



Somatic Mosaicism across Human Tissues (SMAHT) Publication Policy

Version 2.0.0 - Sep 4 2025

SMAHT Website: <https://smaht.org/>

SMAHT Data Portal: <https://data.smaht.org/>

I. Introduction

This document describes the SMAHT network policy for sharing manuscripts from the planning stage through submission to a journal. All manuscripts from SMAHT-funded activities (i.e., manuscripts acknowledging SMAHT support) EXCEPT for manuscripts arising from ELSI-related activities are subject to this policy.

This policy aims to disseminate research outcomes quickly and facilitate collaboration and coordination of research within the network. This policy encourages collaborations and the open sharing of information within the network. In keeping with this principle, sharing manuscript drafts and soliciting feedback from network members before submission is strongly encouraged.

Information shared internally must be confidential and may not be shared outside of the network.

II. Responsibilities

It is the sole responsibility of network members to take appropriate steps to protect any intellectual property rights to their work and to follow any relevant rules of their institutions.

The corresponding author is responsible for complying with this policy, which includes updating the manuscript tracking system when the manuscript's status changes (draft to submitted, submitted to published, etc.), corrections, or retractions after publication.

III. Manuscript Tracking and Concept Sheets

III.(a) Tracking sheet

An internal SMAHT manuscript tracking sheet will share information about manuscripts from network members. The tracking sheet will include specific SMAHT grant information, author lists, corresponding author email, manuscript title, manuscript status, manuscript, preprint submission links, date of publication, and other PubMed information. The primary purpose of the tracking sheet is to record the number and status of published or nearly published manuscripts for evaluating overall SMAHT network production. All publications that cite any SMAHT grant support MUST be entered into the internal SMAHT manuscript tracking system. The SMAHT Organizational Center (OC) will maintain this tracking system.

III.(b) Concept Sheet

Concept sheets are documents summarizing key details of a **proposed** project/publication. Investigators complete the A. SMAHT Concept Sheet Template as early as possible in the life

cycle of a manuscript. The information in a concept sheet is tentative and subject to change as the program evolves. Concept sheets include a project title, the names of involved investigators and SMAHT network centers, an overview of the project, and planned analyses, including material/datasets used, sequencing platforms used, and anticipated manuscript submission dates. Concept sheets are expected to change over time as the effort advances. Submitting a concept sheet does not preclude other network members from using the same dataset or performing analyses for a similar project; however, by identifying similar efforts early, the network hopes to foster collaborative efforts. The SMAHT Google Drive Folder containing the individual concept sheets is available [here](#), and a summary of all concept sheets can be found in the Internal SMAHT Publication Tracking Sheet on the Concept Sheet tab.

Towards the goal of strengthening the scientific impact of all publications arising from the SMAHT network, all **SMAHT members are encouraged to read and review the concept sheets** to identify opportunities for collaboration and possible conflicts. As much as possible, when conflicts or overlapping manuscripts are identified, individual SMAHT network members are encouraged to reach agreement.

Working Group (WG) chairs are responsible for periodically reviewing submitted concept sheets and for identifying potential overlaps or conflicts amongst the proposed manuscripts described in the concept sheets. The WG chairs are responsible for communicating with the parties in potential conflict and attempting to reach an agreement. If no agreement can be reached, a meeting with the Steering Committee (SC) chairs and NIH officers is convened to reach an agreement (**see Section VIII below**).

