ASTROLABE

MASS CYTOMETRY (CYTOF®) BARCODING SURVEY

El-ad David Amir
Astrolabe Diagnostics, Inc.

Mass cytometry is an ever-evolving field: new methods and reagents are released on a regular basis along with novel implementations of existing tools.

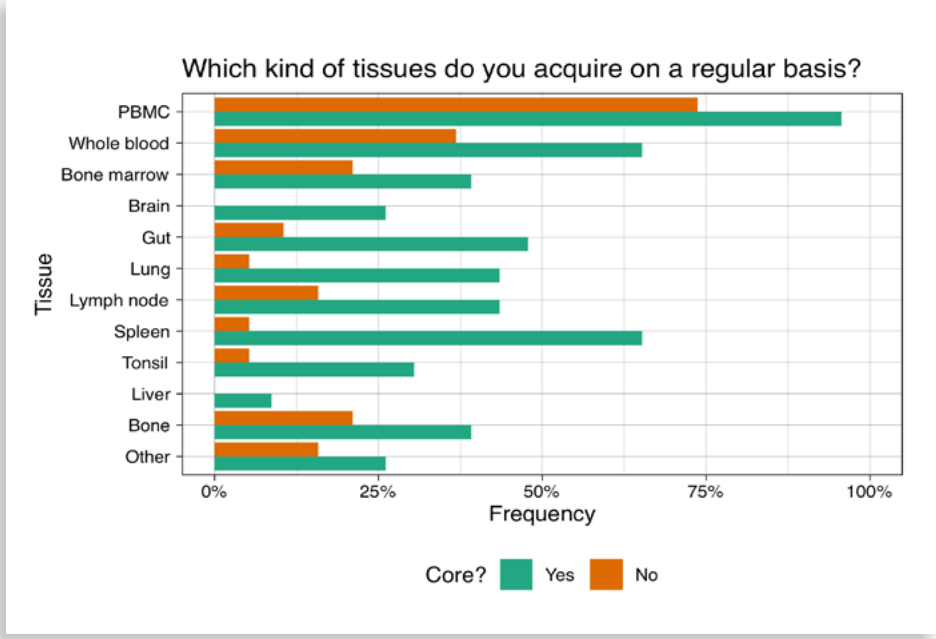
Barcoding has been a mainstay of the technology since its early days: first with variants of Palladium barcoding (see here and here for examples, along with the Fluidigm Cell-ID™ 20-Plex Pd Barcoding Kit), followed by CD45 barcoding, beta-2-microglobulin (B2m) barcoding, and new debarcoding tools (such as 2017 Zunder et al. and premessa).

Like any other protocol, operators have different preferences and variants when it comes to barcoding. We have conducted a survey with the goal of assessing the different conceptions and ideas researchers have of barcoding in the context of mass cytometry.

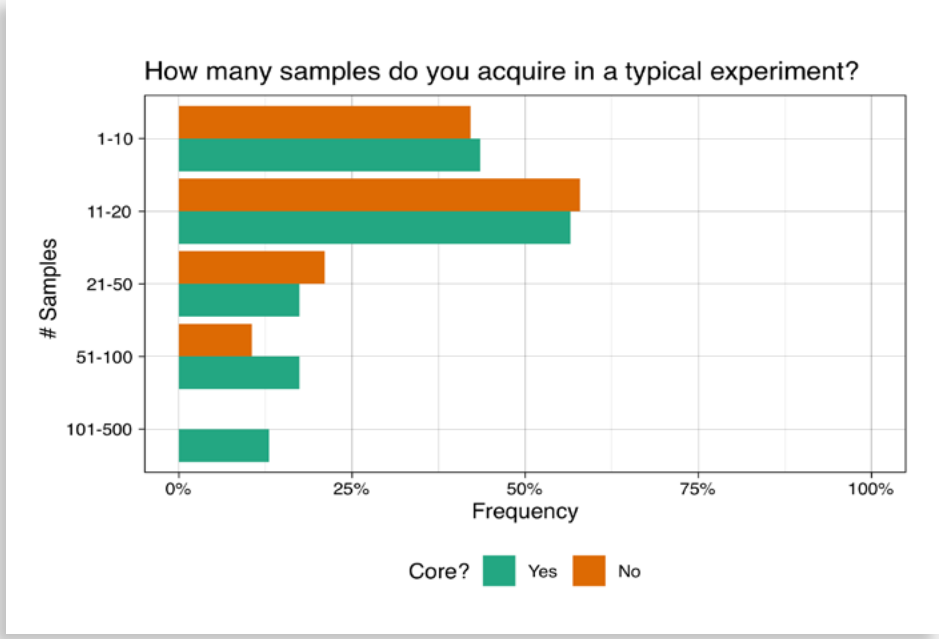
Sign up for the Immune Monitoring Biweekly for more news and content about the world of immune monitoring.

