UNIVERSITY OF CALIFORNIA
SANTA CRUZ

TOOLS FOR LARGE AND DETAILED EXPERIMENTS IN GENOMICS AND TISSUE DEVELOPMENT

A dissertation submitted in partial satisfaction of the
requirements for the degree of

DOCTOR OF PHILOSOPHY

in

BIOMOLECULAR ENGINEERING & BIOINFORMATICS

by

Yohei M. Rosen

September 2022

The Dissertation of Yohei M. Rosen
is approved:

Professor David Haussler, Chair

Professor Benedict Paten

Professor Mircea Teodorescu

Professor Richard Green

Peter Biehl
Vice Provost and Dean of Graduate Studies

Table of Contents

List of Figures

Abstract

Tools for large and detailed experiments in genomics and tissue development

by

Yohei M. Rosen

In this dissertation I present algorithmic and data representation advances in genomics as well as tools for a new bioinformatic approach to mammalian cell culture experiments which I call highly instrumented cell culture. The first section deals with fast variants of the forward algorithm for the Li and Stephens copying model of haplotypes derived from a population. I introduce a direct optimization of the Li and Stephens model forward algorithm which performs the identical calculation, without any approximations, but achieves this in average case sublinear time. This is an improvement over the classical algorithm which is at best linear time. I achieve this by using a sparse representation of the population haplotypes and by introducing an efficient lazy evaluation scheme. I also introduce a generalization of the recombination modeling component of the Li and Stephens model which operates on haplotypes and populations encoded in variation graphs. The second section deals with algebraic representations of genetic sites in variation graphs. I introduce the concept of the *bundle*, a motif in bidirected graphs which leads to a well defined concept of adjacency of sets of nodes. This allows a granular decomposition of the graph into sites which extends prior work on ultrabubbles and snarls previously reported by Paten et al. Lastly, I introduce the concept of highly instrumented cell culture and some technologies to enable it. I demonstrate a low-cost, robust, arbitrarily scalable microscope array for simultaneous parallel continuous time-series microscopy. I demonstrate

new approaches to rapid prototyping of labware and fluidic actuators. I also demonstrate principles and implementation of incubator-free cell culture, which is my approach to cell culture in media containing carbonic acid-carbonate-bicarbonate buffer systems without using any carbon dioxide rich gas chamber. I finally describe how these technologies integrate together to enable the creation of highly instrumented, automated, data rich biology experiments.

Acknowledgments

I am grateful for the support of a Howard Hughes Medical Institute Summer Medical Research Fellowship, which was the catalyst for all of the work described herein. I am also grateful to Jack Baskin and Peggy Downes Baskin and to the Baskin School or Engineering for their support via additional fellowship funding. Work on sequence analysis algorithms and mathematics was also supported by the National Human Genome Research Institute of the National Institutes of Health under Award Number 5U54HG007990, the National Heart, Lung, and Blood Institute of the National Institutes of Health under Award Number 1U01HL137183-01, and grants from the W.M. Keck foundation and the Simons Foundation. Work on devices for cell culture and data acquisition was supported by the Schmidt Futures Foundation under Award Number SF 857, National Human Genome Research Institute under Award Number RM1 HG011543 and the Howard Hughes Medical Institute Investigatorship award of David Haussler.

I would like to thank my mentors, David Haussler, Benedict Paten, and unofficially, Mircea Teodorescu, for their personal, scientific and engineering guidance throughout my projects. They have made the PhD experience extremely fruitful in terms of personal growth and scientific opportunity. I would also like to thank Sofie Salama, who has been my key additional mentor with regards to biological experiments and experimental design.

I would like to thank Jordan Eizenga, David Haussler and Benedict Paten for their helpful discussions throughout the development of all mathematical theory for sequence analysis. Adam Novak and Wolfgang Beyer have also been extremely helpful during my time with

the computational genomics laboratory.

Kateryna Voitiuk, Mason Hargrave, Pattawong Pansodtee, Gary Mantalas, Rob Currie among others were key figures in developing the early ideas and prototypes of what I have come to describe as high throughput biology. The ideas of high throughput biology owe a large debt to the vision of David Haussler. Pierre Baudin, Victoria Ly and Pattawong Pansodtee have also been key contributors to work on automated parallel microscopy which is beyond the scope of this project. Their help and collaboration has also been a great help for the work reported in this dissertation.

I appreciate the direct assistance of the many undergraduate students, graduate students and staff who I have mentored throughout my doctoral studies. Some of these, like Christopher Liu, Kivilcim Doganyigit, Drew Ehrlich and Valeska Victoria have made direct contributions to the projects described in this dissertation and are credited as such. Other students, including Sruthi Shriram, Khushi Shah, Julian Lehrer and Deepro Pasha have been great contributors who have made major contributions to ongoing projects which are not reported in this dissertation.

The scope and pace of the prototyping work which underlies the device-building projects has been a massive logistical undertaking. Kristof Tigyi, Liam Tran, Yvonne Vasquez and Mandeep Singh have been huge helps in making these projects logistically possible. This work also would not have been possible without the support of Mircea Teodorescu's laboratory space and other resources. The work on incubator-independent biology, the beginnings of which are reported herein, has also been a tremendous undertaking which has essentially involved building out a new laboratory from scratch as a graduate student, down to the details of

stocking, waste management, electrical outlets etc. This would not have been possible without the partnership of Mohammed Mostajo-Radji (who has also been a great source of support by means of biological samples and reagents) nor the advice, space and financial support of Sofie Salama and Mircea Teodorescu's laboratories.