We anticipate most concept sheets to fit within the scope of an established WG, which will be the primary level where potential for conflict/opportunity for collaborations is identified. However, if a concept sheet does not clearly fit within the scope of an established WG, all concept sheets will be presented in some capacity to the network, either within a closely related WG or at a Steering Committee meeting. All investigators are encouraged to submit a concept sheet for **any** manuscript that acknowledges SMAHT support, but for some types of manuscripts, concept sheets must be presented for network-wide discussion and review. (**see IV Types of SMAHT Manuscripts, below**)

IV. Types of SMAHT Manuscripts

Coordinating SMAHT publication efforts is critical for maximizing the network's scientific impact. The primary goal of the SMAHT effort is to produce a catalog of somatic variation across a cohort of 150 "production donors." We define production donor data (PDD), as any data produced by a genome characterization center (GCC) or Tool and Technology Development (TTD) group that is derived from sequencing of a tissue designated as one of the 150 "production donors" that are distributed by the Tissue Procurement Center (TPC).

A central goal of this publication policy, and an agreed upon organizing principle from the in-person meeting in Bethesda in June 2025, is to **avoid publication of production donor data ahead of coordinated publication efforts around key SMAHT scientific milestones**.

Publication of PDD ahead of coordinated milestones would dilute the scientific impact of the SMaHT network, reducing chances of receiving support for future funding cycles beyond the first 150 donors.

Key SMaHT scientific milestones include, but are not limited to:

- Benchmark associate papers (**BAPs, see IV.(a) below**)
- First round production (**FRPs, see IV.(b) below**)
- Midway production (timing and donor number not yet determined)
- The final catalogue of 150 donors

A clear example of publishing ahead of key milestones and undercutting the network would be if a group attempted to publish data from the first 9 production donors, when the network decided that the first round would consist of 10 donors.

Other Key SMaHT scientific milestones may be decided upon, and the list above appended, at future SMaHT Steering Committee, in-person, or virtual meetings.

IV.(a) Benchmarking Manuscripts

Per discussion at the in-person meeting in Bethesda in June 2025, a detailed plan for submission of manuscripts arising from benchmarking work was developed. Benchmark associated papers (BAP) include publications arising from work on the HapMap DNA mixture, the COLO829 cell line mixture, the LB fibroblasts with associated induced pluripotent stem cell clones, and the designated benchmark tissue samples (“ST001-ST004”). The plan for benchmark manuscripts is as follows:

- First drafts for all BAPs are due August 1st, 2025. **All BAPs should have concept sheets submitted prior to this deadline**
- The deadline for submission of BAPs is September 15th, 2025
- There will be a single “flagship” publication describing the overarching rationale, benefits, and limitations of each of the benchmarking approaches.
- This flagship paper will include data produced by each of the 7 Working Groups (SNV, SV, MEI, Assembly/Pangenome, Duplex, Single-cell/PTA, and Functional). Representative data from each WG will be determined by the WG, but likely will represent the most important findings from their data and analyses, represented in a single figure. The flagship paper will also highlight differences between the working group analyses, when appropriate.
- A benchmark flagship writing group will include, but is not limited to, representatives from the DAC, the TPC, the OC, each GCC, the NIH, and each of the contributing working groups. **The Flagship benchmarking manuscript will have network-wide Banner authorship only** without individually named authors (see VI. Banner Authorship, below)
- There will be BAPs arising from each of the working groups. These manuscripts may have **individually named authors in addition to a working group-specific banner authorship** to list the names of individuals actively participating in the related working group. (see VI. Banner Authorship, below)

- Any additional BAPs are permitted, but these must conform to the timelines and be accompanied with a concept sheet, in order to be included in the planned singular submission of a compendium of BAPs to a single journal or family of journals

IV.(b) **First Round Production Donor Manuscripts**

Per discussion at the in-person meeting in Bethesda in June 2025, a detailed plan for submission of manuscripts arising from the “first round of production” (FRPs) was developed. Rather than specify a precise number of donors, the consensus was to set deadlines on data and have the number of donors included in the FRPs determined by dates, but we anticipate the FRPs to be based on 10-25 donors. The timeline for FRPs is as follows

