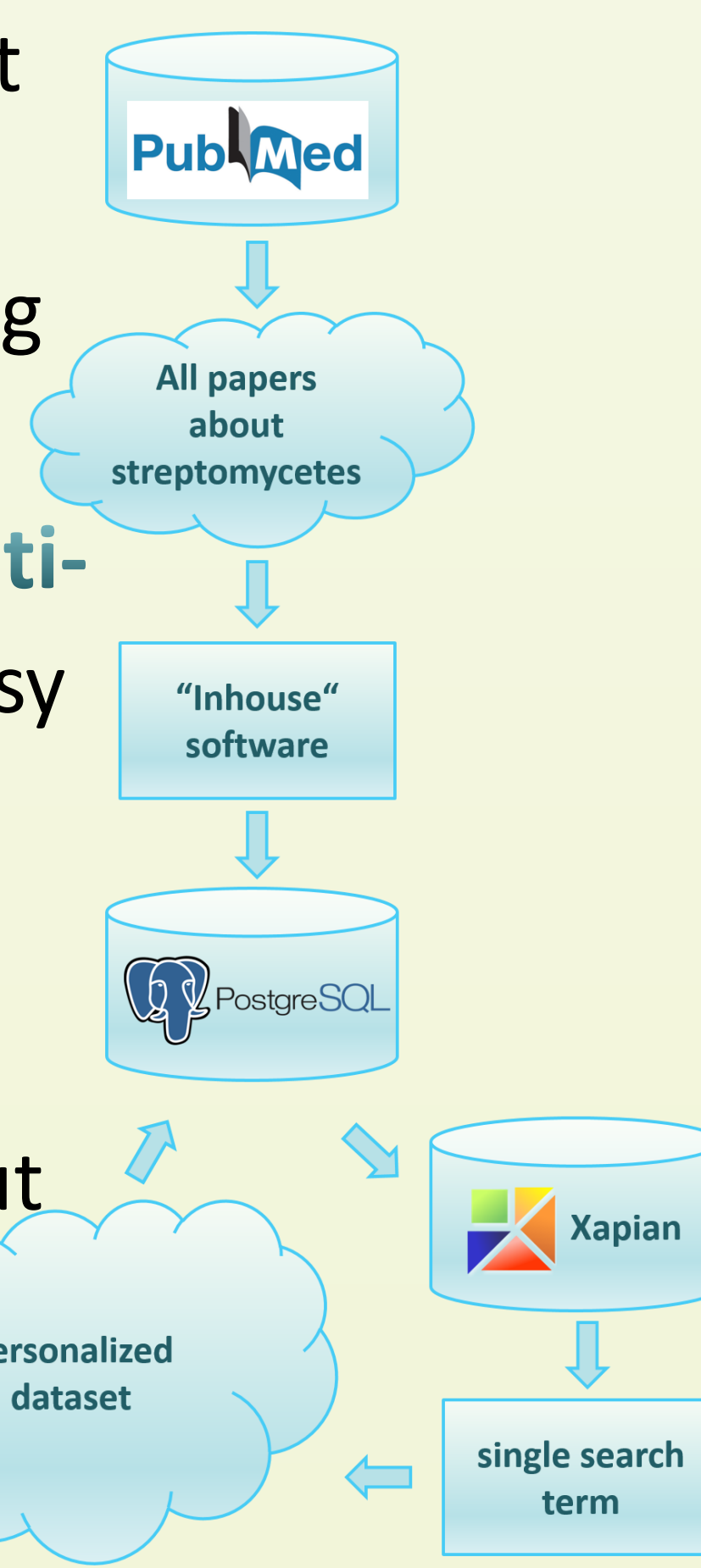
We will also offer an **interactive phylogenetic tree,** which will *e.g.* allow the visualization of the frequency of a given scaffold in an evolutionary context.
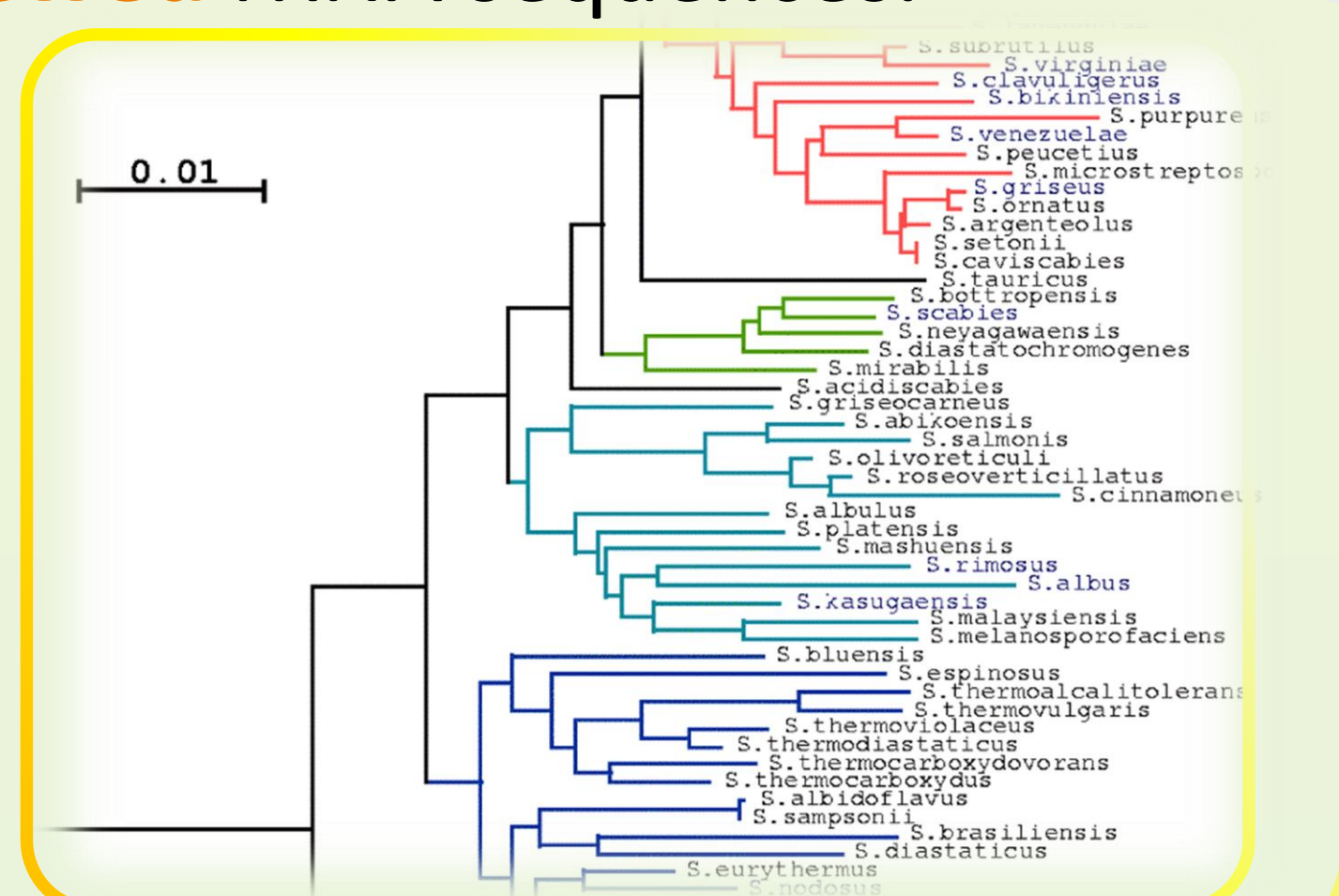
Phylogenetic tree of Streptomyces 16S rRNA [avermitilis.ls.kitasato-u.ac.jp]

