

Bioinformatik: In zwei Wochen Werkzeuge fürs Biologie-Leben sammeln

Röbbe Wünschiers: Wiley-Schnellkurs Bioinformatik für Anwender; Softcover, 256 Seiten, zahlreiche Abbildungen und Programmcodes, Wiley-VCH Verlag Weinheim, 2016, ISBN 978-3-527-53040-3, 16,99 Euro.

Aus Daten Wissen zu generieren: Darum geht es CLB-Autor Röbbe Wünschiers in diesem Buch – und weiter noch: Er zeigt, wie man das Wissen teilt, Visualisierungen für Vorträge erzeugt, es übers Internet verbreitet. Dies führt er an praktischen Beispielen vor. Die Werkzeuge, die er verwendet, sind typischerweise als freie Software verfügbar: Linux, AWK, MariaDB/MySQL, das Programmierpaket R oder auch LaTeX.

In seinem Vorwort macht Diethard Tautz, Evolutionsbiologe am MPI in Plön, die Notwendigkeit eines Buchs wie dieses deutlich: Einerseits sind die Daten in der Biologie schneller gewachsen als die Leistungsfähigkeit der Computer, man denke nur an Genom-Screening etwa bei der Epigenetik oder der Aufklärung des Metabioms. Andererseits ist die Aufnahmefähigkeit herkömmlicher Programme hierfür limitiert, laut Tautz etwa in Excel auf 100 000 Zeilen. Mit den im Buch vorgestellten Werkzeugen soll es jedoch möglich sein, „Big Data“ sogar auf einem Heimcomputer zu verarbeiten.

Das Buch setzt nur ein Basiswissen in Biologie sowie eine spielerisch-leichte Bereitschaft zum Umgang mit dem Computer voraus. Schritt für Schritt wird dann der Gebrauch aller notwendigen Werkzeuge vermittelt, von Linux über die Programmierung bis zu den Datenbanken oder der Visualisierung. Das Ganze geschieht anhand von fünf Beispielen: die Frage nach der Gefährlichkeit von EHEC, die Identifikation von Biogas-produzierenden Bakterien, die Darstellung von Expressionskarten, Untersuchung der Gefährlichkeit des H1N1-Schweinegrippevirus bzw. der Variation bei Ebola-Viren.

Das Buch ist ein Schnellkurs, d. h. ggf. sehr kompakt. Zu Anfang wird in wenigen Zeilen gefordert, Linux auf dem Rechner zu installieren, am besten als virtuelle Maschine... Aber trotzdem keine Angst: Gerade auch die ersten notwendigen Schritte werden beschrieben.

Als Belohnung winkt das Ziel, in der Lage zu sein, eigene Daten bearbeiten zu können sowie neue Methoden zur Bearbeitung zu entwickeln. Damit soll man dann in der Lage sein, die Datenflut in der Biologie zu beherrschen, sei es in Praktika, Abschlussarbeiten oder Forschungsprojekten.

Meiner persönlichen Einschätzung nach – einerseits mit dem Hintergrund eines Studiums, in dem man Rechner noch in Fortran über Lochkarten programmiert hat, andererseits derzeit mit dem Blick auf zwei studierende Kinder in den Fächern Biologie und Informatik – und natürlich der Übersicht eines Wissenschaftsjournalisten ;-) eignet sich das Buch optimal für Studierende der Biologie und verwandter Fächer nach dem ersten oder zweiten Semester. Dann hat man den „Schock“ der universitären Leistungsanforderung – im Vergleich zu der im Gymnasium – überstanden und sieht erste eigene Interessen und Aufgaben auf sich zukommen. Mein Rat: Sich in den Ferien zwei Wochen freinehmen, an den Computer setzen und mit Freude an den Erfolgserlebnissen das Buch durcharbeiten. Dann hat man sich Rüstzeug erarbeitet, das einem viel Zeit bei kommenden Aufgaben erspart.
Rolf Kickuth