The text of this dissertation includes reprints of the following previously published material: 1) *Yohei Rosen, Jordan Eizenga, and Benedict Paten. Modelling haplotypes with respect to reference cohort variation graphs. Bioinformatics, 33(14):i118–i123, 2017*, 2) *Yohei Rosen and Benedict Paten. An average-case sublinear forward algorithm for the haploid Li and Stephens model. Algorithms for Molecular Biology. 2019 Dec; 14(1):1–2* and 3) *Rosen Y, Eizenga J, Paten B. Describing the local structure of sequence graphs. In International Conference on Algorithms for Computational Biology 2017 Jun 5 (pp. 24–46). Springer, Cham*. I served as primary author on all of these publications and I am responsible for the ideas, algorithms, implementation, analysis and writing which they contain. The co-authors listed in these publications, with the exception of Jordan Eizenga, directed and supervised the research which forms the basis for the dissertation. A permission letter from Jordan Eizenga for including his coauthored work is attached.

I would like to thank my family, my friends and especially Nghi Lam for their support during my research here at UC Santa Cruz.

Part I

Graph theory of sequence graphs

Chapter 1

Describing the local structure of sequence graphs

Preamble

This chapter is adapted from my paper "Describing the local structure of sequence graphs" for which I am first author [107]. The paper describes extensions to Benedict Paten et al.'s work on *ultrabubbles* and *snarls* [98], graph features which allow the topological decomposition of sequence graphs. Specifically, I extend the theory of snarls to provide further topological granularity in the case of abutting elements of variation.

I contributed the mathematical concepts, algorithms, proofs and writing which follows. These methods were implemented by Christopher Liu while he was an undergraduate student under my direction. His implementation may be found at https://github.com/chrisliu/vgdecomp Description and characterization of this implementation are outside the scope of this thesis.

1.1 Abstract

Analysis of genetic variation using graph structures is an emerging paradigm of genomics. However, defining genetic sites on general sequence graphs remains an open problem. The invention of the *ultrabubble* and *snarl*, special subgraphs of sequence graphs which can identified with efficient algorithms, represents important first step to segregating graphs into genetic sites. We extend the theory of ultrabubbles to a special subclass where every detail of the ultrabubble can be described in a series and parallel arrangement of genetic sites. We furthermore introduce the concept of *bundle* structures, which allows us to recognize the graph motifs created by additional combinations of variation in the graph, including but not limited to runs of abutting single nucleotide variants. We demonstrate linear-time identification of bundles in a bidirected graph.

1.2 Background

The concept of the genetic site underpins both classical genetics and modern genomics. From a biological perspective, a site is a position at which mutations have occurred in different samples' histories, leading to genetic variation. From an engineering perspective, and using a mathematical graph formalism to describe genetic variations, a site is a subgraph with left and right endpoints where traversals by paths correspond to alleles. This is useful for indexing and querying variants in paths and for describing variants in a consistent and granular manner.

Against a linear reference, it is trivial to define sites, provided that we disallow vari-

ants spanning overlapping positions. This is clearly demonstrated by VCF structure [24]. VCF sites, consisting of any number of possible alleles, are identified by their endpoints with respect to the linear reference.

If we wish to analyze a set of variants containing structural variation, highly divergent sequences or nonlinear references structures, then a linear reference with only non-overlapping variants is no longer a sufficient model. Datasets with one or more of these properties are becoming more common [23, 120], and sequence graphs [90] have been developed as a method of representing them. However, defining sites on graphs is considerably more difficult than on linear reference structures and the creation of methods to fully decompose sequence graphs into sites remains an unsolved problem.

1.3 The challenges of defining sites on graphs

The linear reference definition of a site as a position along the reference and a set of alleles fails to work for several reasons:

1. Sequences which are, in linear position, at the same location may not have comparable contexts. This is a consequence of having variants which cannot be represented as edits to the linear reference but rather as edits to another variant. We illustrate this with an example from *1000 Genomes* polymorphism data, visualized using Sequence Tube Maps [9]

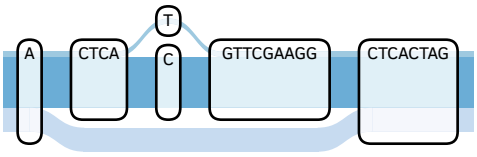


Figure 1.1: The context of the single nucleotide variant shown does not exist in all variants spanning its position.

2. Elements of variation may not be linearly ordered. Parallel structure of the graph is one sort of non-linearity shown in 3. Graphs also accomodate representation of repetitive sequence as well as inverted or transposed elements of sequence, both of which disrupt the linear ordering of the graph.

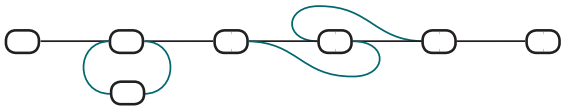


Figure 1.2: A cycle and an inversion in a graph

3. The positions spanned by different elements of variation may overlap. Therefore, multiple mutually exclusive segments of sequence in a region of the graph cannot be considered to be alternates to each other at a well-defined position without having to include extraneous sequence that is shared between some but not all of the "alleles."

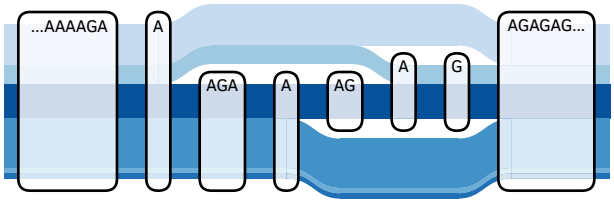


Figure 1.3: Overlapping deletions, from 1000 Genomes polymorphism data [23, 120]

1.4 Mathematical background

1.4.1 Directed and bidirected sequence graphs

In general, the graphs used to represent genetic information are mathematical structures consisting of labelled nodes and edges. Node labels correspond to sequence fragments. Edges between these labelled nodes form paths whose labels spell out allowed sequences. These graphs are of two principal types.

