



INTERNATIONAL CANCER GENOME CONSORTIUM

UPDATES TO
GOALS, STRUCTURE,
POLICIES & GUIDELINES

SECTION E.3 –
PUBLICATION POLICY

NOVEMBER 2011



ICGC

WWW.ICGC.ORG/POLICY_UPDATE_PUBLICATION

E. 3. Publication Policy

POLICY: The individual research groups in the ICGC are free to publish the results of their own efforts in independent publications at any time (subject, of course, to any policies of any collaborations in which they may be participating).

In their individual papers, Consortium participants will not be restricted to describing the methods developed for the project, but can and should expand into describing biological insights that arise from their analyses. To facilitate comparison of data among different groups participating in the ICGC, all publications by Consortium members should include explicit data on quality metrics, possibly including a common reference set of analytes agreed upon by the Consortium, e.g., nucleic acids derived from a common cell line or other source; such papers should also explicitly include a statement that the quality metrics are those that have been adopted by the ICGC (to promote their wide acceptance across the broad research community).

Users of Consortium data, whether members of the Consortium or not, should be aware of the publication status of the data they use and treat them accordingly. For example, all investigators, including other Consortium members, should obtain the consent of the data producers (this term includes clinical contributors and other members of an ICGC collaborative team) before using unpublished data in their individual publications, and the data producers should not unreasonably withhold this consent.

ICGC members will not have privileged access to data from other members of the Consortium. Rather, all data shared by the Consortium members will be obtained from the data that has been released to public databases.

Investigators outside of the ICGC are free to use data generated by ICGC members, either en masse or specific subsets, but are asked to follow the guidelines developed at the Ft. Lauderdale meeting. Specifically, data users should cite the source of the data and should acknowledge the clinical contributors and the data producers from the ICGC. In addition, the data users are asked to recognize the interests of the data producers to publish reports on the generation and analysis of their data, as described previously. Datasets from ICGC members are released to public databases as pre-publication data and remain unpublished until they appear in peer-reviewed publications. Outside investigators who perform an in-depth analysis of data from ICGC members and are interested in publishing a report before the data producers do so should discuss their results with the data producer(s) and are encouraged to establish collaborations. However, ICGC members are not required to collaborate with any outside investigators. All investigators, through their roles as journal and grant reviewers, should enforce a high standard of respect for the scientific contribution of the data producers.

This description of the ICGC data release policy is directed primarily at issues concerning the use of Consortium data in scientific publications. The intent of the policy is to accelerate the use of the data by the global scientific community while, at the same time, allowing the data producers to get appropriate scientific credit for their work through publications. To facilitate this goal, the data producers agree not to restrict the use of the data by others, while the data users are encouraged to act in a manner that is consistent with this policy providing unrestricted access to pre-publication data.

Clarification on the Policies of the ICGC regarding Pre-Publication Data Releases

Members of the ICGC are committed to the principles of rapid data release. However, like other large-scale genomic projects, members do not consider the deposition of its data into its own or public databases to be the equivalent of publication in a peer-reviewed journal. Therefore, although the data are available to others, the members still consider them to be formally unpublished.

The members of the ICGC have adopted data release policies that will accelerate the dissemination of datasets through rapid data releases that will precede publication of the global analyses of ICGC member datasets, anticipating the data will be useful for many investigators. The Consortium anticipates that the Project's data will be used in many ways, such as in developing new analytical methods, in understanding patterns of somatic mutations, and in guiding selection of mutations, genes, and other targets that will be used to develop biomarkers for cancer detection, diagnosis and prognosis, and new therapeutic interventions. Thus, the Consortium recognizes that the data are available to many users for a variety of purposes.

This document was generated for the ICGC members, the wider cancer research community, and journal editors/reviewers as a guide for the use of pre-publication data generated by ICGC members. An important difference between the ICGC and other projects that have adopted early data release principles (such as the Human Genome Project and the HapMap Project) is that the ICGC is not a single project, but a consortium of multiple projects that are led by different research teams, and have different timelines.