EXPERIMENTAL CONTEXT

The seminal mass cytometry paper involved bone marrow samples. Since then, the CyTOF has been used for the analysis of a plethora of tissues, which is reflected in the high variety of responses:



With regards to sample count per experiment (where a sample is defined as a debarcoded FCS file), the responses were homogeneous between core facilities and labs:

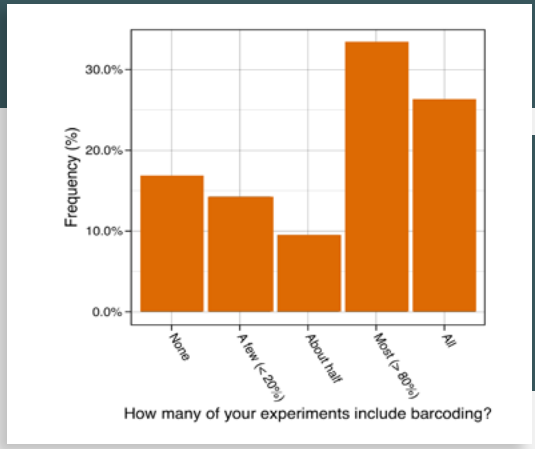


PBMCs and whole blood still dominate, with almost all cores and most labs acquiring them on a regular basis. Blood is straightforward to collect from patients and many studies include it either as a control or as a baseline when comparing to other tissues. In addition, each of brain, gut, lung, lymph node, spleen, tonsil, liver, and bone are being acquired by at least two researchers each; the "other" category included diverse responses such as phytoplankton, cell lines, and tumor-infiltrating lymphocytes (TILs). Not surprisingly, cores deal with a wider variety of tissues as they supply the needs of whole institutions rather than single labs.

When asked about the utilization of mass cytometry in clinical contexts, 21 (50%) of responders stated that they or their team use the CyTOF in the context of clinical trials and 17 (40.5%) use mass cytometry in the context of diagnosis, treatment, or disease monitoring in patients.

Cores and lab seem to acquire experiments of similar sample counts, with over 75% of respondents acquiring between 1 and 20 samples per experiment. Several factors could contribute to this number. One, the Fluidigm Pd Barcoding Kit is limited to a maximum of 20 samples which might guide researchers to these counts. Two, mass cytometry is often used for preliminary or cutting-edge research, where patient and specimen availability might be lower. Three, larger experiments might be financially infeasible to some labs.

BARCODING HABITS



Most researchers barcode their samples on a regular basis:

25 (59.5%) of responders stated that they use barcoding in more than half of their experiments and 11 (26.2%) barcode all of their experiments.

The most popular barcoding scheme is the Fluidigm Pd Barcoding Kit: 25 (71.4%) of the responders who barcode their data use it. The kit is limited to a maximum of 20 samples which is reflected in the sample count: 32 (91.4%) of responders acquire at most 20 samples in barcoded experiments. When asking about other barcoding schemes, the numbers are much lower, with 13 (37.1%) of responders using CD45 barcoding, 5 (14.3%) using B2m barcoding, and 4 (11.4%) using homebrew Palladium barcoding.

We next asked responders which benefits they see in barcoding, listing five potential reasons that we identified based on literature and discussions with researchers:

What are you reasons for including barcoding?	N	%
Better control of technical variability	33	94.3%
Economic use of reagents	23	65.7%
Generally high data quality	18	51.4%
Faster acquisition times	18	51.4%
Inclusion of a reference spike-in	16	45.7%

Not surprisingly, the top reason (which was chosen by almost all responders) for barcoding is better control of technical variability. Barcoding eliminates any signal variability due to pipetting the antibody cocktail into the samples. Furthermore, since cells are mixed prior to acquisition, any signal shifts or instabilities will be distributed across all samples, negating that source of variability as well. Many responders take this a step further by using barcoding for the inclusion of a reference spike-in: PBMCs from a large blood draw that came from a single healthy donor or by mixing multiple healthy donors. The spike-in can then be used for quality control purposes.

Two additional benefits pertain to the economics behind barcoding. One, it allows better utilization and less waste of antibodies and other reagents. Two, it enables faster sample acquisition by eliminating the downtime when switching between samples. Together, these make mass cytometry accessible to a wider crowd, especially when factoring in savings that might be available thanks to a fee-for-service facility.