- Decisions about a data freeze (number of donors to be included in FRP) will be made during the SMaHT Production Data Jamboree, to be held on October 14th, 2025
- Concept sheets for FRPs must be submitted by October 31st, 2025. All FRP related concept sheets must be presented at the Steering Committee meetings in November and December 2025.
- FRP related abstracts, rough drafts, and rough figures are due January 15th, 2026
- FRP-related first drafts are due March 1st, 2026
- FRP-related preprint submissions are due April 1st, 2026
- Exceptions to this timeline may be permitted for “late-breaking” analyses, but only with the approval of the Steering Committee chairs and NIH representatives

IV.(c) **General principles for manuscripts not part of benchmarking and 1st round production publication**

Aside from the explicit plans for benchmarking associated papers (BAPs) and first round production papers (FRPs), the general principles of encouraging collaboration and supporting the greater SMaHT network and its efforts apply. The requirement of completing and sharing concept sheets early in the cycle of developing a manuscript will reduce the chances of a publication that “pre-empts” or undercuts the network. However, the network also recognizes that SMaHT-supported manuscripts may arise from a lab or group supported by a single grant that does not need to be shared network-wide. We also explicitly recognize manuscripts that represent network-wide efforts. The definitions, requirements, and authorships for these three classes of SMaHT-related manuscripts are as follows:

- **Network-Wide Manuscripts (NWMs)**
Examples include the marker paper (PMID 40604182), the planned benchmark and first round production flagship manuscripts, and other network-wide efforts arising from key scientific milestones (e.g., “midway production” and flagship associated with all 150 donors). Other than the initial marker paper (PMID 40604182), **these papers will always have network--wide Banner authorship only** without additional individually named authors.
- **Multi-Grant Manuscripts (MGMs)**

These publications describe efforts and analyses originating from two or more SMaHT-related grants. By definition, **any manuscript describing or reporting sequencing data derived from production donors is a multi-grant manuscript (MGM), and a concept sheet is required to be submitted and distributed for discussion at the working group/steering committee level.** Concept Sheets must be submitted as soon as feasible for these types of manuscripts. All manuscripts arising from working groups are considered MGMs. At the discretion of the WG and parties involved, MGMs may be published as a singular paper or as coordinated, complimentary publications (“joint submission”).

- **Single Project Manuscripts (SPMs)**

Although the majority of SMaHT-associated publications will fall under the MGM category, we recognize that some SMaHT-supported manuscripts do not require broad network communication. Although concept sheets and network distribution prior to publication are not required for SPMs, submitting concept sheets is still encouraged in the interest of collaboration and communication. The primary eligibility criteria for SPMs is that the manuscript would not be perceived by any network member as undermining other, broader network-wide manuscripts, which would diminish the network’s scientific impact.

A clear example of this type of publication would include SMaHT-funded projects publishing a novel method (experimental or computational). The use of production donor data, which by definition is produced by multiple grants, may **only** be used for SPMs if publication of that data could **not** be viewed as undermining other network-wide manuscripts. Examples of the type of data that would undermine broader network-wide manuscripts include, but are not limited to, the number or identity of somatic mutations within a donor, their tissue distribution, or mutational signatures. Examples of manuscripts that would **not** be appropriate for SPMs would include papers describing somatic mutations across two or more tissues from one or more donors.

This restriction on the use of specific production donor datasets is lifted after initial publication of the specific production donor dataset in the context of an MGM or NWM.

V. Preprints

Preprints of submitted manuscripts must be shared with the public via an appropriate public preprint server (such as bioRxiv, arXiv, ChemRxiv) before or concurrently with the time of submission to a journal unless the Executive Committee grants an exception or extension. The pre-print coordination should follow the same guidelines as SPM and MGMs described in Part IV. For example, if two or more labs plan to coordinate the timing of a batch of MGMs, they should also coordinate preprint submission timing. The manuscript tracking system should be updated with a link to the preprint once it becomes available.

Exceptions or extensions may be requested in cases where sharing of preprints is not allowed due to legal requirements related to the protection of intellectual property or when the target journal does not allow preprints to be shared before publication.