ICGC members have agreed to a harmonized approach to informing users about ICGC member publication plans. For convenience, ICGC-derived data analyses that are relevant to publications are divided into three categories:

Global analyses: These include genome-wide descriptions of cancer genomes, transcriptomes, epigenomes, and integrated datasets. Global analyses of such datasets generated by member projects can be subject to a publication moratorium should the member projects wish to do so. In addition to providing ICGC members the opportunity to be the first to publish global analyses of their project, ICGC members will have expert knowledge of the tumour types and are uniquely positioned to describe how the data was generated, including quality metrics.

Targeted analyses: Specific mutations, genes or pathways that can be identified by mining the datasets. These data cannot be protected by a publication moratorium period. While this information is valuable and may be an important motivation of data producers, it will have most value if it can be disseminated rapidly to a wide community of scientists. Targeted datasets included in manuscripts of third parties should be subject to additional analyses beyond simple descriptive summaries that can easily be obtained by browsing ICGC data.

Other analyses: It is possible that researchers testing new bioinformatics methods will require global analyses of pre-publication datasets, even if the intention is not to publish global descriptions of a cancer genome project. If this is the case, users are encouraged to use datasets that are not under a publication moratorium, or to contact the member projects who generated the datasets that are under

a moratorium. Possible outcomes may be 1) an agreement that users limit the publication on a subset of the data analysis (ie. patterns of mutations affecting a single chromosome), 2) collaboration and co-publication, or 3) submission of a third party manuscript that is simultaneous with a global analysis manuscript prepared by the member projects.

Time limits for publication moratoriums: All data shall become free of a publication moratorium when either the data is published by the ICGC member project or **one** year after a specified quantity of data (e.g. genome dataset from 100 tumours per project) has been released via the ICGC database or other public databases. In all cases data shall be free of a publication moratorium two years after its initial release. Since each data release will be identified by unique identifiers that will include a version name and date of release, users will be able to calculate the dates when ICGC project data is no longer subject to a publication moratorium. The quantity of data on which the initial global analysis will be carried out will vary between tumour types but is likely to be data on 100 tumour samples for more common tumour types and fewer for rarer tumour types. Details of any publication moratorium including the quantity of data on which the initial global analysis will be carried out will be specified by individual projects (please see below).

Data users should freely analyze pre-publication data and act responsibly in publishing analyses of those data in a manner that respects the terms described by the data producers. Data users are responsible for accurately citing the source of pre-publication data, including the version of the dataset. Data users are encouraged to contact the data producers to discuss publication plans in cases of uncertainty.

Scientific journal editors and reviewers should be aware that there may be restrictions on data users regarding ICGC datasets. It may be necessary to request that authors provide evidence that communication occurred with data producers, if there is a possibility that the manuscript infringes on the terms set by the data producers.

ICGC Open Access Publication Policy Information for collaborators

The members of the International Cancer Genome Consortium (ICGC) are committed to the principle of rapid data release to the scientific community and to disseminating all results of the Consortium's research and scholarship as widely as possible.

Publishing in open access journals, archiving papers in PubMed Central (PMC), and posting them on host institution websites are ways to provide public access to research findings. ICGC strongly encourages Consortium researchers to submit publications to journals with open access policies and to select a submission option that enables rapid deposit of papers in centralized repositories such as PMC (see <http://www.sherpa.ac.uk/romeo/> to search journals' publication policies).

In advocating rapid and unfettered access to scientific findings, ICGC is joining research organizations around the world that support public access to research results in order to disseminate knowledge as broadly as possible, promote and accelerate continued advances, and demonstrate the outcomes of funding. Funders such as the National Institutes of Health, the Wellcome Trust, and the Canadian Institutes of Health Research all require open access to papers showcasing research they have funded.

Many funders sponsor open access by defraying charges imposed by publishers. In addition, many publishers have adopted policies that allow for open access.

A summary of open access policies given by various research funders as part of their grant awards can be found here: <http://www.sherpa.ac.uk/juliet/>