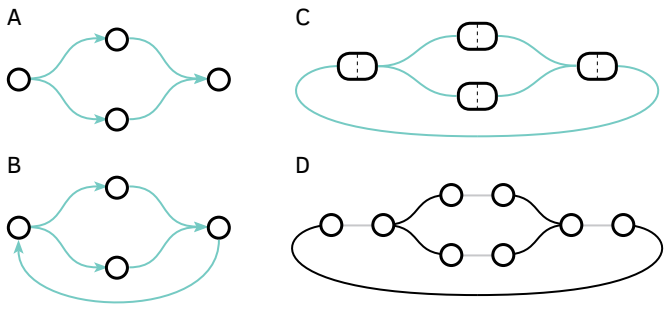


Figure 1.4: **Examples of directed graph types** (A) A directed acyclic graph (B) A cyclic directed graph (C) Graph B represented as a bidirected graph. This cycle is proper. (D) Graph C represented as a biedged graph.

The most simple type is the directed graph. A directed graph (or "digraph") G consists of a set V of nodes and a set E of directed edges. A directed edge is an ordered tuple (x,y), consisting of a *head* $x \in V$ and *tail* $y \in V$. A directed path is a sequence of nodes joined by edges, followed head to tail, in the graph. G is a *directed acyclic graph* (DAG) if it admits no directed path which revisits any node.

Bidirected graphs [82] are conceptually similar to both directed and undirected graphs. A bidirected graph G consists of a set V of vertices and a set E of edges. Each vertex $v \in V$ consists of a pair of *node-sides* $\{v_{left}, v_{right}\}$ and each edge is an unordered tuple of node-sides.

We write N for the set of node-sides in G. The opposite *opposite* $\hat{n}$ of a node-side n is defined to be the other node-side at the same vertex as n.

A sequence $p = x_1, x_2, \ldots, x_k$ of node-sides is a *path* if $\forall x_i$,

1. if $x_{i-1} \neq \hat{x_i}$, then $x_{i+1} = \hat{x_i}$
2. if $x_{i-1} \neq \hat{x_i}$, then $\{x_{i-1}, x_i\} \in E$
3. any contiguous subsequence of p consisting of a node-side x alternating with its opposite $\hat{x}$ must either be even-numbered in length or must be a prefix or suffix of p

A bidirected graph G is *cyclic* if it admits a path visiting a node-side twice. Therefore the self-incident hairpin motif (below, right) is considered a cycle. A bidirected graph G is

properly cyclic if it admits a path which visits a pair $\{n, \hat{n}\}$ twice in the same order.

Figure 1.5: (Left) A properly cyclic graph. (Right) The self-incident hairpin motif of a cyclic but not properly cyclic graph.

Some publications refer to biedged graphs. These are $\{black, grey\}$-edge-coloured undirected graphs, where every node is paired with precisely one other by sharing a grey edge and paths in the graph must alternate between traversing black and grey edges. Paten elaborates on this construction and shows that it is equivalent to a bidirected graph in [98]. We will restrict our language to that of bidirected graphs, recognizing that these are equivalent to biedged graphs.

Acyclic bidirected graphs are structurally equivalent to directed graphs in that

Lemma 1. If G is a bidirected acyclic graph, there exists an isomorphic directed acyclic graph $D(G)$.

Proof. Refer to [98] □

1.4.2 Bubbles, superbubbles, ultrabubbles and snarls

The first use of local graph structure to identify variation was the detection of *bubbles* described by [author] [ref] in order to detect and remove sequencing errors from assembly

graphs. Their bubble is the graph motif consisting of two paths which share a source and a sink but are disjoint between.

In a graph of genetic variation, these bubble motifs correspond to positions in the graph where there are precisely two variants spanning the same positions, with no other variants overlapping them.

The general concept of bubbles was extended by Onodera et al, who defined superbubbles in directed graphs [94]. Brankovic demonstrates an $O(|V|+|E|)$ algorithm to identify them [13], building off work of Sung [122].

We restate the Onodera definition, modified slightly as to be subgraph-centric rather than boundary-centric: A subgraph $S \subseteq G$ of a directed graph is a *superbubble with boundaries* (s,t) if

1. (reachability) t is reachable from s by a directed path in S
2. (matching) the set of vertices reachable from s without passing through t is equal to the set of vertices from which t is reachable without passing through s, and both are equal to S
3. (acyclicity) S is acyclic
4. (minimality) there exists no $t' \in S$ such that boundaries (s,t') fulfil 1,2 and 3. There exists no $s' \in S$ such that (s',t) fulfil 1,2 and 3.

In order to expand the concept of superbubble to bidirected graphs, we prove some consequences of the matching property.

Proposition 1. Let $S \subseteq G$ be a subgraph of a directed graph. If IT posesses the matching property relative to a pair (s,t), then it possesses the following three properties:

1. (2-node separability) Deletion of all incoming edges of s and all outgoing edges of t disconnects S from the remainder of the graph.

2. (tiplessness) There exist no node $n \in S$ other than possibly a boundary, such that n has either only incoming or outgoing edges.

3. S is weakly connected

Proof. $(matching \Rightarrow separability)$ Suppose $\exists x \notin S, y \in S \backslash \{s,t\}$ such that there exists either an edge $x \to y$ or an edge $y \to x$. Suppose wlog that $\exists$ an edge $x \to y$. By matching, there exists a path $y \to \cdots \to t$ without passing through s. We can then construct the path $x \to y \to \cdots \to t$ which does not pass through s. But by matching this implies that $x \in S$, which leads to a contradiction. □

The converse need not be true for directed graphs, but it is on bidirected graphs. An ultrabubble may be seen as an analogue to a superbubble, and a snarl as a more general object which preserves the property of separation from the larger graph without having strong guarantees on its internal structure. The following definitions are due to Paten [98]:

A connected subgraph $S \subseteq G$ of a bidirected graph G is a *snarl* (S,s,t) with boundaries (s,t), $s,t \in N$ if

1. $s \neq \hat{t}$

2. (2-node separability) every path between a pair of node-sides in $x \in S$, $y \in X^C$ contains either $s \to \hat{s}$ or $t \to \hat{t}$ as a subpath.

