TopoIBase: a comprehensive database dedicated to type IA DNA-topoisomserases.

Nicolas DACCORD¹, Eduardo COREL¹,², Damien CORREIA¹,⁴, Anais LOUIS¹, Hélène DEBAT³,⁴,⁵, Vladimir DARIC⁴, Marc NADAL⁴,⁵, Claudine DEVAUCHELLE¹ and Franck SAMSON¹

¹LaMMe, UMR8071 CNRS, 23 bvd de France, Université d’Evry-Val d’Essonne IBGBI, 91037 Evry
²IBPS, UMR 7138, Université Pierre et Marie Curie, 7 quai Saint-Bernard, 75252 Paris
³Université Versailles St-Quentin, 45 avenue des Etats-Unis, 78035 Versailles
⁴IGM, UMR 8621 CNRS, Université Paris-Sud, Bât. 409, 91405 Orsay Cedex
⁵Institut Jacques Monod, UMR 7592 CNRS, Université Paris Diderot
Topoisomerases

- Proteins discovered in 1971 (Wang et al., 1971)
- Function: stabilization of the DNA topology

**Figure:** Mechanisms of action of type I and II topoisomerases.
Diversity

- Many types et sub-types of different origins
- Different activities and dependencies
- Topoisomerasases IA.

**Figure:** Types and sub-types of topoisomerasases and enzymatic activity variations.
Objectives

- Topoisomerases IA are difficult to study
- Lack of protein database dedicated to topoisomerases
- Objective: develop an expertised database dedicated to topoisomerases
- Expertise: data selection, VLD (Variable Length Decoding: Didier et al., 2012 \(^1\)) classification and annotations

---

Figure: Simplified UML diagram of the database.
Data selection and expertise

- Topoisomerasers: TBLASTN on fully sequenced genomes
- Classification: VLD (Didier et al. 2012) + hierarchical ascendant clustering
- Annotations by experts

---

Database statistics

<table>
<thead>
<tr>
<th>Kingdom</th>
<th>Number of topoisomerases</th>
</tr>
</thead>
<tbody>
<tr>
<td>Archaea</td>
<td>250</td>
</tr>
<tr>
<td>Bacteria</td>
<td>1600</td>
</tr>
<tr>
<td>Eukaryota</td>
<td>150</td>
</tr>
</tbody>
</table>

Total: 1750
Database statistics (2)

Figure: Topoisomerasases distribution in domains.
**Graphical interface**

*Figure*: Data browsing on the interface.
**Taxonomy**

*Figure:* Part of the taxonomy tree.
Introduction Database Interface Perspectives

Taxonomy

Figure: Taxon page.
Selections

Figure: Selection page.
**Figure:** BLASTP interface.
Other use-cases

- Taxa and VLD distribution of topoisomerases containing a specific amino acid sequence
- Analysis of a list of uniprot IDs of topoisomerases
- Classification of a new protein compared to stored topoisomerases
• Extension of the database to type IB and II topoisomerases
• Visualization of neighbors genomic elements
• Binding to other databases
Thank you for your attention
stat.genopole.cnrs.fr/topodb

Acknowledgements

Claudine Devauchelle and Franck Samson
Marc Nadal, Hélène Debat, Eduardo Corel, Damien Correia,
Vladimir Daric, Anaïs Louis and Florence Vogliolo