

Open access genomic resources for terrestrial arthropods

Vanessa L González¹, Amanda M Devine¹, Mike Trizna²,
Daniel G Mulcahy¹, Katharine B Barker¹ and
Jonathan A Coddington¹



Genome sequencing initiatives like the Arthropod i5k project and other biodiversity genomics research rely on access to high quality DNA and/or tissue. Global collection initiatives such as the Smithsonian Global Genome Initiative (GGI) and its partner network, the Global Genome Biodiversity Network (GGBN) aim to provide access to these resources at high-quality standards. Here, we review progress toward providing genomic resources (tissues, DNA, genomes) for terrestrial arthropods, a megadiverse animal group, and compare progress in genome sequencing to all other animals.

Addresses

¹Global Genome Initiative, National Museum of Natural History, Smithsonian Institution, Washington, DC 20013, USA

²Smithsonian DNA Barcode Network, National Museum of Natural History, Smithsonian Institution, Washington, DC 20013, USA

Corresponding author: González, Vanessa L (gonzalezv@si.edu)

Current Opinion in Insect Science 2018, 25:91–98

This review comes from a themed issue on **Insect genomics**

Edited by **Stephen Richards, Christopher Childers and Anna Childers**

<https://doi.org/10.1016/j.cois.2017.12.003>

2214-5745/© 2017 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Introduction

Biodiversity genomics in the last five years is an exploding field [1[•]]. Unlike the previous decade, when researchers usually sequenced genomes of ‘model’ organisms or other species for which large quantities of high molecular weight (HMW) DNA was readily available [2–6], current projects are increasingly ‘*de novo*’ sequencing of wild species, often of small body size and challenging to collect ideal genomic samples. Genomic research requires reliable and rapid access to HMW DNA and tissue samples. This stricture increasingly limits the rate of production of high quality genomes. Despite the importance of such samples, information about them is fragmented across the many repositories that preserve them.

The net implication is that coordinated sampling efforts, storage and documentation strategies, information on sample quality, and even simple databases of publically available tissue and collections are more urgent than ever. Cryo-collections are widely distributed around the world, held by institutions as well as individual researchers. It is imperative to adopt a global, single, data model for genetic resources so that users can understand where genetic resources are, their attributes, and how to access them. More strategically, global efforts to sample and sequence genomes can be coordinated phylogenetically, ecologically, and for myriad applied purposes so that a synoptic sample of the genomic diversity of life is preserved and available for scientific study.

Although many institutions recognize the importance of this mission, the Smithsonian Institution explicitly created the Global Genome Initiative (GGI, URL: ggi.si.edu) to foster an international collaboration to achieve these goals. Hereafter we will outline the data models and resources necessary to move toward accessible standardized biodiversity genomic collections as well as review the status of current collections in the context of terrestrial arthropods.

