Galaxy Training: A powerful framework for teaching!

Saskia Hiltemann, Helena Rasche, Simon Gladman, Hans-Rudolf Hotz, Delphine Larivière, Daniel Blankenberg, Pratik Jagtap, Thomas Wollmann, Anthony Bretaudeau, Nadia Goué, et al.

▶ **To cite this version:**

Saskia Hiltemann, Helena Rasche, Simon Gladman, Hans-Rudolf Hotz, Delphine Larivière, et al.. Galaxy Training: A powerful framework for teaching!. PLoS Computational Biology, 2022, 19 (1), pp.e1010752. 10.1371/journal.pcbi.1010752 . hal-04069041

HAL Id: hal-04069041
https://hal.inrae.fr/hal-04069041

Submitted on 14 Sep 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons CC0 - Public Domain Dedication 4.0 International License

PLOS COMPUTATIONAL BIOLOGY

EDUCATION

Galaxy Training: A powerful framework for teaching!

Saskia Hiltemann[1☯], Helena Rasche[1,2,3☯], Simon Gladman[4†], Hans-Rudolf Hotz[5,6], Delphine Larivière[7], Daniel Blankenberg[8], Pratik D. Jagtap[9], Thomas Wollmann[10], Anthony Bretaudeau[11,12], Nadia Goué[13], Timothy J. Griffin[9], Coline Royaux[14], Yvan Le Bras[14], Subina Mehta[9], Anna Syme[4], Frederik Coppens[15,16], Bert Droesbeke[15,16], Nicola Soranzo[17], Wendi Bacon[18], Fotis Psomopoulos[19], Cristóbal Gallardo-Alba[2], John Davis[20], Melanie Christine Föll[21,22,23], Matthias Fahrner[21,23], Maria A. Doyle[24,25], Beatriz Serrano-Solano[2,26,27], Anne Claire Fouilloux[28], Peter van Heusden[29], Wolfgang Maier[2], Dave Clements[30], Florian Heyl[2], Galaxy Training Network[¶], Björn Grüning[2‡], Bérénice Batut[2‡*]

1 Clinical Bioinformatics Group, Department of Pathology, Erasmus Medical Center, Rotterdam, the Netherlands, **2** Bioinformatics Group, Department of Computer Science, Albert-Ludwigs-University Freiburg, Freiburg, Germany, **3** Academie voor de Technologie van Gezondheid en Milieu, Avans Hogeschool, Breda, the Netherlands, **4** Melbourne Bioinformatics, University of Melbourne, Parkville, Victoria, Australia, **5** Friedrich Miescher Institute for Biomedical Research, Basel, Switzerland, **6** SIB Swiss Institute of Bioinformatics, Basel, Switzerland, **7** Eberly College of Science, Pennsylvania State University, State College, Pennsylvania, United States of America, **8** Genomic Medicine Institute, Lerner Research Institute, Cleveland Clinic, Cleveland, Ohio, United States of America, **9** Department of Biochemistry, Molecular Biology and Biophysics, University of Minnesota, Minneapolis, Minnesota, United States of America, **10** Biomedical Computer Vision Group, BioQuant, IPMB, Heidelberg University, Heidelberg, Germany, **11** IGEPP, INRAE, Institut Agro, Univ Rennes, Rennes, France, **12** GenOuest Core Facility, Univ Rennes, Inria, CNRS, IRISA, Rennes, France, **13** AuBi, Mesocenter, Clermont Auvergne University, Aubière, France, **14** PNDB French Biodiversity e-infrastructure, PatriNat, French Museum of Natural History, Concarneau, France, **15** Department of Plant Biotechnology and Bioinformatics, Ghent University, Ghent, Belgium, **16** VIB Center for Plant Systems Biology, Ghent, Belgium, **17** Earlham Institute, Norwich Research Park, Norwich, United Kingdom, **18** School of Life, Health and Chemical Sciences, The Open University, Milton Keynes, United Kingdom, **19** Institute of Applied Biosciences, Centre for Research and Technology Hellas, Thessaloniki, Greece, **20** Department of Biology, Johns Hopkins University, Baltimore, Maryland, United States of America, **21** Institute for Surgical Pathology, Medical Center—University of Freiburg, Faculty of Medicine, Freiburg, Germany, **22** Khoury College of Computer Sciences, Northeastern University, Boston, Massachusetts, United States of America, **23** German Cancer Consortium (DKTK) and Cancer Research Center (DKFZ), Freiburg, Germany, **24** Peter MacCallum Cancer Centre, Melbourne, Victoria, Australia, **25** Sir Peter MacCallum Department of Oncology, The University of Melbourne, Melbourne, Victoria, Australia, **26** ELIXIR Germany, Institute of Bio- and Geosciences (IBG-5)—Computational Metagenomics, Forschungszentrum Jülich GmbH, Jülich, Germany, **27** Euro-BioImaging ERIC Bio-Hub, European Molecular Biology Laboratory (EMBL), Heidelberg, Germany, **28** Simula Research Laboratory, Oslo, Norway, **29** South African Medical Research Council Bioinformatics Unit, South African National Bioinformatics Institute, University of the Western Cape, Bellville, South Africa, **30** Anaconda Inc, Austin, Texas, United States of America

☯ These authors contributed equally to this work.
† Deceased.
‡ BG and BB also contributed equally to this work.
¶ Membership list can be found in the Acknowledgments section.
* berenice.batut@gmail.com

Check for updates

OPEN ACCESS

Citation: Hiltemann S, Rasche H, Gladman S, Hotz H-R, Larivière D, Blankenberg D, et al. (2023) Galaxy Training: A powerful framework for teaching! PLoS Comput Biol 19(1): e1010752. https://doi.org/10.1371/journal.pcbi.1010752

Editor: Francis Ouellette, McGill University, CANADA

Published: January 9, 2023

Copyright: This is an open access article, free of all copyright, and may be freely reproduced, distributed, transmitted, modified, built upon, or otherwise used by anyone for any lawful purpose. The work is made available under the Creative Commons CC0 public domain dedication.

Data Availability Statement: All training materials mentioned in this paper are stored on GitHub (https://github.com/galaxyproject/training-material/), freely available online (https://training.galaxyproject.org/). The infrastructure code (templates, automation, scripts) behind the project is stored on GitHub (https://training.galaxyproject.org/). Most figures in this paper have been generated using Jupyter notebooks available on the GitHub repository for this paper (https://github.com/galaxyproject/GTN-community-paper-2020).

Funding: The authors would like to acknowledge funding from different funders: SH, HR, BB, FP, AB, and YLB from the Erasmus+ programme of the European Union (2020-1-NL01-KA203-064717),

Abstract

There is an ongoing explosion of scientific datasets being generated, brought on by recent technological advances in many areas of the natural sciences. As a result, the life sciences have become increasingly computational in nature, and bioinformatics has taken on a

including salary; SH and HR from the European Union's Horizon 2020 research and innovation programme under grant agreement No 825775, including salary; BB, BG, WM, TW from the German Federal Ministry of Education and Research BMBF grant 031 A538A de.NBI-RBC, including salary; BD, FC from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824087 (EOSC-Life) and the Research Foundation - Flanders (FWO) for ELIXIR Belgium (I002819N); FH from DFG (322977937/GRK2344); PVH from South African Research Chairs Initiatives of the Department of Science and Technology and National Research Foundation of South Africa grant UID 64751 and South African Medical Research Council flagship program MRC-RFA-UFSP-01-2013/COMBAT-TB; TJG, PDJ and SM from NIH NCI U24CA199347; JD from NIH NCI 5U24CA231877; JD, DC from NIH NHGRI 5U24HG010263, NIH NHGRI 2U24HG006620' DC from NIH NIAID 1R01AI134384 and National Science Foundation U. S. 1445604; HH from the Novartis Research Foundation; BB, BG, WM from European Union's Horizon Europe Research and Innovation Programme under agreement No 101046203 (BY-COVID); ACF from the European Union's Horizon 2020 programme under grant agreement No 857652 (EOSC-Nordic) and No 101017501 (RELIANCE); DB from NIH NHGRI U24HG006620 and NIH NHGRI U24CA231877; WB from the UKRI Medical Research Council (MR/S035931/1); PDJ from ASM-IUSSTF Visiting Teaching Fellowship; NS from the Biotechnology and Biological Sciences Research Council (BBSRC), part of UK Research and Innovation, Core Capability Grant BB/CCG1720/1 and the National Capability BBS/E/T/000PR9814. The funders had no role in study design, data collection, and analysis, decision to publish, or preparation of the manuscript.