AVOIDING BARCODING

17 (40.5%) of responders stated that they use barcoding in half of less of their experiments, and 7 (16.7%) don't include it at all. There are various reasons for avoiding barcoding in a given experiment:

What are your reasons for not including barcoding?	N	Freq
Low sample count that does not justify it	18	52.9%
I acquire fresh samples as they come in	10	29.4%
Cost	9	26.5%
Technically difficult to use (due to pipetting, protocol, etc.)	7	20.6%
Concerns over cross-contamination of samples	7	20.6%
Barcoding leads to lower-quality results	6	17.6%
High frequency of dead cells	6	17.6%
Low barcoding efficiency	5	14.7%
Concern about sample clumping	4	11.8%
High frequency of debris	4	11.8%
Debarcoding is time-consuming or inaccurate	3	8.8%

First and foremost, a researcher might have a low sample count that does not justify barcoding. 18 (42.9%) of responders indicated that they acquire between 1 and 10 samples in a typical experiment. Many do not consider that enough to barcode, especially when factoring in other potential costs. Several responders picked the financial cost of barcoding as a hindrance.

Next, there are several experimental constraints that could make barcoding more difficult or even impossible. Chiefly, many researchers acquire fresh samples as they come in and do not or cannot freeze them in order to allow pooling. This is especially critical with certain tissues (such as brain) or when looking for specific subsets (such as myeloids) which are sensitive to cryopreservation. Other experimental issues include high frequency of dead cells or debris, which could lead to spurious debarcoding and reagent waste, or sample clumping which could complicate debarcoding.

Barcoding carries various technical challenges. Several responders find the protocol difficult or time consuming, due to pipetting, debarcoding software, or other reasons. Barcoding could have low efficiency in some settings, again making debarcoding tricky. Finally, several responders do not trust the technology, either due to code cross-contamination or lower-quality results.

DISCUSSION

I was pleasantly surprised by the high percentage of responders who use barcoding in most, if not all, of their experiments. From an analysis perspective, barcoding addresses two significant confounding effects. First, it does an excellent job of identifying debris and doublet from the data, often removing all debris and over 90% of doublets. The alternatives, such as using event length, Ir191/193, or the Helios Gaussian channels, are not as effective. Additionally, barcoding provides an elegant solution to batch effects, completely eliminating that source of cross-sample variation. Whether your follow-up involves traditional gating or bioinformatics, barcoding improves the quality of the data you will start with.

The Fluidigm Pd Barcoding Kit is a powerful tool for researchers: It does not require any changes to the panel, comes pre-titrated by the vendor, follows a well-established protocol, and includes an analysis solution for debarcoding the resulting samples as part of the CyTOF acquisition software. Furthermore, while it does carry a non-trivial price tag, that is offset by the savings in reagents and machine time. When factoring that along with the measurable improvements in data quality, I think that even more experiments could integrate barcoding into their design.

Furthermore, I suspect that many of the objections to barcoding come from lack of familiarity with alternative schemes. For example, a common paradigm that involves fresh samples includes multiple tissues from a single donor, such as blood, tumor, and neighboring tissue. In such cases, the researcher could use CD45 or B2m barcoding to pool the samples and benefit from barcoding. The same techniques can be used for low-sample experiments – a 6-sample experiment can be barcoded by a four-channel code. Of course, none of the above could help with some technical challenges, such as digested tissue with a high debris count.

Overall, barcoding encompasses several protocols which are highly diverse and can fit a wide variety of experiment designs. Researchers should consider the pros and cons for it as an integral part when designing their CyTOF experiment.

ACKNOWLEDGEMENTS

I would like to thank Michael Leipold for his advice on the survey and for his permission to post the survey on CyTOForum. I would like to thank Paul Robinson for his permission to post the survey on the Purdue cytometry mailing list. Last but not least, I would like to thank the community members who replied to the survey. Your contribution is greatly appreciated.

We have disseminated the survey through CyTOForum, the Purdue cytometry mailing list, Astrolabe customers, and through LinkedIn and Twitter. It was available for two weeks, starting on September 11, 2019.

Overall, 42 researchers answered the survey. 37 (88.1%) of the responders work in academia, 3 (7.1%) in a government organization, and 2 (4.8%) in industry. 23 (54.8%) work for a core or a similar facility (where they provide services to other labs or groups in the organization). The following table describes the various positions of the survey responders:

Which of the following best describes your position?	N	%
Core operator	1	2.4%
Director	1	2.4%
Manager	7	16.7%
Post-baccalaureate Fellow	1	2.4%
Post-doctoral Fellow	5	11.9%
Principal Investigator	4	9.5%
Research Assistant/Associate	6	14.3%
Scientist	11	26.2%
Student (non-graduate)	1	2.4%
Student (Ph.D.)	3	7.1%
Technician	2	4.8%