The Global Genome Initiative

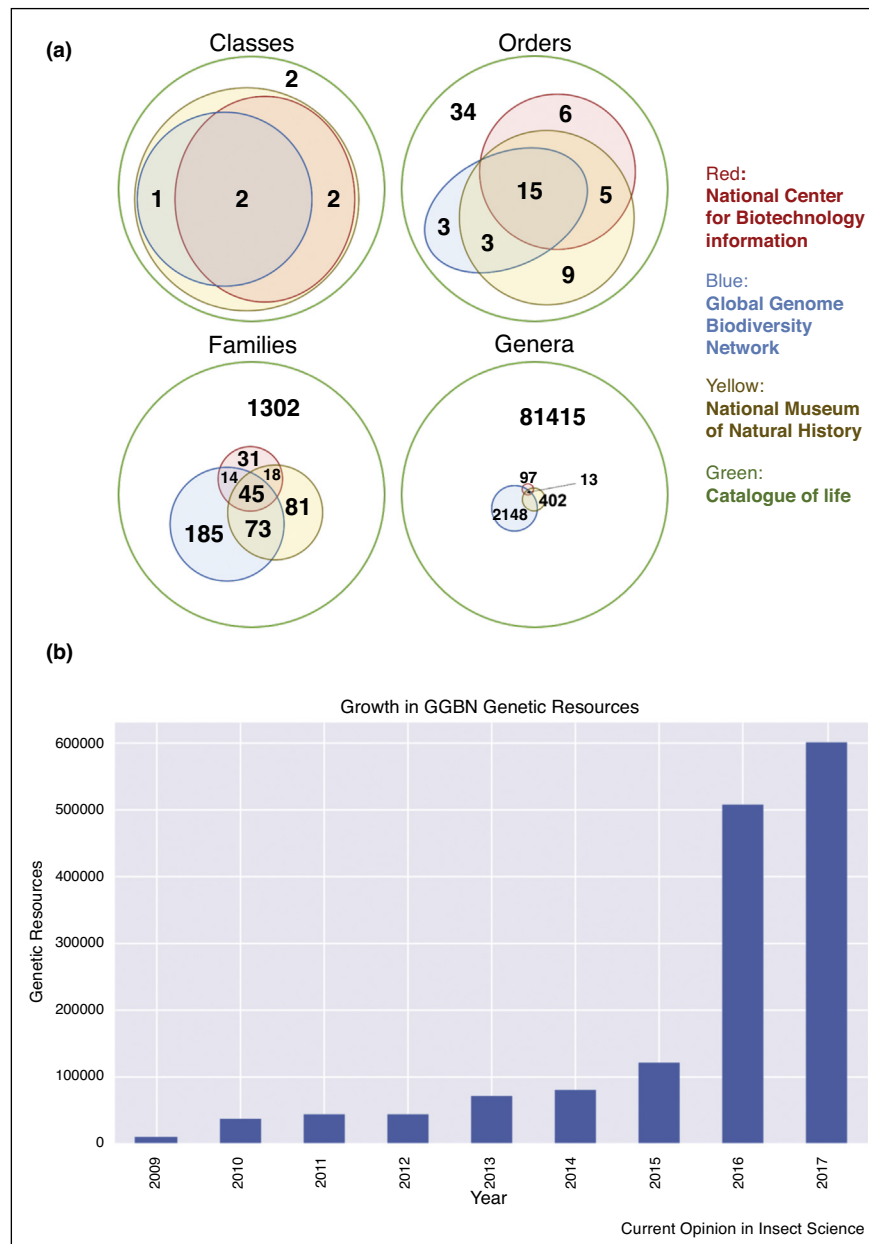
GGI’s mission is ‘to preserve and understand the genomic diversity of life.’ With partners, it aims to preserve a vouchered genomic sample of every taxonomic family on Earth and half of all genera. Partners include universities, research centers, government agencies, industry, museums, culture collections, botanical gardens, zoos, and, importantly, the many ambitious genome sequencing initiatives underway.

GGI’s first step was to engage with international partners to organize a network of biorepositories; the Global Genome Biodiversity Network (GGBN) [7[•]]. GGBN aims to be the federated index to all scientific genomic samples that exist, serving as the infrastructure for GGI programs and a resource for sequencing projects and other research.

Toward open access resources: the Global Genome Biodiversity Network

Formed in 2011, the Global Genome Biodiversity Network (GGBN, ggbn.org) aims to make high quality,

Figure 1



Available genomic resources: **(a)** counts of terrestrial arthropod taxa in four databases. The area of each circle is proportional to its count in each database. The counts in each overlap are accurate, but the overlap areas are approximate. All zero and some small overlaps were omitted for clarity; **(b)** growth of genetic resources for all of life discoverable on the Global Genome Biodiversity Network (GGBN) since 2009.

well-documented, legal, vouchered collections of DNA or tissue samples discoverable for research through its data portal (GGBN, URL: <http://data.ggbn.org>). GGBN has 68 members as of September 2017, from every continent except Antarctica, of which 18 are currently providing data to the portal. GGBN samples (all of life) have increased about 380% since January 2015, to 604k samples (September, 2017). GGBN and GGI's goals broadly overlap: GGI concerns itself more with the strategic questions posed by its mission, and GGBN prioritizes

professional care and informatics of non-human genomic samples, including both legacy as well as future biodiversity collections.

GGBN is a member-driven organization with the governance structure of a scientific society. Crucially, it contributed the data model for genetic resources as an extension to the two existing models for classical specimens, DarwinCore Archive [8] and ABCD [9]. GGBN provides a platform for members to collaborate to ensure

Download English Version:

<https://daneshyari.com/en/article/8878542>

Download Persian Version:

<https://daneshyari.com/article/8878542>

[Daneshyari.com](https://daneshyari.com)