Competing interests: We have read the journal's policy and the authors of this manuscript have the following competing interests: DB has a significant financial interest in GalaxyWorks, a company that may have a commercial interest in the results of this research and technology. This potential conflict of interest has been reviewed and is managed by the Cleveland Clinic. This does not alter our adherence to all the PLOS Computational Biology policies on sharing data and materials.

central role in research studies. However, basic computational skills, data analysis, and stewardship are still rarely taught in life science educational programs, resulting in a skills gap in many of the researchers tasked with analysing these big datasets. In order to address this skills gap and empower researchers to perform their own data analyses, the Galaxy Training Network (GTN) has previously developed the Galaxy Training Platform (https://training.galaxyproject.org), an open access, community-driven framework for the collection of FAIR (Findable, Accessible, Interoperable, Reusable) training materials for data analysis utilizing the user-friendly Galaxy framework as its primary data analysis platform. Since its inception, this training platform has thrived, with the number of tutorials and contributors growing rapidly, and the range of topics extending beyond life sciences to include topics such as climatology, cheminformatics, and machine learning. While initially aimed at supporting researchers directly, the GTN framework has proven to be an invaluable resource for educators as well. We have focused our efforts in recent years on adding increased support for this growing community of instructors. New features have been added to facilitate the use of the materials in a classroom setting, simplifying the contribution flow for new materials, and have added a set of train-the-trainer lessons. Here, we present the latest developments in the GTN project, aimed at facilitating the use of the Galaxy Training materials by educators, and its usage in different learning environments.

Introduction

Education is a fundamental human right (e.g., the Universal Declaration of Human Rights (UDHR), the International Covenant on Economic, Social and Cultural Rights (ICESCR)). Indicators of the achievement of education as a right are outlined in the ICESCR, and further developed into what is known as the "4 As framework" [1], which specifies Availability, Accessibility, Acceptability, and Adaptability as essential metrics. The 4 As, therefore, provide a concrete set of ideals to strive towards in any global educational endeavor.

Over time, education has increasingly come to rely on technology and online resources, a trend accelerated greatly by the Coronavirus Disease 2019 (COVID-19) pandemic, when much of the world's education was forced to occur online. This comes with a novel set of challenges, such as feelings of isolation and lack of engagement [2–5], but also some notable advantages, such as increased flexibility [6] and self-paced and explorative learning [7,8]. As noted by Gallardo and colleagues [9], a constructivist approach to education is most suited to this mode of learning. In this approach, an emphasis is placed on active experimentation by learners in real-world contexts. This encourages the selection, organization, and integration of learner's experiences with previous knowledge in a social context [10]. This approach also emphasizes the importance of establishing learning communities for successfully implementing online learning practices [11].

Galaxy [12] is an open source platform for accessible, reproducible, and transparent computational research, driven by an inclusive and diverse worldwide community. Researchers are able to access a wealth of tools (>8,500 tools in the Galaxy ToolShed [13] as of March 2022), datasets, and high-performance compute resources, through a standard web browser, without requiring informatics expertise. However, comprehensive training is still required to adequately understand the data analyses and to accurately interpret the results. A survey of 704 NSF-BIO-funded investigators revealed that training topics were the top 3 of 13 most unmet

data analysis needs, above HPC/cloud facilities, workflows/pipelines analysis or storage facilities needs [14].

To address this large demand for training, a community of practice, the Galaxy Training Network (GTN; https://training.galaxyproject.org) was founded to provide both learners and instructors with free online training resources, connect them with a global community, and help promote open data analysis practices worldwide [15]. The aims of the Galaxy project as a whole, as well as those of the GTN, are well aligned with the 4 As. Furthermore, the GTN training resources are compatible with a constructivist approach to education and can be used with both synchronous and asynchronous learning, be it online or in a classroom setting.

GTN training materials are openly developed, maintained, and community-reviewed via GitHub (https://github.com/galaxyproject/training-material). The materials cover an increasing spread of topical domains, such as life sciences, computational chemistry, climate sciences, data visualization, statistics, and machine learning.

By utilizing Galaxy as its primary data analysis platform, the GTN also addresses the clear demand for easily accessible compute resources for research and training in science [14]. Galaxy is an ideal platform for effective learning and teaching. It takes care of all the heavy lifting behind the scenes, such as installing software and managing the compute infrastructure, leaving learners free to focus on the scientific analysis concepts at hand, rather than on the details of the technical implementations. It encourages active participation of learners [9].

Given this observation of the value of Galaxy and the GTN in teaching activities, we have aimed to optimize the GTN framework not only for learners, but also to support educators in the development and reuse of training materials, e.g., by integrating the FAIR (Findable, Accessible, Interoperable, Reusable) principles for training materials [16] directly into the framework. In this paper, we describe the aims and objectives of the GTN and its framework, and highlight some recent efforts by the GTN community to expand and improve the project. We also present how this project supports both educators and lesson developers to bring availability, accessibility, acceptability, and adaptability to bioinformatics education for trainers and learners alike. Finally, we showcase some example user stories of how the GTN materials are being used across all levels of education.

Much life science training is driven and hosted by academic institutions, such as universities (e.g., the UCSC Genome Browser), institutes (e.g., NIH-hosted NCBI), and funded consortia (e.g., ELIXIR, The Australian BioCommons), who typically offer synchronous online training and in-person workshops. By virtue of its geographic breadth in users and training contributors, the Galaxy training ecosystem must accept challenges posed by this globally (and temporally) distributed network. These challenges have been met with particular success in automation and peer-review. This approach has provided seamless integration with similar initiatives such as Software Carpentry [17] to produce the highly successful Gallantries project [18], which in turn took inspiration from the Australians BioCommons approach to hybrid teaching [19].

The Galaxy Training Network is a collaborative effort to develop and maintain state-of-the-art Galaxy-based training in the life sciences and beyond. This community of practice [20] helps to develop and implement Open Science practices and FAIR principles in training, with an open infrastructure that other education initiatives can reuse.

In this paper, we provide an overview of the GTN training framework, our ongoing efforts to keep our materials FAIR and sustainable, and we present the latest development within the GTN, with a focus on novel features supporting teachers and instructors in the reuse of our materials. Finally, we provide a set of user stories showcasing the broad range of educational settings the GTN training materials and framework are used for by the community.

Results

GTN background

The Galaxy Training Network and its framework have grown significantly over the past years (Fig 1), and the GTN website has seen a steady increase in usage since its creation (Fig 1D), with in average 28,000+ visits per month. Since September 2018, 1,800+ individual tutorial feedback forms have been submitted with a high satisfaction rate (more than 88% had a satisfaction rate of 4 or 5, per Fig 2D).

By using one of the worldwide public Galaxy servers, learners and educators alike gain free access to high-performance compute resources without any of the systems administration burden. These include the UseGalaxy.* servers with Galaxy Main (https://usegalaxy.org), Galaxy Europe (https://usegalaxy.eu), and Galaxy Australia (https://usegalaxy.org.au), but also 160+ others [21]. Therefore, over the past years, Galaxy has become increasingly attractive as a training platform for instructors around the world [22] (Fig 1A). For example, between June 2018 and May 2022, UseGalaxy.* servers have been used for 330+ registered training events reaching 17,000+ learners. In addition to these short-running training events, GTN tutorials have also been included as part of formal undergraduate and graduate courses (see User stories section for more details).