3. (minimality) there exists no $t' \in S$ such that boundaries (s,t') fulfil 1 and 2. There exists no $s' \in S$ such that (s',t) fulfil 1 and 2

We furthermore say that a snarl is an *ultrabubble* if

4. S is acyclic

5. S contains no tips

Three examples of ultrabubbles are shown below.

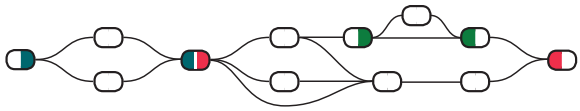


Figure 1.6: **Three ultrabubbles, boundaries coloured blue, pink and green** These illustrate the non-overlapping property

The following is important property of snarls.

Proposition 2 (Non-overlapping property)**.** If two distinct snarls share a vertex (node-side pair) then either they share a boundary node or one snarl is included in the other's interior.

Proof. Let S be a snarl with boundaries s,t. Let T be another snarl, with boundaries u,v. Suppose that $u \in S \backslash \{s,t\}$ but $v \notin S$, and $s \notin T$.

Consider the set $S \cap T$. It is nonempty since it contains u. Let $x \in S \cap T$. Let $y \notin S \cap T$. Suppose that there exists a path $p = x \leftrightarrow \cdots \leftrightarrow y$ which neither passes through u nor t.

Since $y \notin S \cap T$, either $y \notin S$ or $y \notin T$. Wlog, assume $y \notin T$. Then due to the separability of T, since the path p does not pass through u, it must pass through v before leaving T to visit y. But $v \notin S$ so p must also pass through s before leaving S to visit v since it does not pass through t. But it must pass through v before leaving T to visit s, which leads to an impossible sequence of events. Therefore any path $x \leftrightarrow \cdots \leftrightarrow y$ for $x \in S \cap T, y \notin S \cap T$ must pass through either u or t. This contradicts the minimality of both S and T. □

The equivalent property for superbubbles was stated without proof by Onodera in [94]. Our proof also constitutes a proof of the statement for superbubbles, due to the following proposition, proven by Paten in [98]:

Proposition 3. Every superbubble in a directed graph corresponds to an ultrabubble in the equivalent (see Lemma 1) bidirected graph.

Identifying all superbubbles in a directed graph or all snarls in a bidirected graph introduces a method of compartmentalizing a graph into partitions whose contents are all in some sense at the same position in the graph, and for which the possible internal paths are independent of what path they continue on beyond their boundaries. We will use this concept to define sites for certain specialized classes of graphs.

1.5 New results: decomposing graphs by ultrabubbles

1.5.1 Simple bubbles and nested simple bubbles

An (s,t)-*traversal* of S is a path in S beginning with s and ending with t. An (s,s)-*traversal* and a (t,t)-*traversal* are analogously defined. Presence of an (s,s)- or (t,t)-traversal implies cyclicity. Two traversals of a snarl are *disjoint* if they are disjoint on $S\backslash\{s,t\}$.

Paten's [98] snarls and ultrabubbles are 2-node separable subgraphs whose paired boundary nodes isolate their traversals from the larger graph. We can state this with more mathematical rigor:

Claim 1. Consider a snarl (S,s,t) in a bidirected graph G. The set of all paths in G which contain a single (s,t)-traversal as contiguous a subpath is isomorphic to the set-theoretic product $P(s) \times Trav(s,t) \times P(t)$ consisting of the three sets

1. $P(s) := \{\text{paths in } G\backslash S \text{ terminating in } \hat{s}\}$
2. $Trav(s,t) := \{(s,t)\text{-traversals of } S\}$
3. $P(t) := \{\text{paths in } G\backslash S \text{ beginning with } \hat{t}\}$

with the isomorphism being the function mapping $p_1 \in P(s), p_2 \in Trav(s,t), p_3 \in P(t)$ to their concatenation $p_1p_2p_3$.

Definition 1. An ultrabubble (S,s,t) is a simple bubble if all (s,t)-traversals are disjoint.

Simple bubbles are structurally equivalent to (multiallelic) sites consisting of disjoint substitutions, insertions or deletions, with all alleles spanning the same boundaries.

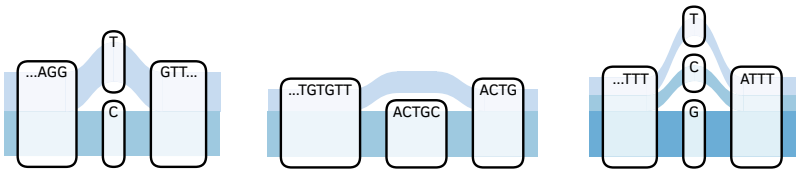


Figure 1.7: **Three examples of simple bubbles from the 1000 Genomes graph**

Proposition 4 below demonstrates that we can identify simple bubbles in $O(|V|)$ time given that we have found all snarl boundaries. Paten has shown [98] that identification of snarl boundaries is achieved in $O(|E|+|V|)$ time.

Lemma 2. All nodes in an ultrabubble are orientable with respect to the ultrabubble boundaries.

Proof. This is a corollary to Lemma 1. □

Therefore, we can refer to each vertex in an ultrabubble as having an *s-side* and a *t-side*.

Proposition 4 (Simple bubbles have unbranching interiors)**.** Let (S,s,t) be an ultrabubble. Then all traversals are disjoint iff every interior node-side has precisely one neighbor.