VI. Authorship & Acknowledgements

VI.(a) Authorship Criteria & Order

Authors of all manuscripts should meet the common criteria for authorship used by journals:

- Must have contributed significantly to scholarly effort in conception, design, data generation or analysis.
- Was involved in drafting, reviewing, and/or revising the manuscript.
- Must approve of the manuscript.

Given the role of SMAHT network Program Consultants, these individuals are in conflict and are not eligible to be included as authors on manuscripts arising from the SMAHT network.

Lead authors may exercise discretion in deciding whether an author meets all criteria for authorship, and may include additional authorship criteria if necessary for their manuscript. Authorship order is based on criteria consistent with editorial policies published by major journals in biomedical sciences and the Vancouver Protocol. A major consideration for authorship order is the effort and contribution of participating members.

VI.(b) Banner Authorship

“Banner authorship” refers to the practice of indexing the names of more than one author under a single name, emphasizing the group or network contribution. This section describes the different types of banner authorships recognized by the SMAHT network, and explicitly lists procedures for determining who is eligible to be included in a banner, and how listing of BOTH a banner and individually named authors is handled.

Consistent with the National Library of Medicine (NLM) authorship verification guidelines, **full authors' names** in the banner must be available in the online PDF version of the article. Banner author names are placed in the **published full text** of the article when the journal permits, but can be placed in an associated published supplementary file (not linked out). In this case, statements about how to find this supplementary file **must** be placed (1) in the byline area, (2) at the bottom of the first page of an article, or (3) near the acknowledgements area at the end of an article.

Currently recognized banner authorships include, but are not limited to:

1) Network-wide banner

This banner includes all members of the SMAHT network recognized by the OC **at the time of manuscript submission**. The format of this banner will be “The Somatic Mosaicism across Human Tissues Network”, and the names of the included authors will

be included somewhere within the manuscript, organized under center, contact principal investigator, and corresponding grant activity. NIH members are also included within this banner, which will only be used for Network Wide Manuscripts (NWMs) and will NOT list any additionally named authors. NIH staff must obtain approval according to Institute and NIH procedures prior to manuscript submission in order to be included in Banner Authorship.

--- First (title) page of paper ---

The Somatic Mosaicism across Human Tissues Network*

*A full list of SMaHT Network authors and their affiliations is available at the end of the article.

--- Separate section at the end of the paper ---

The Somatic Mosaicism across Human Tissues Network

National Institutes of Health

Author a¹, Author d¹

Organizational Center - Ting Wang (U24NS132103)

Author a¹, Author d¹

Tissue Procurement Center - Tom Bell (U24MH133204)]

Author a¹, Author d¹

Data Analysis Center- Peter Park (UM1DA058230)

Author a¹, Author d¹

Genome Characterization Centers

University of Washington and Seattle Children's Research Institute – James T. Bennett (UM1DA058220)

Author a¹, Author d¹

<... other GCCs listed in same manner...>

Tool and Technology Development

University of Michigan – Ryan E. Mills (UG3NS132084)

Author a¹, Author d¹

<... other TTDs listed in same manner...>

2) Working group banners

These include all members of a working group (WG) recognized by the OC **at the time of manuscript submission**. It is the responsibility of the WG chairs to ensure that the OC listed WG members are up to date and accurate. The names of the included authors will be included somewhere within the manuscript. Since WGs are, by nature, collaborations across different centers, these authors will be listed in alphabetical order, by last name and will not be organized by center or grant activity. The names of the currently recognized WG banners are listed below, but additional WGs may arise over time.

- SMaHT Network Single Nucleotide Variant (SNV) Working Group

- SMAHT Network Structural Variant (SV) Working Group
- SMAHT Network Mobile Element Insertion (MEI) Working Group
- SMAHT Network Assembly/Pangenome Working Group
- SMAHT Network Duplex Working Group
- SMAHT Network Functional Working Group
- SMAHT Network Single-cell PTA Working Group?