Scaffolds

Chemical scaffolds play a major role in computational high-throughout **screening methods**[3], *e.g.* for drug discovery. 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.

Level 0 scaffold · *Level 1 scaffold* · *Daunorubicin* · *Doxorubicin*
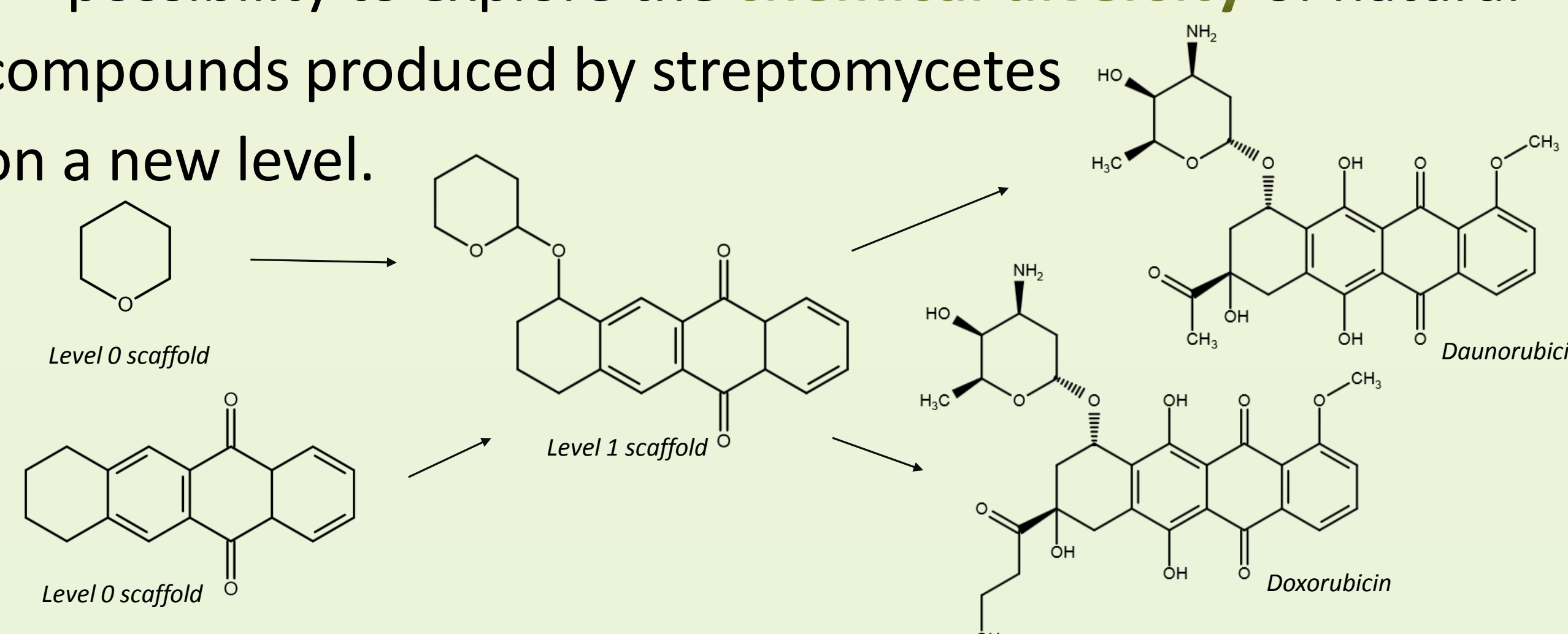
Conclusion

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

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