Proof. $(\Rightarrow)$ Suppose that all such traversals are disjoint. Suppose the existence of an interior node-side n with multiple neighbors.

Since n is orientable with respect to (s,t), suppose, without loss of generality, that it is s-sided. Then there exist distinct paths from s to n passing through each of its neighbors. Continuing these with a path from n^{opp} to t produces two nondisjoint traversals of S.

$(\Leftarrow)$ Suppose that every interior node-side has precisely one neighbor. Suppose that there exist two distinct nondisjoint traversals of S. For no node-side to have multiple neighbors,

they must coincide at every node-side, contradicting the assumption that they are not the same traversal. □

We seek to extend this simple property to more complex graph structures. We will take advantage of the nesting of nondisjoint ultrabubbles property from Proposition 2 to define another structure on which, despite being nondisjoint, all traversals are easily indexed.

Definition 2. An ultrabubble $(S,s,t) \subseteq G$ is *decomposable into nested simple sites* if either:

1. S is a simple bubble
2. if, for every ultrabubble S' contained in the interior of S, you replace the ultrabubble with a single edge $s-t$ whenever S' is decomposable into simple sites, then S becomes a simple bubble

The following figure demonstrates decomposability into nested simple sites.

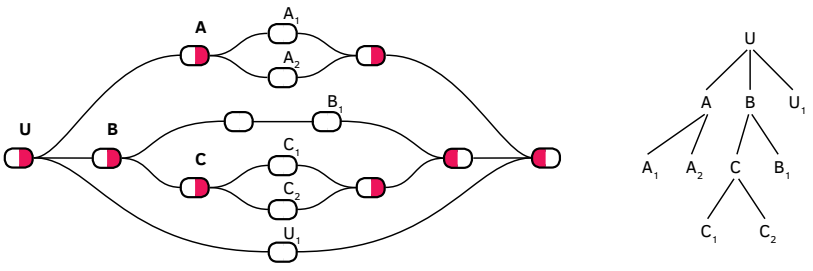


Figure 1.8: Left: A nesting of ultrabubbles. Right: The tree structure to index traversals implied by Proposition 5.

Proposition 5. If an ultrabubble (U,s,t) is decomposable into nested simple sites, then the complete node sequence of any (s,t)-traversal can be determined only by specifying the path

it takes inside those nested ultrabubbles within which the traversal does not visit any further nested ultrabubble.

Proof. Let p be a (s,t)-traversal of an ultrabubble U which is decomposable into nested simple sites. Let V be a nested ultrabubble inside U. If p traverses, V, write $p|_V$ for the traversal p restricted to V.

Suppose that $t|_V$ intersects no nested ultrabubbles within V. Then $t|_V$ is disjoint of all other traversals within V due to U begin decomposable into nested simple sites. Therefore specifying any node of $t|_V$ uniquely identifies it.

Suppose that $t|_V$ intersects some set of ultrabubbles nested within V. Since U is decomposable into nested simple sites, the nodes of $t|_V$ must be linear and disjoint of all other paths if we replace all ultrabubbles nested in V with edges joining their boundaries. Therefore specifying which ultrabubbles are crossed uniquely determines the paths taken outside of the nested ultrabubbles in V.

The statement of the proposition follows from the two arguments above by induction.

□

Proposition 6. An ultrabubble is decomposable into nested simple sites iff every node side is either the interior ultrabubble boundary or has precisely one neighbor.

Proof. This is a corollary to Proposition 2. □

This property allows $O(|V|+|E|)$ evaluation of whether a graph is decomposable into nested simple sites, since it takes one such linear time pass to identify the snarl boundaries, after which it remains to count the neighbors of all remaining node sides.

1.5.2 A partial taxonomy of graph motifs which do not admit decomposition into sites

In section 4.3, we will show that we can detect whether we can decompose a graph into nested simple sites as defined in the previous section if it lacks a certain forbidden motif. We will begin with examples of three graph motifs, and the biological events which might produce them.

We describe some of these features which prevent decomposition into nested sites below.

1. Two (or more) substitutions or deletions against a linear sequence which overlap, but not completely.

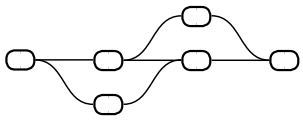


Figure 1.9: **Overlapping substitutions**

2. A substitution (or deletion) which spans elements of sequence on the interior of two disjoint ultrabubbles. Addition of such an edge joining two ultrabubbles which were decomposable into nested simple sites will consolidate the two into a single ultrabubble which is not decomposable into nested simple sites.

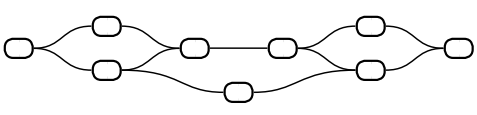


Figure 1.10: **An edge crossing bubble boundaries.**

3. Two SNVs or other simple elements of variation at adjacent positions. This will be the focus of our Section 5.

1.5.3 The Relationship Between Nested Simple Sites and Series Parallel Graphs

The structure of ultrabubbles decomposable into nested simple sites, and their tree representation (see Fig 7) might be familiar to the graph theorist familiar with series-parallel digraphs. The fact that the digraphs equivalent to ultrabubbles form a subclass of the two-terminal series-parallel digraphs is interesting due to the computational properties of the latter class of graphs.