Unlike Network Wide Banner authorship, **manuscripts utilizing working group banners may choose to list individually named authors in addition to the WG banner**, to better represent some author's contributions. Individually named authors will, in most cases, also be listed under the WG banner.

VI.(c) Acknowledgements & Citations

All papers from the SMAHT network and/or NIH members need to cite the Common Fund award that supported the science. For example:

This research is supported by the NIH Common Fund, through the Office of Strategic Coordination/Office of the NIH Director under awards U24 MH133204, U24 NS132103, UG3

NS132024, UG3 NS132061, UG3 NS132084, UG3 NS132105, UG3 NS132127, UG3 NS132128, UG3 NS132132, UG3 NS132134, UG3 NS132135, UG3 NS132136, UG3 NS132138, UG3 NS132139, UG3 NS132144, UG3 NS132146, UM1 DA058219, UM1 DA058220, UM1 DA058229, UM1 DA058230, UM1 DA058235, and UM1 DA058236.

All papers from the SMAHT network must acknowledge the data access and availability of the SMAHT data, similar to the example statements below:

Example 1:

All protected (e.g., sequencing data, germline variant calls, and complete donor metadata for the SMAHT Production donors) and open (e.g., gene expression quantification, tables, etc.) access data are available in dbGaP (phs004193 for the SMAHT Benchmarking data; phs004104 for the SMAHT Production data). The SMAHT Data can be accessed from the SMAHT Data Portal at [SMAHT Data Portal](#).

Example 2:

This study was conducted as part of the NIH Common Fund consortium initiative, Somatic Mosaicism across Human Tissues (SMAHT). The [benchmark | production] datasets described in this study are available through dbGaP (<http://www.ncbi.nlm.nih.gov/gap>) under the study accession number, phs[004193 for the SMAHT Benchmarking data; 004194 for the SMAHT Production data]. The data used in this work was provided by the SMAHT Data Analysis Center (DAC) [1UM1DA058230] on behalf of the SMAHT network. More information about the SMAHT network is available online at <https://smaht.org/>, about the SMAHT Data Portal at [SMAHT Data Portal](#), and types of data generated by the Network at <https://data.smaht.org/about/consortium/data>

Publications and/or presentations using SMaHT data must also (a) cite the SMaHT Marker Paper (**PMID 40604182**), (b) indicate which released SMaHT dataset was used, and (c) cite all relevant publications and preprints describing the data being used. SMaHT data producers who generated SMaHT data used in publications and/or presentations must be acknowledged. An example of acknowledgment text that can be used is below:

Some data used in this work is from a NIH Common Fund-supported effort, the Somatic Mosaicism across Human Tissues (SMaHT) network, and is provided by the SMaHT Data Analysis Center (DAC) [1UM1DA058230] on behalf of the SMaHT Network. More information about the SMaHT Network is available online at [SMaHT Network](#).

VII. Disagreements and Conflict Resolutions

The Steering Committee (SC) co-chairs & NIH program officers will endeavor to resolve authorship disagreements, including banner and non-banner authorship status, authorship order, and any other publication-related disagreements through *ad hoc* discussions

If the SC co-chairs cannot reach a satisfactory resolution, as a last resort the issue will be addressed through the dispute resolution process stated in the SMaHT notices of funding which is as follows:

Any disagreements that may arise in scientific or programmatic matters (within the scope of the award) between award recipients and the NIH may be addressed by convening a Dispute Resolution Panel. It will be composed of three members: a designee of the SC chosen without NIH staff voting, one NIH designee, and a third designee with expertise in the relevant area who is chosen by the other two; in the case of disagreement for one award, the first member may be chosen by that recipient. This special dispute resolution procedure does not alter the recipient's right to appeal an adverse action that is otherwise appealable in accordance with PHS regulation 42 CFR Part 50, Subpart D and DHHS regulation 45 CFR Part 16.