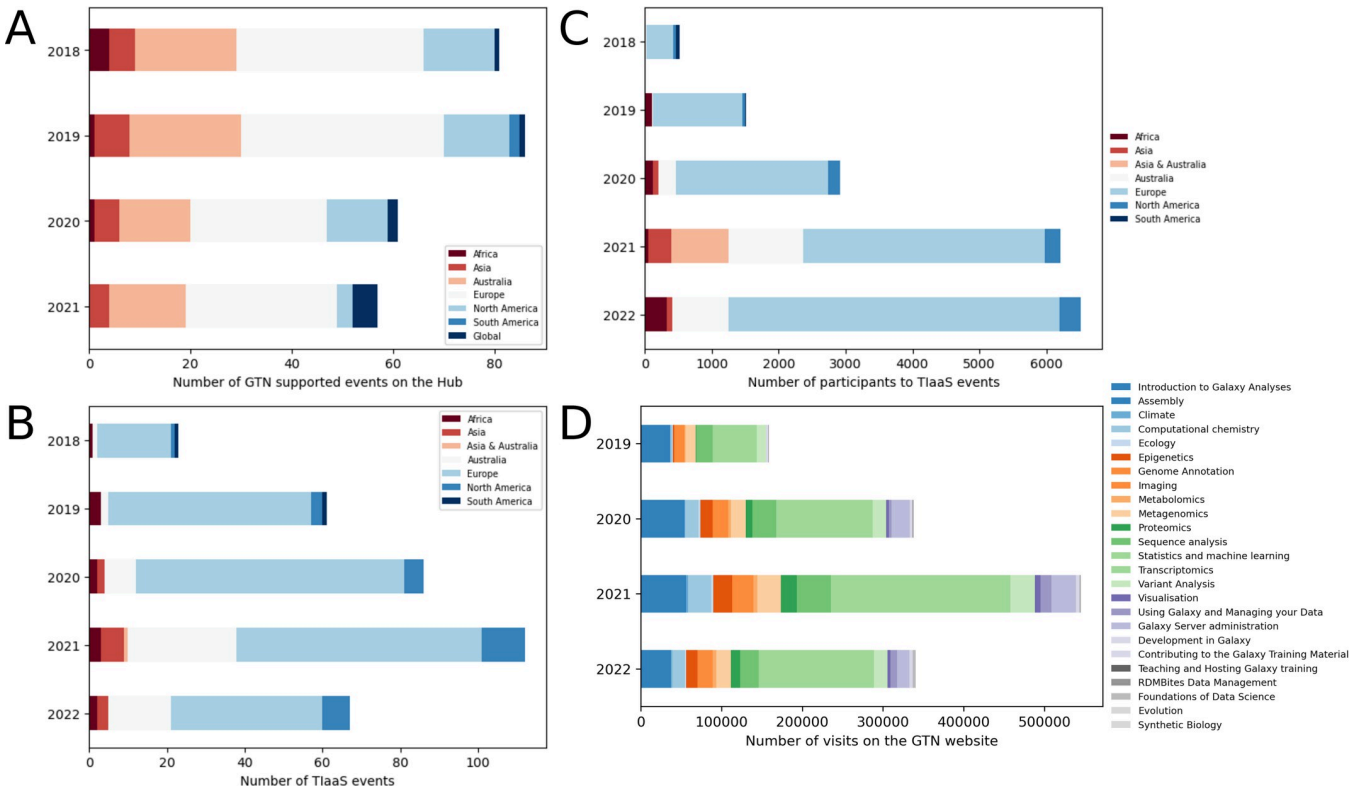


Fig 1. Statistics about training events organized by the Galaxy community and visits of the GTN website over years. (A) Number of events supported GTN materials, registered on the Galaxy Community Hub website [22], per year between 2018 and 2021. (B) Number of TIaaS events per year on Galaxy Europe (collected in May 2022). (C) Number of participants to TIaaS events per year on UseGalaxy.* (collected in May 2022). (D) Number of visits per year on the GTN website and per topics, initially tracked by Google Analytics and later with Plausible (collected in May 2022). The latest usage statistics are publicly available from https://plausible.galaxyproject.eu/training.galaxyproject.org.

https://doi.org/10.1371/journal.pcbi.1010752.g001

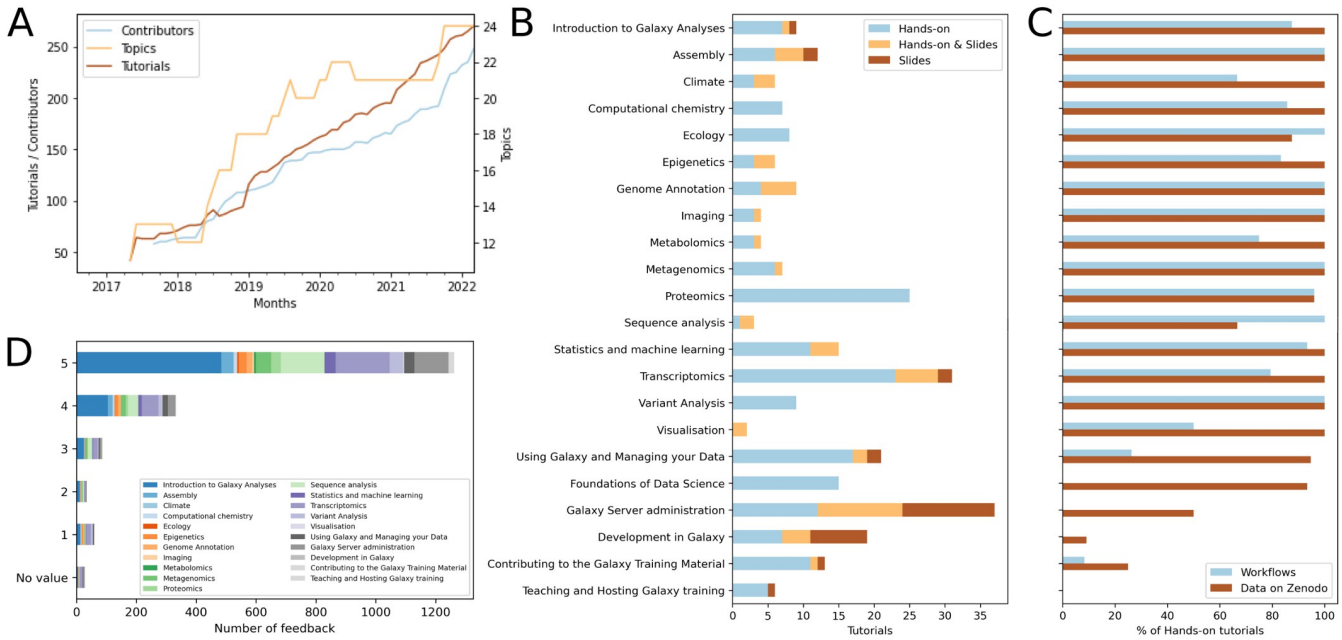


Fig 2. Content of material available on the GTN and feedback from learners. (A) Evolution of number of topics, tutorials, and contributors over the months between 2017 and 2022. (B) Number and type of tutorials per topics available on the GTN on April 2022. The latest statistics are publicly available from https://training.galaxyproject.org/stats. (C) Type of supporting materials for tutorials per topics available on the GTN on April 2022. (D) Score of the embedded feedback in the tutorials per topics. Three questions are asked in the form: "How much did you like this tutorial?" (from 1 (bad) to 5 (great)), "What did you like?", "What could be improved?". The latest feedback results are publicly available from https://training.galaxyproject.org/feedback.

https://doi.org/10.1371/journal.pcbi.1010752.g002

GTN overview: Materials and features

The GTN Materials are continually growing, both in terms of the number of tutorials as well as its community (Fig 2A). The GTN Materials in May 2022 has 260+ tutorials covering 23 topics (17 scientific and 6 technical), developed by over 260+ contributors (Fig 2A).

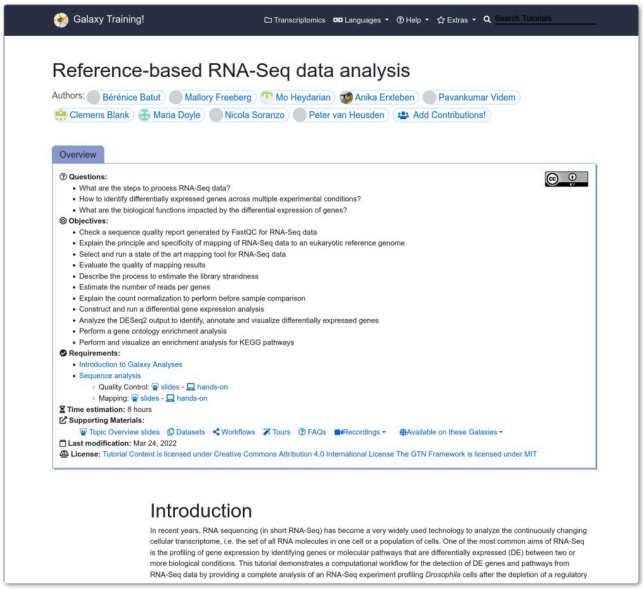
Tutorials. GTN Tutorials aim to follow best practices in course design [23] and are typically constructed around a real-world *research story*, e.g., a published journal article describing an analysis workflow or dataset. The tutorial starts by introducing the relevant scientific background, and then proceeds to describe the analysis, providing details on the relevant scientific and computational concepts involved at each step. The tutorials are composed of alternating theory and practical sections, interspersed with formative assessment questions and exercises to enhance learning [24], and may be supported by a set of introductory slides (Fig 3). Our tutorials are self-contained; everything needed to complete them is bundled with the tutorial; this includes all the datasets, lecture slides, videos, the workflows, and the list of public Galaxy servers supporting the tutorial. No software installation is needed to follow these tutorials; the only technical requirement is access to internet.

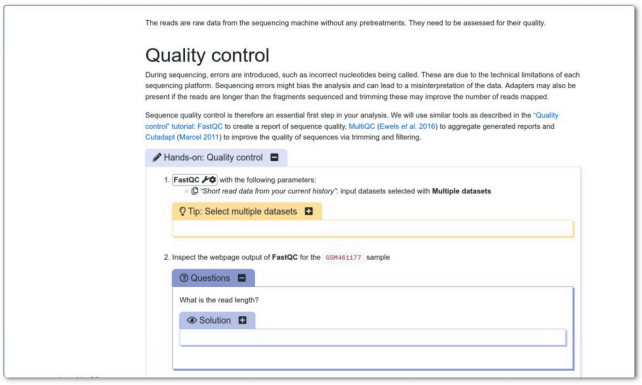
A survey of trainers identified a desire of increased interactivity between tutorials and Galaxy. We have developed a new Galaxy feature "Tutorial Mode" that allows following the training materials directly in Galaxy. With that mode activated, hands on steps are now interactive and clicking tools brings you directly to the tool in Galaxy.