Definition 3. A directed graph G is two-terminal series parallel (TTSP) with source s and sink t if either

1. G is the two-element graph with a single directed edge $s \to t$

2. There exist TTSP graphs G_1, G_2 with sources s_1, s_2 and sinks t_1, t_2 such that G is formed from G_1, G_2 by identification of s_1 with s_2 as s and identification of t_1 with t_2 as t (Parallel addition)

3. There exist TTSP graphs G_1, G_2 with sources s_1, s_2 and sinks t_1, t_2 such that G is formed from G_1, G_2 by identification of t_1 with s_2 (Series addition)

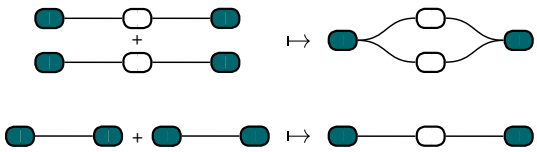


Figure 1.11: Top: parallel addition. Bottom: series addition.

Two terminal series parallel digraphs have a useful forbidden subgraph characterization.

Proposition 7 (due to [131])**.** A directed graph G is two terminal series parallel if and only if it contains no subgraph homeomorphic to the graph W shown below

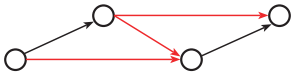


Figure 1.12: **The W motif** An illustration of the forbidden subgraph in question.

Proof: Refer to Valdes [131] and Duffin [27]

Proposition 8. If an ultrabubble (U,s,t) is decomposable into nested simple sites, then the equivalent directed graph is TTSP with source s and sink t.

Proof. Suppose that the directed graph $D(U)$ equivalent to U (which exists by Lemma 1) contains a subgraph homeomorphic to W. Then there must be a node-side u in U with two neighbours a_1, a_2 which are the beginnings of disjoint paths p_1, p_2 ending on node-sides b_1, b_2 which are neighbours of a node-side v. By Proposition 6, u and v must be ultrabubble boundaries. Since p_1, p_2 are disjoint, u and v must be opposing boundaries of the same ultrabubble. But

the presence of a subgraph homeomorphic to W also implies that there exists a pair q_1, q_2 of disjoint paths, one from a node x to $\hat{u}$ and the other from x to v, both not passing through u or $\hat{v}$. But this is not possible since it would contradict 2-node separability of (u, v). □

We highlight the middle "Z-arm" of the W-motif in our first two examples of ultrabubbles which are not decomposable into nested simple sites.

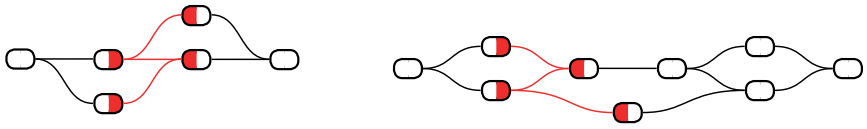


Figure 1.13: **Portions of the ultrabubbles 1. and 2. of the previous section, showing the nodes which project to the forbidden subgraph W.**

1.6 Abutting variants

We wish to decompose the graph structure of adjacent variants whose boundaries abut such that they are not segregated between them by ultrabubble boundaries. We will describe a graph motif called the *balanced recombination bundle* which describes this graph structure, and can be rapidly detected in the graph.

1.6.1 Snarl boundaries and bundles

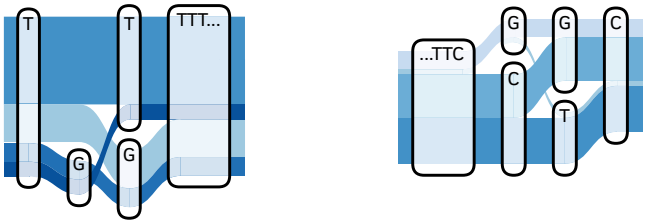


Figure 1.14: Two examples of abutting SNVs in the 1000 Genomes graph

Given a node-side n, write $Nb(n)$ for the set of all neighbors of n. Note that $a \in Nb(b) \Leftrightarrow b \in Nb(a)$.

Definition 4. An *internal chain* $n_1 \to n_2 \to \cdots \to n_k$ is a sequence of node-sides such that $\forall i, 2 \leq i \leq k, n_i \in Nb(n_{i-1})$.

Definition 5. We say that a tuple (L,R) of sets of node-sides is a *bundle* if

1. (Matching) $\forall \ell \in L, Nb(\ell) \subseteq R$ and $Nb(\ell) \neq \varnothing$; $\forall r \in R, Nb(r) \subseteq L$ and $Nb(r) \neq \varnothing$
2. (Connectedness) $\forall \ell \in L, r \in R$, there exists an internal chain $\ell \to r_1 \to \ell_1 \to \cdots \to r_k \to \ell_k \to r$ such that $\forall i, 1 \leq i \leq k, r_i \in R$ and $\ell_i \in L$

Definition 6. We say that a tuple (L,R) of sets of node-sides is a *balanced recombination bundle* (R-bundle for short) if

1. (Complete matching) $\forall \ell \in L, Nb(\ell) = R$ and $\forall r \in R, Nb(r) = L$
2. (Acyclicity) $L \cap R = \varnothing$

Lemma 3. A balanced recombination bundle is a bundle.

Proof. Complete matching $\Rightarrow$ matching.

Complete matching $\Rightarrow$ connectedness by the chain $\ell \to r$ for all $\ell \in L, r \in R$ $\square$

Definition 7. An *unbalanced bundle* is a bundle which is not a balanced recombination bundle. An unbalanced bundle is acyclic if $L \cap R = \varnothing$.

Definition 8. We say that two bundles $(L_1, R_1), (L_2, R_2)$ are *isomorphic* if either $L_1 = L_2$ and $R_1 = R_2$ or $L_1 = R_2$ and $R_1 = L_2$.

We will describe a $O(|V| + |E|)$ algorithm to detect and categorize bundles exhaustively for all node-sides in a bidirected graph. To establish the validity of this algorithm, we need several preliminary results:

Lemma 4. Every $q \in N$ is either a tip or an element of a bundle.

