

Limbform: a functional ontology-based database of limb regeneration experiments

Daniel Lobo, Erica B. Feldman, Michelle Shah, Taylor J. Malone and Michael Levin*

Center for Regenerative and Developmental Biology, Department of Biology, Tufts University, 200 Boston Avenue, Suite 4600, Medford, MA 02155, USA

Associate Editor: Jonathan Wren

ABSTRACT

Summary: The ability of certain organisms to completely regenerate lost limbs is a fascinating process, far from solved. Despite the extraordinary published efforts during the past centuries of scientists performing amputations, transplantations and molecular experiments, no mechanistic model exists yet that can completely explain patterning during the limb regeneration process. The lack of a centralized repository to enable the efficient mining of this huge dataset is hindering the discovery of comprehensive models of limb regeneration. Here, we introduce Limbform (Limb formalization), a centralized database of published limb regeneration experiments. In contrast to natural language or text-based ontologies, Limbform is based on a functional ontology using mathematical graphs to represent unambiguously limb phenotypes and manipulation procedures. The centralized database currently contains >800 published limb regeneration experiments comprising many model organisms, including salamanders, frogs, insects, crustaceans and arachnids. The database represents an extraordinary resource for mining the existing knowledge of functional data in this field; furthermore, its mathematical nature based on a functional ontology will pave the way for artificial intelligence tools applied to the discovery of the sought-after comprehensive limb regeneration models.

Availability and implementaion: The Limbform database is freely available at <http://limbform.daniel-lobo.com>.

Contact: michael.levin@tufts.edu

Received on June 18, 2014; revised on August 3, 2014; accepted on August 24, 2014

1 INTRODUCTION

Most phyla include species with the remarkable ability to regenerate lost body sections, the regeneration of limb appendages being one of the most complex and fascinating processes (Brookes and Kumar, 2008). Hundreds of experiments based on physical manipulations and molecular disturbances have identified many key genetic elements necessary for limb regeneration; however, no model exists yet that can explain the sufficient mechanisms that enable salamanders and other organisms to carry out the perfect regeneration of amputated limbs (Nacu and Tanaka, 2011). Understanding the process of regeneration of complex 3D shape from cell growth is a problem of fundamental importance in biology, and this goal should not get lost in the search for the minute details of genetic and molecular

interactions if we ever hope to find comprehensive testable models required to implement the sought-after promising biomedical applications (Lobo *et al.*, 2014b).

The lack of formalization tools to describe functional experiments and a centralized repository to store and analyze the vast experimental literature of limb regeneration are hindering our ability to find a comprehensive model of limb regeneration. Papers are written in natural language, which is ambiguous and difficult to systematically mine and analyze, whereas proposed text-based phenotypic ontologies (Robinson and Webber, 2014) are not ideal for describing qualitative (upper-arm versus forearm), quantitative (number of fingers), topological (ectopic hand in forearm) and geometrical (segment length) information key in regenerative experiments.

In contrast, novel functional mathematical ontologies based on graphs have been recently proposed for the formalization of morphologies and manipulations for planarian worm (Lobo *et al.*, 2013b) and limb (Lobo *et al.*, 2014a) regenerative experiments. In contrast to natural language, a mathematical graph can unambiguously specify morphological data (Lobo *et al.*, 2011). A graph can represent both complete jointless organisms such as planaria (where regions are represented by graph nodes connected by graph edges) and joint-based structures such as limbs (where segments connected by joints are represented by graph edges connected by graph nodes).

Here, we present a database based on a functional ontology system that unambiguously centralize the knowledge of limb regeneration experiments in a diverse set of model organisms, enabling the unambiguous description, centralize storage and manual and automated mining of this vast dataset.

2 METHODS

2.1 Functional graph-based ontology for limb morphologies and manipulations

To unambiguously describe the main characteristics of a limb morphology in the database, we used a novel functional ontology for limb regeneration experiments (Lobo *et al.*, 2014a). A mathematical graph represents the topological structure of a limb, where nodes specify limb segments and links specify the connection between them. Labels in the graph define the segment type and its geometrical characteristics, including connection angles and overall size and shape. Manipulations are defined in the ontology with a tree-like mathematical graph: external nodes (leaves of the tree) describe starting limb morphologies, and internal nodes describe elemental actions performed to the limb (remove,

*To whom correspondence should be addressed.

crop, join, relocate, reverse or irradiate). Links define the interrelation between actions.

2.2 Database implementation and curation

The Limbform database is implemented with the database engine SQLite (public domain). A single file stores both the database schema and the data, facilitating the administration and sharing of databases. The centralized database was manually curated using a freely available specific software tool (Lobo *et al.*, 2014a), which uses interactive graph representations and automatically generated cartoon diagrams as an optimal user-interface for regeneration experiments (Lobo *et al.*, 2013a).

3 RESULTS

We created the Limbform database of limb regeneration experiments based on a functional ontology of mathematical graphs. The database currently contains published regeneration experiments performed on the main model organisms in limb regeneration, including salamanders, frogs, insects, crustaceans and arachnids. An experiment in the database links a manipulation procedure (any combination of cuts, amputations, reversals, grafts and irradiations) to the resultant limb morphologies at specific regeneration times. In this way, the database can record morphologies with partial penetrance, as well as temporal events during the regeneration process. In addition, an experiment also includes information about the publication details (including a Web link to the original publication), the species used and the pharmacological and genetic treatments administered and their timing during the experiment.