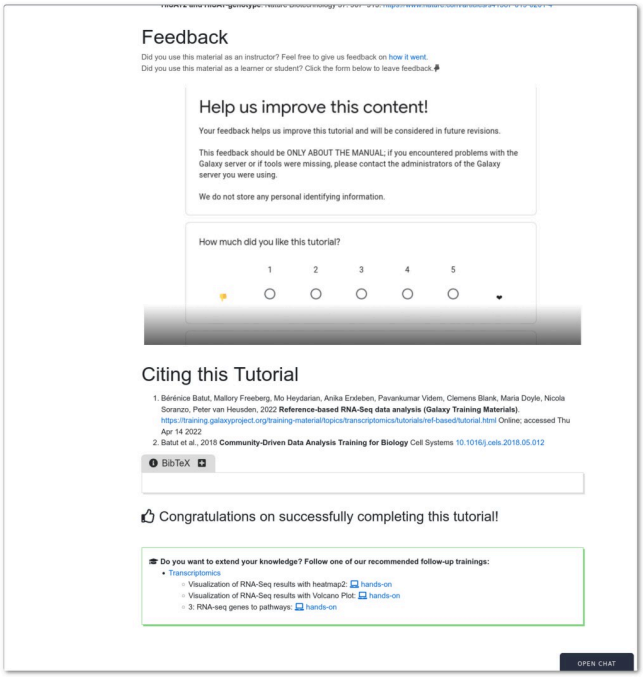
Recording and automated video lectures. In order to further expand the formats in which our training materials are available, several tutorials have been recorded by community members, are available on GTN Training Video Library [25] and accessible directly from the tutorial itself. Recording a video is a laborious and time-intensive activity. To save a significant

A

Top







Bottom

B

Rich metadata

- Title
- License
- Contributors
- Learning objectives
- Requirements
- Duration for teaching
- Audience
- Date of last revision
- Suggested further training

Structured content

- Introduction to the relevant scientific background
- Theory and analysis explanation with details on the relevant scientific and computational concepts involved
- Hands-on to apply the concepts
- Formative assesment question

Strong (technical) support

- List of possible Galaxy servers with needed tools
- Data on Zenodo and available as data libraries on usegalaxy.*
- Workflow
- Recordings
- Frequently Asked Questions

Community oriented

- Translation (automatic and currated) of content
- Highlighted contributors and funding and encourage contribution
- Giving credit information
- Embeded freedback form with automatic submission to GitHub and aggregation in a page
- Community support chat

Fig 3. Structure of a GTN tutorial and its features. (A) Screenshots for parts of a GTN tutorial to show its structure. Title and authors are listed, followed by an overview box containing metadata about the tutorial (target audience, learning objectives, prerequisites, supporting materials, time estimate, etc). The tutorial content itself is a mix of theory (text) and so-called "hands-on" boxes describing practical steps to be performed in Galaxy. Question and answer boxes may be added at any point in the tutorial to allow students to self-evaluate as they progress. The end of a tutorial provides a box with take-home messages, references, and suggestions for further reading and follow-up tutorials. Additionally, a feedback form is embedded at the end of every tutorial. Access to FAQs, support channels, page view statistics, and translated versions of the materials are provided via the top menu of the webpage. (B) Features of a GTN tutorial grouped in 4 categories: rich metadata, structured content, strong (technical) support, and community oriented.

https://doi.org/10.1371/journal.pcbi.1010752.g003

amount of instructor time, we automatically generate videos based on lecture slide decks. The slides are narrated using automated text-to-speech (TTS), and the script is based on the speaker's slide notes, which makes for an extremely easy video update process—updating the slides automatically re-records the videos.

Coding tutorials. Jupyter and RStudio [26] can both be run within Galaxy in the form of an *interactive tool*, allowing for interactive tutorials that offer a combination of Galaxy-based analysis and coding-based R or Python analysis steps. In addition to covering a wide range of analyses in Galaxy, the GTN also supports coding-oriented tutorials in the form of Jupyter [27] and RMarkdown [28] notebooks that are automatically generated from the tutorial content. A large set of introductory coding tutorials have recently been developed in a new "Foundations of data science" topic.

FAIR training. The GTN infrastructure has been developed in accordance with the FAIR principles for training materials [16] (Table 1) and following the 10 simple rules for collaborative lesson development as defined by Deveny and colleagues [29] (Table 2). Following these principles enables trainers and trainees to find, reuse, adapt, and improve the available tutorials.

GTN for educators

The GTN training platform as described before helps minimize the amount of time and effort required for instructors to prepare for and run their training courses and workshops, by providing tutorials and a complete training infrastructure.

Given the observation that different instructors have different approaches to teaching, GTN materials are designed to be flexible in their use. That is, instructors may embrace the constructivist approach to learning that our materials are designed to be compatible with, but the materials can also be used in a more instructivist approach if desired. For example, tutorials could either be performed by an instructor as a demonstration, or be left for learners to work through at their own pace (individually or in small groups), or a mix of these 2 styles, where instructors perform the steps of the tutorial centrally while the learners follow along, discussing the steps in detail as a group along the way. Moreover, most GTN tutorials also consist of both theoretical lecture slides and practical hands-on tutorials, but these are independent materials designed to complement each other, and each may be used in isolation as well (e.g., if one prefers to focus only on the theoretical background). Furthermore, we provide a large number of exercises and formative assessment questions in our tutorials, but these are independent from the main flow of the tutorial and are presented in collapsed sections that can be easily skipped by instructors when desired. To further increase this flexibility of GTN materials, we have recently introduced support for "choose your own adventure" style tutorials, where choices can be made by the learner, which affect the contents of the tutorial, for example, to change the level of technical detail provided, or which analysis tools, steps or datasets are used, allowing the tutorials to be easily adapted to a variety of audiences.

Train-the-trainer (TtT) tutorials. To support teachers and trainers, we have developed a dedicated topic in the GTN for TtT tutorials, covering all aspects of using the GTN materials

Table 1. Implementation of the "Ten simple rules for making training materials FAIR" [16] in the GTN.

	Rule	Implementation in the GTN
1	Plan to share your training materials online	Online training material portfolio (https://training.galaxyproject.org/), managed via a public GitHub repository (https://github.com/galaxyproject/training-material).
2	Improve findability of your training materials by properly describing them	Rich metadata associated with each tutorial that are visible and accessible via schema.org on each tutorial webpage.
3	Give your training materials a unique identity	URL persistency with redirection in case of renaming of tutorials. Data used for tutorials stored on Zenodo and associated with a Digital Object Identifiers (DOI).
4	Register your training materials online	Tutorials automatically registered on TeSS, the ELIXIR's Training e-Support System [30].
5	If appropriate, define access rules for your training materials	Online and free to use without registration.
6	Use an interoperable format for your training materials	Content of the tutorials and slides written in Markdown. Metadata associated with tutorials stored in YAML and workflows in JSON.
7	Make your training materials (re-)usable for trainers	Online. Rich metadata associated with each tutorial: title, contributor details, license, description, learning outcomes, audience, requirements, tags/keywords, duration, date of last revision. Strong technical support for each tutorial: workflow, data on Zenodo and also available as data libraries on UseGalaxy.*, tools installable via the Galaxy Tool Shed, list of possible Galaxy instances with the needed tools. (Fig 3). Extensive GTN train-the-trainer materials.
8	Make your training materials (re-)usable for trainees	Online and easy to follow tutorials. Rich metadata with "Specific, Measurable, Attainable, Realistic and Time bound" (SMART) learning outcomes following Bloom's taxonomy. Requirements and follow-up tutorials to build learning path. List of Galaxy instances offering needed tools, data on Zenodo and also available as data libraries on UseGalaxy.*. Support chat embedded in tutorial pages, video walkthroughs for many tutorials (Fig 3).
9	Make your training materials contribution friendly and citable	Open and collaborative infrastructure with contribution guidelines, a CONTRIBUTING file and a chat. Topic maintainers. How to cite tutorials and give credit to contributors available at the end of each tutorial (Fig 3).
10	Keep your training materials up-to-date	Open, collaborative, and transparent peer-review and curation process. Feedback collected from learners and instructors. Short time between updates.

https://doi.org/10.1371/journal.pcbi.1010752.t001

in education. This includes best practices providing pedagogical, technical and logistical recommendations, accessible on the website with a series of tutorials (available in "Teaching and Hosting Galaxy training" category). For first-time instructors, the GTN has come together to collect different instructors' experiences ("Training Philosophies" [31]).

Preparing a workshop. During preparation of a workshop, organizers and instructors work together to identify relevant tutorials for their event. Our tutorials are roughly divided into "topics" such as *Transcriptomics* or *Climate Science*, with a search function available on the GTN website (https://training.galaxyproject.org). Within topics, multiple tutorials can be found. Each tutorial starts with a list of metadata such as learning objectives following Bloom's taxonomy [32], prerequisites, time estimate, or questions addressed by the tutorial (Fig 3). This but also tags of tutorials enables instructors to identify the best tutorials for their audience. Reusable slide decks with speaker notes are also available to introduce the topics prior to a tutorial.

To ensure the success of practical sessions, the tutorials rely on specific datasets (included in the tutorial metadata) and tools that the GTN analyzes in order to provide a list of

Table 2. Implementation of the "Ten simple rules for collaborative lesson development" [29] in the GTN.