Proof. Suppose that q is not a tip. Define a function W that maps a tuple (L, R) of nonempty sets of node-sides to a tuple $W(L), W(R)$ where

$$W(R) := R \cup \bigcup_{\ell \in L} Nb(\ell)$$

$$W(L) := L \cup \bigcup_{r \in W(R)} Nb(r)$$

$\forall n \in \mathbb{N}$ define

$$W^n((L,R)) := \underbrace{W \circ \cdots \circ W((L,R))}_{n \text{ times}}$$

$$W^\infty((L,R)) := W^k((L,R)) \text{ for } k \text{ such that}$$

$$W^{k+i}((L,R)) = W^k((L,R)) \forall i \in \mathbb{N}$$

W^∞ exists since W^n is nondecreasing with respect to set inclusion and our graphs are finite. Now define $\overline{W}(q) := W^\infty((\{q\}, Nb(q)))$, noting that $Nb(q) \neq \varnothing$ since $\{q\}$ is not a tip. Let us write L_{W^∞} and R_{W^∞} for the respective elements of $\overline{W}(q)$. We claim that $\overline{W}(q)$ is a bundle.

Proof of matching: let $\ell \in L_{W^\infty}, r \in R_{W^\infty}$. By construction of W,

$$Nb(\ell) \subseteq W(R_{W^\infty}) = R_{W^\infty}$$

$$Nb(r) \subseteq W(L_{W^\infty}) = L_{W^\infty}$$

Proof of connectedness: let $\ell \in L_{W^\infty}, r \in R_{W^\infty}$. We will show that for any $r \in R_{W^\infty}$, $\exists$ an internal chain $q \to r_1 \to \ell_1 \to \cdots \to r_k \to \ell_k \to r$ such that $\forall i, 1 \leq i \leq k, r_i \in R_{W^\infty}$ and $\ell_i \in L_{W^\infty}$.

Suppose that $r \in Nb(q)$, then we are done. Otherwise, since $r \in R_{W^\infty}$, there exists some minimal $n \in \mathbb{N}$ such that $r \in$ the R-set R_{W^n} of some $W^n((\{q\}, Nb(q)))$. It is straightforward to see that we can then construct an internal chain $q \to r_0 \to \ell_1 \to r_1 \to \ldots \ell_{n-1} \to r$ such that $\forall i, 1 \leq i \leq n-1, r_i \in R_{W^i}, \ell_i \in L_{W^i}$. By an analogous argument, we can do the same for an internal chain $\ell \to \cdots \to r'$ for some $r' \in Nb(q)$. Concatenation of the first chain with the

reverse of the second gives our chain $\ell \to \cdots \to r$, proving connectedness. □

Proposition 9. If $q \in L$ for a bundle (L,R), then $(L,R) = \overline{W}(q)$

Proof. Suppose that $\overline{W}(q) \neq (L,R)$. Then either $L \neq L_{W^\infty}$ or $R \neq R_{W^\infty}$. First, suppose the latter. Suppose that $\exists r \in R$ such that $r \notin R_{W^\infty}$. Since (L,R) is a bundle, we know that there is an internal chain $q \to r_0 \to \ell_1 \to r_1 \to \cdots \to r_k \to \ell_k \to r$ with all $r_i \in R$, $\ell_i \in L$. But, using the same shorthand as before, it is also evident that $r_i \in R_{W^i}, \ell_i \in L_{W^i} \forall i, 1 \leq i \leq k$. But since $\ell_k \in Nb(r)$, we can deduce that $r \in R_{W^{k+1}}$, which leads to a contradiction since $r \notin R_{W^\infty}$.

Suppose otherwise that $\exists r \in R_{W^\infty}$ such that $r \notin R$. Consider an internal chain $c = q \to r_0 \to \ell_1 \to r_1 \to \cdots \to r_k \to \ell_k \to r$ fulfilling the conditions needed to prove connectedness of $\overline{W}(q)$. Note that $q \in L$ and by matching $r_0 \in Nb(q)$. But $r \notin R$, which leads to a contradiction since it means that two consecutive members somewhere in the chain c cannot be neighbors. □

We say that a node-side n is *involved in* a bundle (L,R) if $n \in L$ or $n \in R$.

Corollary 1. Every non-tip node-side is involved in precisely one bundle.

1.6.2 An algorithm for bundle-finding

The following two diagrams demonstrate our algorithm for finding the balanced recombination bundle containing a query node-side q if it is contained in one, and discovering that it is not if it is not. This is written in pseudocode below, with an illustration following.

In order to prove that this is a valid algorithm for detection of balanced recombination bundles, we need the following lemma.

Algorithm 1: BalancedRecombinationBundleFinder

Data: Node-side q

Result: Bundle containing q if it is in a balanced recombination bundle, $\varnothing$ if q is in an unbalanced bundle or is a tip

begin

- **if** $Nb(q) = \varnothing$ **then**
  - return $\varnothing$
- $A \longleftarrow Nb(q)$
- $B \longleftarrow Nb(R[0])$
- **if** $A \cap B \neq \varnothing$ **then**
  - return $\varnothing$
- **else**
  - **for** $a \in A \backslash \{R[0]\}$ **do**
    - **if** $Nb(a) \neq B$ **then**
      - return $\varnothing$
  - **for** $b \in B \backslash \{q\}$ **do**
    - **if** $Nb(b) \neq A$ **then**
      - return $\varnothing$
- return $MakeBundle(A, B)$

Lemma 5. Let (L,R) be a tuple of sets of node-sides. If $\exists q \in L$ such that $\forall a \in Nb(q), \forall b \in Nb(a), Nb(b) \subseteq Nb(q)$ but $Nb(q) \subset R$, then (L,R) cannot be connected.