Currently, the centralized database contains 829 experiments manually curated from 90 selected publications spanning from the 1950s to the present day. The selected experiments cover a wide range of surgical manipulations and treatments and a great variety of resultant morphologies.

Figure 1 shows screenshots of the database, accessed with the specific software application. The database is logically divided into experiments, manipulations, morphologies, treatments, species and publications (Fig. 1A). The details of an experiment include its name, publication, species, pharmacological and genetic treatments, the manipulation performed and the resultant morphologies obtained at specific times after amputation (Fig. 1B). Limb morphologies contain segments specific to the species used, their topological connectivity and their general shapes (Fig. 1C). Complex manipulations are made of a combination of nested elemental actions, such as crops and joins (Fig. 1D).

4 DISCUSSION AND CONCLUSIONS

Limbform is the first repository centralizing the main experiments of limb regeneration in multiple model organisms. We curated hundreds of limb regeneration experiments, resultant morphologies and manipulations from the published literature into a centralized database. We have followed a careful curation methodology to assure the consistency of the data; however, we encourage the publications' authors to verify the accuracy of the curated experiments and report through our dedicated Web site any possible inaccuracy in the data. We are continuously publishing revised versions of the database, with new and corrected data, and will incorporate user revisions as they are submitted.

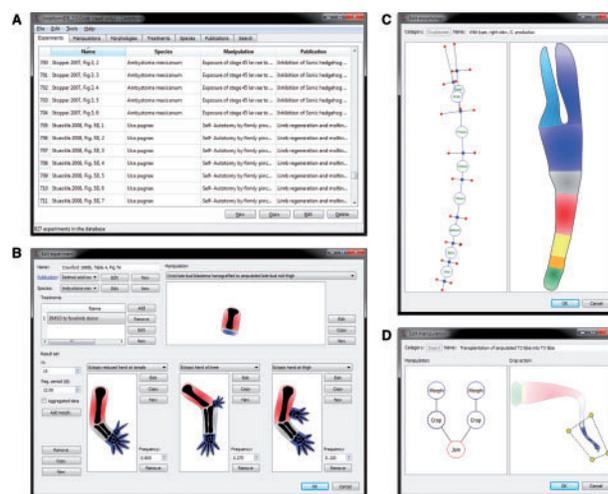


Fig. 1. Components of the database of limb experiments, visualized with the dedicated software tool. (A) Window listing all the experiments stored in the database. (B) A regeneration experiment describes the treatments, manipulations and resultant morphologies. (C) A limb morphology is defined with a graph representing the segments topology and their overall shape. (D) A surgical manipulation is defined with a hierarchy of elemental actions

The rapid accumulation of regeneration experiments and the difficulty of human scientists to glean algorithmic (constructive) models of pattern formation (not merely genetic pathways) necessitate the use of computational approaches. The Limbform database, owing to the mathematical and functional nature of its ontology, will pave the way for the application of computational automated tools based on artificial intelligence algorithms that will assist in the discovery of comprehensive models of regeneration.

The Limbform database of limb regeneration experiment represents an invaluable resource for the community and establishes the first necessary steps of a new bioinformatics of shape. This in turn will assist in the discovery and development of innovative biological models driving transformative applications in regenerative biomedicine.

ACKNOWLEDGEMENTS

The authors thank the Levin Lab members for valuable discussions and the four anonymous reviewers for useful suggestions.

Funding: National Science Foundation (EF-1124651), National Institutes of Health (GM078484), W. M. Keck Foundation and G. Harold and Leila Y. Mathers Charitable Foundation.

Conflict of interest: none declared.

REFERENCES

- Brockes, J.R. and Kumar, A. (2008) Comparative aspects of animal regeneration. *Annu. Rev. Cell. Dev. Biol.*, **24**, 525–549.
- Lobo, D. *et al.* (2014a) A bioinformatics expert system linking functional data to anatomical outcomes in limb regeneration. *Regeneration*, **1**, 37–56.
- Lobo, D. *et al.* (2014b) A linear-encoding model explains the variability of the target morphology in regeneration. *J. R. Soc. Interface*, **11**, 20130918.

Lobo,D. *et al.* (2011) Graph grammars with string-regulated rewriting. *Theor. Comp. Sci.*, **412**, 6101–6111.

Lobo,D. *et al.* (2013a) Planform: an application and database of graph-encoded planarian regenerative experiments. *Bioinformatics*, **29**, 1098–1100.

Lobo,D. *et al.* (2013b) Towards a bioinformatics of patterning: a computational approach to understanding regulative morphogenesis. *Biol. Open*, **2**, 156–169.

Nacu,E. and Tanaka,E.M. (2011) Limb regeneration: a new development? *Annu. Rev. Cell. Dev. Biol.*, **27**, 409–440.

Robinson,P.N. and Webber,C. (2014) Phenotype Ontologies and Cross-Species Analysis for Translational Research. *PLoS Genet.*, **10**, e1004268.