	Rule	Implementation in the GTN
1	Clarify audience	Tutorial metadata includes level indicators (introductory, intermediate, advanced) and a list of prerequisite tutorials as recommended prior knowledge. This information is rendered at the top of each tutorial.
2	Make lessons modular	Development of small tutorials linked together via learning paths. Recommended prior and follow-up tutorials as part of the tutorial metadata.
3	Teach best practice lesson development	A topic *Contributing to the Galaxy Training Material* including 10 tutorials describing how to create new content. Furthermore, quarterly online collaboration fest (CoFests) are organized, where contributors can get direct support. Development of a Train the Trainer program and a mentoring program for instructors, in which lesson development is taught.
4	Encourage and empower contributors	Involve them in reviews. Mentor them. Encourage them to become maintainers.
5	Build community around lessons	Quarterly online collaboration fest (CoFests) and Community calls. Chat (Matrix channel).
6	Publish periodically and recognize contributions	Author listed on tutorials. Hall of fame listing all contributors. Full tutorial citation at the end of the tutorial. Tweet about new or updated tutorials. List of new or updated tutorials in Galaxy Community newsletter. Soon: publication of tutorials via article.
7	Evaluate lessons at several scales	Tutorial change (Pull Request) review. Embedded feedback form in tutorials for trainee feedback. Instructor feedback. Automatic workflow testing.
8	Reduce, reuse, recycle	Sharing content between tutorials, specially using snippets. Development of small modular tutorials linked by learning paths.
9	Link to other resources	Links to original paper and data, documentation, external tutorials, and other material.
10	You can't please everyone	But we can try (several different Galaxy introduction tutorials for different audience). Aim to clearly state what the tutorial does and does not cover, at the start.

https://doi.org/10.1371/journal.pcbi.1010752.t002

compatible public Galaxy servers. Automated workflow testing [33] provides reassurances to instructors, letting them know that a given tutorial keeps working on their selected server.

We have added support for FAQs to be added to tutorials, in order to help instructors prepare their lesson. These FAQs typically cover common questions or frequently observed trainee mistakes encountered during a particular tutorial.

Preparing workshop infrastructure. A major challenge in computational workshop organization is the identification of affordable and reliable compute infrastructure. The needs of training infrastructure are significantly different from regular research; analysis steps must complete in relatively short time periods in order to not disrupt the flow of the lesson. The GTN uses small (sub-sampled) training datasets for tutorials in order to reduce the run time of the analysis step, but a second factor to consider is the waiting time of the compute infrastructure. In order to support training events, we developed Training Infrastructure as a Service (TIaaS) [34]. This free service provides a dedicated job queue for the participants of training events, in order to reduce waiting times on the cluster and ensure courses run smoothly and efficiently. Furthermore, the TIaaS service also provides instructors with a dashboard, enabling them to monitor the progress of the participants. To date, TIaaS has been widely used: 17,000 + students were taught over the course of 330+ events between June 2018 and May 2022, with 65% using the GTN materials (Fig 1B and 1C).

Teaching a workshop. During the workshop, instructors can introduce the topic using the available slide decks, supported by detailed speaker notes. After this theoretical introduction, instructors may either use a *live-demo* approach, guiding students through the tutorials

in a step-by-step fashion, or alternatively let students work through the tutorials at their own pace while providing support. The tutorials are a fully self-contained teaching resource as described in Fig 3, and therefore, support both these training modalities.

Feedback from instructors. In January 2020, we conducted a survey asking how trainers used the available training material and infrastructure. We received answers from 33 trainers, 88% of which had conducted a training event in the last 3 years. This sample consisted of a continuum of trainers, from occasional to seasoned trainers. Of those recently giving training events, 79% of them used GTN resources, and some even developed new GTN materials for the occasion.

The GTN tutorials are well liked and shared; 91% of respondents indicating they would recommend them. The use of GTN material allows them to have more time to focus on the fundamental principles of analyses and the specific needs of the trainees. Of the surveyed trainers, 69% are contributors to the GTN, and among those who are not, 55% declare that they plan to become one in the future.

Contributing to the GTN

The GTN framework also aims to ease the burden of tutorial development, contribution, and maintenance by offering a comprehensive set of tools and standards for tutorial authors.

Tutorials about contributing. To get acquainted with the tutorial development process, a contributor can start by consulting a set of dedicated tutorials in the *Contributing* topic in the GTN. These tutorials provide a step-by-step guide to creating a tutorial in the GTN framework, covering everything from technical guides about the framework to pedagogical best practices and submission of tutorials to GitHub.

Fig 4 depicts the typical process a tutorial contributor will go through when developing a new tutorial. The first step is usually to develop the analysis workflow in Galaxy. The challenge

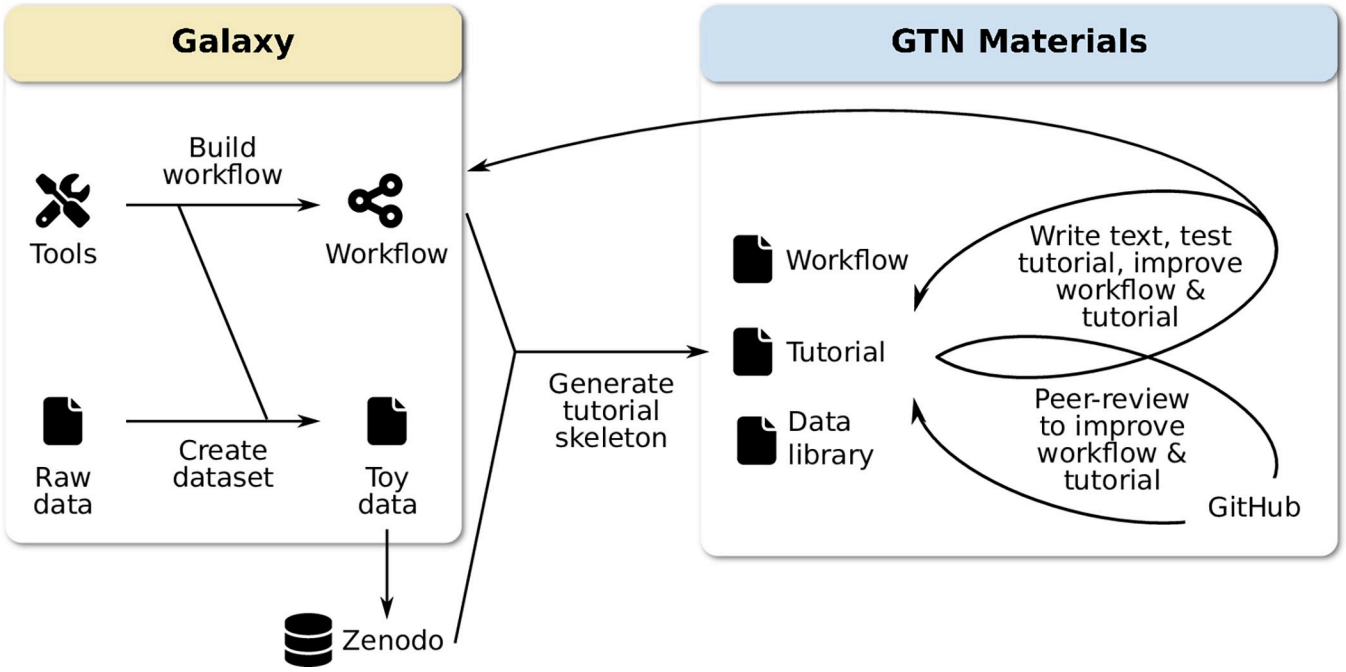


Fig 4. Typical process to create a new tutorial. Authors usually start by identifying suitable input datasets and developing the workflow in Galaxy. Thus, a workflow can be automatically converted into a tutorial skeleton using the PTDK described in the next section. The tutorial is then tested, reviewed, and finally merged into the GTN.

https://doi.org/10.1371/journal.pcbi.1010752.g004

is to identify or create suitable input datasets. The selected data must be informative enough to illustrate the meaning of a given analysis, but not too large as to require long waiting times for upload or processing during a training event. The selected data could be a toy dataset created from scratch, or (preferably) an informative subset of a real-life dataset, for example using a single chromosome of a whole-genome dataset.

Planemo Tutorial Development Kit (PTDK). To simplify the tutorial development process, we have created a *Tutorial Development Kit* within Planemo, the Galaxy development kit [35]. This automatically generates a tutorial skeleton from a Galaxy workflow as starting point for tutorial authors. This skeleton contains the overall structure of the tutorial: metadata section, example question boxes for formative assessments, instructions on how to proceed with adding scientific and pedagogical content, and more important auto-generated hands-on boxes for every analysis tool in the workflow with the parameters to select. This process greatly reduces the development time and allows tutorial creators to focus on the scientific content of the tutorial, rather than the technical details and style guidelines.

In order to further lower the contribution barrier, this *Planemo Tutorial Development Kit* has been encapsulated into a web service (https://ptdk.apps.galaxyproject.eu/), where tutorial authors can provide a link to a public workflow on one of the UseGalaxy.* servers and obtain the tutorial skeleton based on this workflow in Markdown, but also the workflow and data library file if a Zenodo link is provided. This removes the need for contributors to install and run Planemo locally. We have found that this approach saves time for not only the contributors but also for reviewers, since the tutorials contributed using this method adhere well to the GTN style guide.

Content reuse. To avoid duplication of content between tutorials, we have developed a set of modular tutorial components, called *snippets*, which can be easily reused across tutorials. For example, common tasks in Galaxy such as starting a workflow or creating a new history will be a part of most tutorials. Contributors can include these snippets at any place in a tutorial with a simple import statement. If instructions for this common task change, e.g., due to changes in the Galaxy user interface, the changes need to only be made in the snippet itself and will automatically propagate to any tutorial using them.