VIII. Citations

The network acknowledges that network members consider public preprints as publications for all practical purposes and should be cited and credited like papers. Preprints establish priority in the field and can be included in progress reports, CVs, and lists of publications.

IX. Open Access:

At a minimum, publications must comply with the NIH Public Access policy (<https://publicaccess.nih.gov/>). To maximize the impact of the network's efforts, authors should endeavor to publish their work using open-access policies whenever possible.

X. Policy Updates

This policy will be reviewed and updated semi-annually if necessary.

XI. Contributors

First Name	Affiliation	Group	Email Address
Jimmy Bennett, Co-Chair	SCRI-UW	GCC	jtbenn@u.washington.edu
Lucinda Fulton, Co-Chair	Washington University	OC	lfulton@wustl.edu
Heather Lawson, Co-Chair	WUSM	OC	lawson@wustl.edu
Alexej Abyzov	Mayo Clinic	TTD	abyzov.alexej@mayo.edu alexej.abyzov@yale.edu
Tom Bell	NDRI	TPC	tbell@ndriresource.org
Elizabeth Chun	Harvard Medical School	DAC	elizabeth_chun@hms.harvard.edu
Richard Conroy	OSC/NIH	NIH	richard.conroy@nih.gov
Melissa Faith	TPC/Johns Hopkins	ELSI Co-PI	mfaith1@jhmi.edu
Fei Chen	Broad Institute	TTD	chenf@broadinstitute.org
Amy Lossie	NIDA/NIH	NIH	lossieac@nih.gov
Stephen Montgomery	Stanford (Broad GCC)	GCC	smontgom@stanford.edu
Jill Morris	NINDS	NIH	jill.morris@nih.gov
Peter Park	Harvard Medical School	DAC	peter_park@hms.harvard.edu
Dena Procaccini	NIH	NIH	dena.procaccini@nih.gov
Raquel Hernandez	TPC/Johns Hopkins	ELSI Co-PI	raquel.hernandez@jhmi.edu
Alexi Runnels	New York Genome Center	GCC	arunnels@nygenome.org
Geetha Senthil	NCATS/NIH	NIH	senthilgs@mail.nih.gov
Alex Urban	Stanford University	TTD	aeurban@stanford.edu
Jeffrey Ou	UW-SCRI	GCC	Jeffrey.Ou@seattlechildrens.org
Kristin Ardlie	Broad Institute	GCC	kardlie@broadinstitute.org
Mehrnoosh Ahmadi	NIH	NIH	mehrnoosh.ahmadi@nih.gov
Ron Johnson	NCI	NIH	rjohnso2@mail.nih.gov
Pamela Birriel	OSC-Common Fund	NIH	pamela.birriel@nih.gov

XII. SMaHT Concept Sheet Template

Submission Date	
Date of most recent revision/update to concept sheet	
Project Title	
Working Group(s)	
Tentative Project Leads & Contact Email	
Tentative Investigators Involved	
Are you using data from SMaHT GCCs/TTDs other than your own?	☐ Yes ☐ No
Tentative GCC / DAC / TPC / OC / TTD Involvement	
Manuscript Summary	
Planned Analyses	
Material / Datasets Used	☐ COLO829 Mixture (BLT-50, -BL, -T) ☐ Benchmarking Donors ☐ HapMap Mixture ☐ Production Donors ☐ iPSC & Fibroblast ☐ Other (describe here)
Sequencing Platforms	☐ Illumina short-read WGS ☐ Single-cell Sequencing ☐ ONT Standard/UL WGS ☐ Duplex Sequencing ☐ PacBio HiFi ☐ RNA Sequencing ☐ Other (describe here)
Milestones <i>(List dates for project duration, draft completion, and manuscript submission)</i>	Project Duration:
	Draft Completion:
	Manuscript Submission:

