<u>Data and Resource Sharing Policies</u>

<u>CBG Pricing: In-Kind</u>

The CBG is committed to the principles of rapid release and sharing of genomic data adopted by the international barcode community. Material that is processed in-kind through CBG internal funding will be considered part of a "Community Resource Project". As the producer of a community resource (the DNA barcode library), the CBG is expected to make genomic data publicly available through GenBank upon submission to BOLD (https://boldsystems.org/).

Provenance information of sampling sites, (i.e., collection dates and GPS co-ordinates that correspond to sample numbers assigned) submitted for DNA barcoding is critical for meeting CBG goals and a prerequisite for initiating molecular analyses.

All generated sequence data, as well as provenance and taxonomic data will be released immediately and will be used by BOLD data aggregation, analytical and reporting tools.

Recipient and Providers maintain editing control of project data but sequence, provenance, and taxonomic data will be visible to the public. Any errors in taxonomic assignment will be corrected through an ongoing iterative data validation process.

All sequence data contained in BOLD (including unpublished projects) are used by the BOLD identification engine to provide DNA-based taxonomic identifications to public users. Reports generated by the BOLD identification engine include similarity scores and tree-based identification with branch labels containing detailed taxonomy, broad geographic localization (to province level), and corresponding BOLD Process ID's and Sample ID's. Information on individual specimens (museum catalogue numbers and place of voucher deposition) and their detailed geographic origin are not disclosed through the BOLD identification engine.

Once specimen data have been submitted to BOLD, provenance information and images become partially available to the public online through the BOLD Taxonomy Browser at http://v3.boldsystems.org/index.php/TaxBrowser_Home. This information is used to generate summary statistics and illustrative distribution maps and does not disclose the contents of individual research projects and specimen data records.