Tutorial preview and testing. GTN tutorials are written in Markdown format for ease of contribution and subsequently converted to HTML web pages automatically by our framework. The HTML web page can be locally previewed if desired by using simple commands. Additionally, contributors who do not wish to install and run the GTN locally, can also generate a preview of their in-development tutorials online using GitPod [36]. We have integrated the GTN GitHub repository with GitPod, enabling contributors to obtain an online tutorial development environment, complete with online preview, with a single click of a button. GitPod has proven particularly useful for collaboration between multiple lesson developers with varying degrees of familiarity with git or GitHub.

In addition to a visual inspection of the generated web pages, our framework offers a suite of testing tools, allowing contributors to check that their tutorial meets the technical and style guidelines. These tests include checks of whether all required metadata is present, whether links within the tutorial are valid, and whether files are correctly formatted. This helps contributors and reviewers to quickly identify and correct potential problems with a tutorial.

Peer review. Once a contributor is happy with their tutorial, they can create a *pull request* (PR) to the GTN GitHub repository (https://github.com/galaxyproject/training-material), where each contribution will then undergo automatic quality assurance tests and a peer-review process.

This review process is completely open, and any volunteers from the community may participate. For each topic within the GTN, we have encouraged several prominent community

members who are regular contributors to act as topic *maintainers* that help safeguard the quality of the content in the topic and are empowered to review, approve, and merge contributions. Maintainers and other reviewers will check the proposed contribution, both in terms of formatting and scientific content. They can make suggestions for updates and start a discussion with the contributor(s). Typically, there will be 2 or 3 reviewers for any given PR. Based on these reviews, the contributor(s) will update the tutorial, and this cycle of update and review will continue until both contributor and reviewer are happy with the result. In case of disagreements, topic maintainers will decide on the path forward.

This open strategy for content creation and updates is paying off. Since the beginning of the project in the middle of 2016, over 2,500 PRs have been created to add or update tutorials.

The review process is quite fast: it takes 20 days between the opening of a pull request and its merge. Thanks to a fast turnaround, tutorials are regularly updated and iteratively improved over time, keeping the material relevant and of a high quality. Indeed, each tutorial undergoes (on average) a PR every quarter to update it. Some highly used tutorials, like the "Reference-based RNA-seq" tutorial, have received 100+ PRs over the last 3 years, on average one every 3 weeks.

Maintenance. Tutorials are dynamic entities; underlying analysis tools receive updates, new tools are developed, and the Galaxy interface itself changes regularly. As a result, tutorials must undergo regular updates as well to reflect the state-of-the-art in the scientific domain and correctly reflect the latest available version of the tools and the Galaxy interface. Instructors preparing for their next workshop check the tutorials as preparation, and in the process, identify the places where the tutorial should be updated. If the instructors feel comfortable making these changes themselves, they can open a PR proposing the changes. If not, they can create an *issue* on GitHub to request the changes.

Feedback. Learner feedback is one of the most valuable resources to inform training improvements [37]. An anonymous feedback form is embedded at the end of every tutorial, where users can provide feedback and make suggestions for enhancement. All feedback is transparently collected on the GTN website, enabling the community to view and address the feedback (Fig 2D).

User stories: GTN across all levels of education

GTN in higher education. Galaxy and the GTN are an integral part of Bioinformatics programs at the undergraduate level (e.g., Clermont Auvergne University, Texas A&M University, Avans Hogeschool, University of Freiburg, University of Frankfurt).

In these undergraduate courses, Galaxy is often used to teach the concept of data analysis, pipelines/workflows, and also reproducibility. Students learn the importance of the tool versions, as tools can change subtly or significantly over time and this may impact their analyses. Teachers can likewise showcase the evolution of algorithms over time using different versions of tools (e.g., bowtie and bowtie2) to help students understand the advances in the field. Without the need to spend resources on teaching command line skills and tool installation, the time can be fully used to introduce each step (data cleaning, data analysis, and visualization) and to go into details of each tools and algorithms. Galaxy is also convenient to explain how data is structured (data types, metadata) and its connection to tools. Finally, Galaxy can be used to illustrate complete analytical workflows paying attention to input and output data.

GTN for research scientists. Postgraduate courses (e.g., Agrocampus Ouest, Rennes University, Brest University, Clermont Auvergne University, Station Biologique de Roscoff, Melbourne University, University of Freiburg) as well as internal bioinformatics short-format training sessions (e.g., Friedrich Miescher Institute, French Bioinformatics Institute, Erasmus

MC, University of Freiburg) are relying on Galaxy as the main teaching tool. These courses are often directed at learners who are not yet comfortable working with the command line. However, we often notice that after getting familiar with each analysis step and understanding the meaning of parameters of each tool, some may feel confident to learn command line or scripting (possibly within Galaxy using RStudio or Jupyter notebook) and to move to command line environments.

GTN materials are also often used to provide supplemental training to later-career research scientists, for example, at the Erasmus Medical Center (EMC), Earlham Institute, and as part of the Gallantries project [18]. These trainings often consist of short-running workshops, usually aimed at familiarizing researchers with novel analysis techniques. Since many of the GTN tutorials are centered around an analysis pipeline described in a recently published article, they are ideally suited to provide researchers with an update on the latest state-of-the art analysis pipelines in their domain. GTN tutorials are also frequently created as part of scientific publications by authors presenting a novel analysis method, as an additional form of documentation for the readers [38–43].

Training in underrepresented communities. Reliable internet access is often taken for granted by researchers. Bottlenecks in infrastructure may, however, cause significant issues for bioinformatics training events, as many students try to upload or download data at once. This is especially problematic in low- or middle-income countries (LMICs) where internet access may be intermittent, restricted, or completely unavailable.

Trainers are often brought in from afar, meaning that the teaching takes places in—for the trainer—an unfamiliar setting and on a strict time schedule limited by the return trip(s) of those involved. It is therefore important to avoid or minimize unforeseen delays caused by incompatibilities with local infrastructure, connectivity failures or unforeseen updates forcing new software to be downloaded or queries to remote servers to fail.

Based on these needs, the eBioKit [44] was developed: it is an assembly of open source or free-for-academic-use software, along with key databases and selected material for bioinformatics. It can be installed beforehand, brought on a portable server to the local training, and made available on the local network. Students can then access it through the network and work directly on the server, which avoids installation issues for the students, unforeseen updates of web services, or other failures. Galaxy has been a part of the eBioKit for most of its existence, specially in the eB3Kit that includes a Galaxy-based bioinformatics platform with a specialized workflow interface [45].

Various versions of the eBioKit has been used for over a decade by organizations such as EMBnet, H3Abionet, SANBio, and BECA/ILRI to train hundreds of researchers and bioinformatics trainers in LMICs.

Teaching based on Galaxy is also valuable in LMICs with limited internet access. The training datasets are all available online, either on the Galaxy server, in shared data libraries, or in a third-party service such as Zenodo. Thus, the learners own (poor) internet connections are bypassed for these bandwidth-heavy tasks. The internet is only used to launch the analysis steps and for the teacher to monitor the progress of the analysis in Galaxy via TIaaS.

For cases where internet access is frequently absent, the GTN offers Docker [46] images, allowing for completely offline training. Every topic has its own Docker image, preconfigured with all the necessary tools, workflows, tours, and data-libraries to complete the tutorials within that topic. A drawback of this solution is that all required compute resources to complete the tutorials must be available locally.

Citizen science and education. Thanks to its graphical user interface, Galaxy can also be used to introduce bioinformatics to a general audience and include them into scientific projects.

The Street Science Community [47] offers workshops to introduce biology, genomic sciences, and bioinformatics to the public. For example, participants extract yeast DNA out of beer bottles and sequence it using a MinION, the Oxford Nanopore sequencing device. The generated sequencing data are then processed by participants inside Galaxy: uploading data, running a metagenomic analysis workflow, and visualizing results. Combining lab work and bioinformatics data analysis with Galaxy vividly demonstrates the challenges and possibilities that genomics brings to our society.

The GTN has also provided an opportunity for the ecology community to experiment with several ways to extend citizen science schemes, for example, through 2 monitoring schemes from Vigie-Nature [48], French citizen science programs about common birds (Suivi Temporel des Oiseaux Communs) and bats (Vigie-Chiro). Volunteers use Galaxy to analyze data they gathered, as researchers would do, through user-friendly tools and workflows, documented in GTN. Such solutions increase the motivation of participants, especially volunteers, as they can analyze and visualize data about species and ecological systems they monitor, sometimes for several years.

Training in the COVID-19 era: Remote learning. Due to the global COVID-19 pandemic, many educators have found themselves forced to change the modality of their training activities to completely virtual events. Galaxy and the GTN cater to remote learners and teachers with a set of features to facilitate the online learning process [9]. It provides easy access to data and the possibility to share the progress and achievements, both student-to-student and student-to-instructor [6]. This has been extensively tested during the COVID-19 pandemic. For example, the GTN community organized the GTN Smörgåsbord events [49], global, 5 day, 24/7 events. These were fully asynchronous events, where instructors prerecorded GTN tutorials, which the 2,000+ registered participants could work through at their own pace, with 120 + instructors available online for support on Slack across all time zones. This approach has been repeated for the Galaxy Community Conference training week [50], as well as various other virtual training events, including (but not limited to):

1. A 5-day workshop on "Machine Learning using Galaxy" organized by the Galaxy Europe team in June 2020, with 400+ registrations and 200+ participants on the first day [51].
2. A 5-day plant transcriptomics workshop organized by the Galaxy Europe team [52].
3. A Microbiome Informatics 5-day workshop organized by the Galaxy-P team with Galaxy India [53].
4. A SARS-CoV-2 data analysis workshop [54].
5. A case study in COVID-19 data analysis at the Great Lakes Bioinformatics Conference (GLBIO 2021) [55].
6. Metatranscriptomics analysis using microbiome RNA-seq data in Galaxy [56].
7. The Spanscriptomics workshop, a pilot study into using Spanish-translated GTN tutorials to teach single-cell analysis to native Spanish speakers [57].

