Title	Bionode - Modular and universal bioinformatics
Author	*Bruno Vieira*, Yannick Wurm
Affiliation	http://sbcs.qmul.ac.uk
Contact	mail@bmpvieira.com
URL	http://bionode.io
License	MIT

The exponential growth of biological data generated from sequencing in the last 10 years is putting a lot of pressure in the bioinformatics analysis downstream. However, most tools were developed in an unsustainable way, with a monolithic architecture, complex APIs, custom file formats and without good software development practices. This makes it very hard to build flexible and reproducible analysis pipelines that reuse those tools. In addition, the migration of most utilities to the web has led many developers to rewrite some functionality in JavaScript (a language that can be directly interpreted by browsers).

Bionode aims to solve bioinformatic problems while building highly reusable tools and code, by following the best development practices coming from web startups. Each Bionode module aims to be both a pipeable command line tool that follows the UNIX philosophy (*"try to do just one thing well"*) but also a JavaScript module that can be integrated in web applications or server-side JavaScript (Node.js). This while using standards like JSON file format and Google's Protocol Buffers.

This allows for an unprecedented amount of flexibility in the usability of Bionode. Unlike other similar bio* libraries that require the usage of a specific language, Bionode only requires the usage of JavaScript for web development use cases, and can thus be used via command line in other languages or workflow projects (e.g., Galaxy, NoFlo, and Node-RED). The architecture based on Node.js Streams allows to write complex pipelines that scale without running out of memory by piping chunks of data around.

The high level of modularity of the project also makes it easier for development and contributions, due to the split of functionality into small modules, individually hosted on GitHub/NPM with continuous integration, testing and code coverage. Everything is 100% Open Source (MIT license) and development discussions happen publicly. We also collaborate with Dat (*"git for data"*) and BioJS (*"represent biological data on the web"*).