Hybrid learning in geographically sparse locations. Even before the pandemic, learners in remote areas have faced significant barriers to traveling to in-person training events. In such areas, a so-called *hybrid training* [19] approach offers a solution. In such an approach, learners gather in classrooms in several geographically distinct locations. The training is live-broadcast to each of these locations and communication with the instructor happens in real-time.

GTN resources have been successfully used for such hybrid training events, for example, by Galaxy Australia, which organized a large number of hybrid training events, with up to 11 satellite classrooms across the region participating, and reaching over 800 learners [19]. The Gallantries project [18] has successfully applied a similar approach in the European region.

Conclusion

From the start of the Galaxy project, education has been an important focus. Since 2018, more than 350 training events have been carried out. Recipients of these trainings included undergraduate and postgraduate students, research scientists, underrepresented communities, and citizens. The foundation for most of these events were the Galaxy tutorials provided by the GTN using its well-structured and easy to maintain Galaxy-based training resource.

Since then, the GTN has grown steadily and today more than 260 tutorials developed by more than 260 contributors across 23 topics are available. The GTN infrastructure is in accordance with the 4 As framework and the FAIR principles. It furthermore enables remote learning also in settings with intermittent, restricted, or unavailable internet access. Teachers are empowered to prepare and run their trainings by specialized training resources for educators and technical features such as TIaaS. Also, contribution to GTN has been facilitated by the creation of dedicated tutorials and technical tools such as the Planemo tutorial development kit (PTDK) and GitPod.

To support the instructors and build a community of instructors, we are also collaborating with the Gallantries project and ELIXIR to build a Train the Trainer (TtT) [58–60] program and a mentoring program for instructors. For contributors, we are also working with publishers to implement a system for publication of tutorial via article with limited extra work. Following the feedback from learners, instructors, and contributors, new features are currently in development.

The Galaxy Training Network is an example of a robust, effective Community of Practice [20] where fellowship within a domain promotes iterative knowledge sharing and ongoing professional development.

Acknowledgments

In loving memory of Simon Gladman (1970 to 2022), brilliant teacher and system administrator, beloved Galaxy community member, who contributed heavily to the Galaxy Training Network over the past 7 years.

This project would not be possible without the tireless efforts and valuable contributions from the worldwide GTN community. All contributors to the GTN; 2022. Available from: https://training.galaxyproject.org/training-material/hall-of-fame.

Author Contributions

Conceptualization: Saskia Hiltemann, Helena Rasche, Hans-Rudolf Hotz, Dave Clements, Florian Heyl, Björn Grüning, Bérénice Batut.

Formal analysis: Florian Heyl, Björn Grüning.

Funding acquisition: Saskia Hiltemann, Helena Rasche, Pratik D. Jagtap, Frederik Coppens, Dave Clements, Björn Grüning, Bérénice Batut.

Investigation: Bérénice Batut.

Methodology: Saskia Hiltemann, Helena Rasche, Bérénice Batut.

Project administration: Saskia Hiltemann, Helena Rasche, Yvan Le Bras, Bérénice Batut.

Software: Saskia Hiltemann, Helena Rasche, Simon Gladman, Hans-Rudolf Hotz, Delphine Larivière, Daniel Blankenberg, Pratik D. Jagtap, Thomas Wollmann, Anthony Bretaudeau, Nadia Goué, Coline Royaux, Yvan Le Bras, Subina Mehta, Anna Syme, Bert Droesbeke, Nicola Soranzo, Wendi Bacon, Fotis Psomopoulos, Cristóbal Gallardo-Alba, John Davis, Melanie Christine Föll, Matthias Fahrner, Maria A. Doyle, Beatriz Serrano-Solano, Anne Claire Fouilloux, Peter van Heusden, Wolfgang Maier, Florian Heyl, Björn Grüning, Bérénice Batut.

Supervision: Saskia Hiltemann, Helena Rasche, Simon Gladman, Daniel Blankenberg, Dave Clements, Bérénice Batut.

Validation: Saskia Hiltemann, Helena Rasche, Simon Gladman, Hans-Rudolf Hotz, Delphine Larivière, Daniel Blankenberg, Pratik D. Jagtap, Thomas Wollmann, Anthony Bretaudeau, Nadia Goué, Timothy J. Griffin, Coline Royaux, Yvan Le Bras, Subina Mehta, Frederik Coppens, Nicola Soranzo, Wendi Bacon, Fotis Psomopoulos, Cristóbal Gallardo-Alba, John Davis, Melanie Christine Föll, Matthias Fahrner, Maria A. Doyle, Beatriz Serrano-Solano, Anne Claire Fouilloux, Peter van Heusden, Wolfgang Maier, Florian Heyl, Björn Grüning, Bérénice Batut.

Visualization: Saskia Hiltemann, Helena Rasche, Bérénice Batut.

Writing – original draft: Saskia Hiltemann, Helena Rasche, Hans-Rudolf Hotz, Daniel Blankenberg, Coline Royaux, Nicola Soranzo, Bérénice Batut.

Writing – review & editing: Saskia Hiltemann, Helena Rasche, Hans-Rudolf Hotz, Daniel Blankenberg, Pratik D. Jagtap, Nadia Goué, Timothy J. Griffin, Coline Royaux, Bert Droesbeke, Nicola Soranzo, Cristóbal Gallardo-Alba, John Davis, Melanie Christine Föll, Matthias Fahrner, Beatriz Serrano-Solano, Peter van Heusden, Florian Heyl, Björn Grüning, Bérénice Batut.

References

1. Tomaševski K. Human rights obligations: making education available, accessible, acceptable and adaptable. Raoul Wallenberg Institute of Human Rights and Humanitarian Law; 2001.
2. Bambara CS, Harbour CP, Davies TG, Athey S. Delicate Engagement. Community Coll Rev. 2009; 36 (3):219–238. https://doi.org/10.1177/0091552108327187
3. Hara N, Kling R. A case study of students' frustrations with a web-based distance education. 1999.
4. Jaggars SS. Choosing between online and face-to-face courses: Community college student voices. Am J Dis Educ. 2014; 28(1):27–38.
5. Xu D, Jaggars SS. Performance gaps between online and face-to-face courses: Differences across types of students and academic subject areas. J Higher Educ. 2014; 85(5):633–659.
6. Serrano-Solano B, Erxleben A, Gallardo-Alba C, Rasche H, Hiltemann S, Föll M, et al. Fostering Accessible Online Education Using Galaxy as an e-learning Platform. Preprint. 2020. https://doi.org/10.20944/preprints202009.0457.v1
7. Huang HM. Toward constructivism for adult learners in online learning environments. Br J Educ Technol. 2002; 33(1):27–37.
8. Liaw SS. Considerations for developing constructivist web-based learning. Int J Instr Media. 2004; 31 (3):309.
9. Gallardo-Alba C, Grüning B, Serrano-Solano B. A constructivist-based proposal for bioinformatics teaching practices during lockdown. PLoS Comput Biol. 2021; 17(5):e1008922. https://doi.org/10.1371/journal.pcbi.1008922 PMID: 33983931
10. Bangert AW. The development of an instrument for assessing online teaching effectiveness. J Educ Comput Res. 2006; 35(3):227–244.

11. Garrison DR, Anderson T, Archer W. Critical inquiry in a text-based environment: Computer conferencing in higher education. Internet High Educ. 1999; 2(2–3):87–105.
12. Jalili V, Afgan E, Gu Q, Clements D, Blankenberg D, Goecks J, et al. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2020 update. Nucleic Acids Res. 2020; 48 (14):8205–8207. https://doi.org/10.1093/nar/gkaa554 PMID: 32585001
13. The Galaxy Community. The Galaxy ToolShed. 2022. Available from: https://galaxyproject.org/toolshed/.
14. Barone L, Williams J, Micklos D. Unmet needs for analyzing biological big data: A survey of 704 NSF principal investigators. PLoS Comput Biol. 2017; 13(10):e1005755. https://doi.org/10.1371/journal.pcbi.1005755 PMID: 29049281
15. Batut B, Hiltemann S, Bagnacani A, Baker D, Bhardwaj V, Blank C, et al. Community-Driven Data Analysis Training for Biology. Cell Systems. 2018; 6(6):752–758.e1. https://doi.org/10.1016/j.cels.2018.05.012 PMID: 29953864
16. Garcia L, Batut B, Burke ML, Kuzak M, Psomopoulos F, Arcila R, et al. Ten simple rules for making training materials FAIR. PLoS Comput Biol. 2020; 16(5):e1007854. https://doi.org/10.1371/journal.pcbi.1007854 PMID: 32437350
17. Wilson G. Software Carpentry: lessons learned. F1000Res. 2016; 3:62. https://doi.org/10.12688/f1000research.3-62.v2 PMID: 24715981
18. The Gallantries Project. The Gallantries Project. 2022. Available from: https://gallantries.github.io/.
19. Hall CR, Griffin PC, Lonie AJ, Christiansen JH. Application of a bioinformatics training delivery method for reaching dispersed and distant trainees. PLoS Comput Biol. 2021; 17(3):e1008715. https://doi.org/10.1371/journal.pcbi.1008715 PMID: 33735276
20. Li LC, Grimshaw JM, Nielsen C, Judd M, Coyte PC, Graham ID. Evolution of Wengers concept of community of practice. Implementation Sci. 2009; 4(1). https://doi.org/10.1186/1748-5908-4-11 PMID: 19250556
21. The Galaxy Community. Galaxy Platform Directory: Servers, Clouds, and Deployable Resources. 2022. Available from: https://galaxyproject.org/use/.
22. The Galaxy Community. Galaxy Community Hub website. 2022. Available from: https://galaxyproject.org/.
23. Via A, Palagi PM, Lindvall JM, Tractenberg RE, Attwood TK, Foundation TG. Course design: Considerations for trainers–a Professional Guide. 2020. Available from: https://f1000research.com/documents/9-1377.
24. Irons A, Elkington S. Enhancing Learning through Formative Assessment and Feedback. Routledge; 2021. Available from: https://doi.org/10.4324/9781138610514.
25. The G alaxy Community and the Gallantries Project. Gallantries & GTN Training Video Library. 2022. Available from: https://gallantries.github.io/video-library/.
26. Allaire J. RStudio: integrated development environment for R. Boston, MA. 2012; 770(394):165–171.
27. Ragan-Kelley M, Perez F, Granger B, Kluyver T, Ivanov P, Frederic J, et al. The Jupyter/IPython architecture: a unified view of computational research, from interactive exploration to communication and publication. In: AGU Fall Meeting Abstracts. vol. 2014; 2014. p. H44D–07.
28. Baumer B, Udwin D. R markdown. Wiley Interdiscip Rev Comput Stat. 2015; 7(3):167–177.
29. Devenyi GA, Emonet R, Harris RM, Hertweck KL, Irving D, Milligan I, et al. Ten simple rules for collaborative lesson development. PLoS Comput Biol. 2018; 14(3):e1005963. https://doi.org/10.1371/journal.pcbi.1005963 PMID: 29494585
30. ELIXIR. TeSS. 2022. Available from: https://tess.elixir-europe.org/.
31. The Galaxy Training Network. Training Philosophies. 2022. Available from: https://training.galaxyproject.org/topics/instructors/philosophies/.
32. Bloom BS. Taxonomy of educational objectives: The classification of educational goals. Cognitive Domain. 1956.
33. The Galaxy Europe Community. Galaxy Workflow Testing. 2022. Available from: https://github.com/usegalaxy-eu/workflow-testing/.
34. Rasche H, Gruening BA. Training Infrastructure as a Service. BioRxiv. 2020.
35. Bray S, Bernt M, Soranzo N, van den Beek M, Batut B, Rasche H, et al. Planemo: a command-line toolkit for developing, deploying, and executing scientific data analyses. bioRxiv. 2022.
36. GitPod Community. GitPod. 2022. Available from: https://gitpod.io/.
37. Ahea M, Ahea M, Kabir R, Rahman I. The Value and Effectiveness of Feedback in Improving Students' Learning and Professionalizing Teaching in Higher Education. J Educ Pract. 2016; 7(16):38–41.

38. Mehta S, Crane M, Leith E, Batut B, Hiltemann S, Arntzen MØ, et al. ASaiM-MT: a validated and optimized ASaiM workflow for metatranscriptomics analysis within Galaxy framework. F1000Res. 2021; 10:103. https://doi.org/10.12688/f1000research.28608.2 PMID: 34484688

39. de Koning W, Miladi M, Hiltemann S, Heikema A, Hays JP, Flemming S, et al. NanoGalaxy: Nanopore long-read sequencing data analysis in Galaxy. GigaScience. 2020; 9(10). https://doi.org/10.1093/gigascience/giaa105 PMID: 33068114

40. Tekman M, Batut B, Ostrovsky A, Antoniewski C, Clements D, Ramirez F, et al. A single-cell RNA-sequencing training and analysis suite using the Galaxy framework. GigaScience. 2020; 9(10). https://doi.org/10.1093/gigascience/giaa102 PMID: 33079170

41. Rasche H, Hiltemann S. Galactic Circos: User-friendly Circos plots within the Galaxy platform. GigaScience. 2020; 9(6). https://doi.org/10.1093/gigascience/giaa065 PMID: 32530465

42. Fahrner M, Föll MC, Grüning BA, Bernt M, Röst H, Schilling O. Democratizing data-independent acquisition proteomics analysis on public cloud infrastructures via the Galaxy framework. GigaScience. 2022:11. https://doi.org/10.1093/gigascience/giac005 PMID: 35166338

43. Bray S, Dudgeon T, Skyner R, Backofen R, Grüning B, von Delft F. Galaxy workflows for fragment-based virtual screening: a case study on the SARS-CoV-2 main protease. J Cheminformatics. 2022; 14 (1):1–13. https://doi.org/10.1186/s13321-022-00588-6 PMID: 35414112

44. Hernández-de Diego R, de Villiers EP, Klingström T, Gourlé H, Conesa A, Bongcam-Rudloff E. The eBioKit, a stand-alone educational platform for bioinformatics. PLoS Comput Biol. 2017; 13(9): e1005616. https://doi.org/10.1371/journal.pcbi.1005616 PMID: 28910280

45. Klingström T, de Diego RH, Collard T, Bongcam-Rudloff E. Galaksio, a user friendly workflow-centric front end for Galaxy. EMBnet J. 2017; 23(0):897. https://doi.org/10.14806/ej.23.0.897

46. Boettiger C. An introduction to Docker for reproducible research. ACM SIGOPS Oper Syst Rev. 2015; 49(1):71–79. https://doi.org/10.1145/2723872.2723882

47. The Street Science Community. The Street Science Community. 2022. Available from: https://streetscience.community.

48. Vigie Nature. Vigie Nature. 2022. Available from: https://www.vigienature.fr/.

49. The Gallantries Project. GTN Smörgåsbord: A Global Galaxy Course. 2021. Available from: https://gallantries.github.io/posts/2021/03/01/sm%C3%B6rg%C3%A5sbord/.

50. GCC2021 Organizers and the Global GTN Community. GCC2021 Training Week. 2021. Available from: https://galaxyproject.org/events/gcc2021/training/.

51. Freiburg Galaxy Team. Remote training using Galaxy. Lessons learned from our ELIXIR Galaxy Machine Learning Workshop. 2020. Available from: https://docs.google.com/document/d/1_sQocj98DxhgnyvtXbRvcXlV84T_I3K1rFmWrMuw6x0/preview.

52. Freiburg Galaxy Team. Plant Transcriptomics Analysis using Galaxy. 2021. Available from: https://galaxyproject.eu/posts/2021/05/03/plant-summary/.

53. CSIR-IMTech Galaxy-P. Analysis of Functions Expressed by Microbiomes. 2021. Available from: https://gallantries.github.io/galaxy-workshop/events/functional-microbiome-2021/.

54. Freiburg Galaxy Team. SARS-CoV-2 Data Analysis and Monitoring with Galaxy. 2021. Available from: https://galaxyproject.eu/event/2021-06-21-sars-cov-2-data-analysis-monitoring-training/.

55. Galaxy-P. Mass Spectormetry (MS)-based multi-omics analysis using the Galaxy-P bioinformatics platform: A case study in COVID19 data analysis. 2021. Available from: https://youtu.be/Ihu9a84nM78.

56. VIB Galaxy-P. Metatranscriptomics analysis using microbiome RNA-seq data in Galaxy. 2021. Available from: https://training.vib.be/all-trainings/metatranscriptomics-analysis-using-microbiome-rna-seq-data-galaxy.

57. Bacon W. Spanscriptomics: Análisis de células únicas usando Galaxy. 2021. Available from: https://gallantries.github.io/galaxy-workshop/events/spanscriptomics/.

58. Morgan SL, Palagi PM, Fernandes PL, Koperlainen E, Dimec J, Marek D, et al. The ELIXIR-EXCELERATE Train-the-Trainer pilot programme: empower researchers to deliver high-quality training. F1000Res. 2017; 6:1557. https://doi.org/10.12688/f1000research.12332.1 PMID: 28928938

59. Via A, Attwood TK, Fernandes PL, Morgan SL, Schneider MV, Palagi PM, et al. A new pan-European Train-the-Trainer programme for bioinformatics: pilot results on feasibility, utility and sustainability of learning. Brief Bioinform. 2017; 20(2):405–415. https://doi.org/10.1093/bib/bbx112 PMID: 29028883

60. McGrath A, Champ K, Shang CA, van Dam E, Brooksbank C, Morgan SL. From trainees to trainers to instructors: Sustainably building a national capacity in bioinformatics training. PLoS Comput Biol. 2019; 15(6):e1006923. https://doi.org/10.1371/journal.pcbi.1006923 PMID: 31246949