Proof. Let $B = \bigcup_{a \in Nb(q)} Nb(a)$. We know that $\forall b \in B, Nb(b) \subseteq Nb(q)$. Suppose that (L,R) is connected. Choose $r \in R \backslash Nb(q)$. Then $\exists$ an internal chain $c = q \to r_1 \to \ell_1 \to \cdots \to r_k \to \ell_k \to r$ with $r_i \in R, \ell_i \in L \forall i$. Since $q \in B$, $Nb(b) \subseteq Nb(q) \forall b \in B$, and $Nb(a) \subseteq B \forall a \in Nb(q)$, it is impossible that the sequence of node-sides c is both a valid internal chain and ends with r. Therefore (L,R) cannot be connected. □

Proposition 10 (Validity of Algorithm 1)**.** This algorithm detects all balanced recombination bundles, and rejects all unbalanced recombination bundles.

Proof. Suppose q is involved in a balanced recombination bundle (L,R). W.l.o.g. suppose that $q \in L$. Due to the complete matching property, the set $Nb(q)$ in the algorithm is guaranteed to be equal to R. Due to the completeness property, the set $Nb(R[0])$ in the algorithm is guaranteed to be equal to L. It is evident that the algorithm directly verifies complete matching and acyclicity.

Suppose otherwise. Assuming we have eliminated all tips, which can be done in $|\mathcal{V}|$ time, Lemma 4 proves that q is involved in an unbalanced bundle B. If B fails acyclicity but not complete matching, then checking that $A \cap B = \varnothing$ will correctly detect that $L \cap R \neq \varnothing$.

Otherwise, suppose that B fails complete matching. Suppose first that $Nb(q) \subset R$. We assert that $\exists a \in Nb(q)$ such that $\exists b \in Nb(a)$ such that $\exists c \in Nb(b)$ such that $c \notin Nb(q)$. This event will be detected by the second loop of the algorithm. This follows from the connectedness of B and Lemma 6.

Suppose otherwise that $Nb(q) = R$ but $\exists r \in R$ such that $Nb(r) \subset L$. Let $c \in L \backslash Nb(r)$.

By matching, $\exists r' \in R$ such that $r' \in Nb(c)$. Therefore $Nb(r)$ and $Nb(r')$ will be found to be unequal in the first loop of the algorithm.

Suppose otherwise that $Nb(q) = R, Nb(r) = L \forall r \in R$, but $\exists \ell \in L$ such that $Nb(\ell) \subset R$. Then we will find in the second loop that $Nb(\ell) \neq Nb(q)$. □

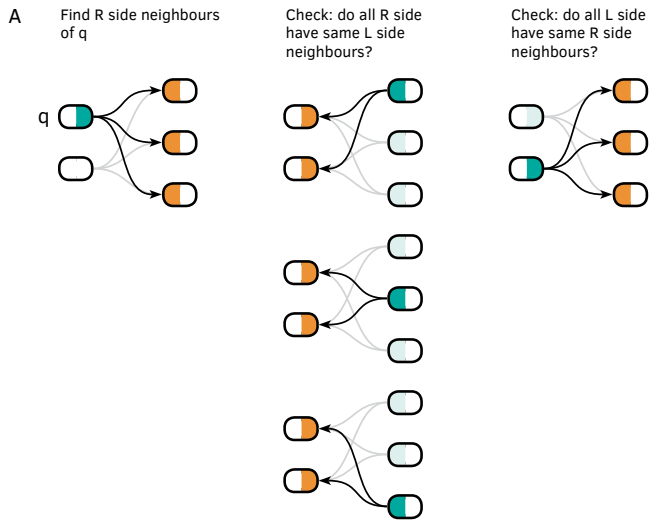


Figure 1.15: **Illustration of Algorithm 1 returning a positive result**

Proposition 11 (Speed of Algorithm 1)**.** We can identify all balanced recombination bundles, all unbalanced bundles and all tips in $O(|E| + |V|)$ time.

Proof. We depend on a neighbor index giving us $O(|Nb(n)|)$ iteration across neighbors of a node-side n.

We begin by looping over all node-sides and identifying all tips, which is achieved in $O(|V|)$ time. We then loop again over all remaining node-sides. At each node-side q, we run the function describe above, which, if q is involved in a balanced recombination bundle, will return

the bundle $B = \overline{W}(q)$. It is evident that this function runs in $O(|E_B|)$ time, seeing as it loops over each edge of B twice—once from each side—each time making an $O(1)$ set inclusion query. After B is built, all nodes are marked such that they are skipped when they are encountered in the global loop. This gives overall $O(|E_B| + |V_B|)$ exploration of B.

If q is involved in an unbalanced bundle $B = \overline{W}(q)$, this fact is detected by the same function in $O(|E_B|)$ time. In this case, we can find all nodes of B by performing a breadth-first search. Examination of the W-function will convince the reader that Lemma 5 proves that a breadth-first search will find all node-sides of B in $O(|E_B| + |V_B|)$ time. We follow the same procedure of marking all these node-sides to be skipped in the global loop.

This proves that, after eliminating tips in $O(|V|)$ time, we can build the set $\mathbb{B}$ of all non-isomorphic bundles B, and decide whether they are balanced recombination bundles, in time proportional to $\sum_{B\in\mathbb{B}} |E_B| + |V_B|$. But Corollary 2 tells us that $V = \{v : v \text{ is a tip}\} \cup \bigcup_{B\in\mathbb{B}} V_B$, and that all components of this union are disjoint. Furthermore, due to the matching property of bundles, $E = \bigcup_{B\in\mathbb{B}} E_B$, and all elements of this union are disjoint. Therefore, our method is $O(|V| + |E|)$. □

1.6.3 Bundles and snarl boundaries

Definition 9. Given a boundary node-side $b = s$ or t of a snarl (S, s, t), we call the tuple $(b, Nb(b))$ a *snarl comb*. A snarl comb is called proper if $\forall n \in Nb(b), Nb(n) = \{b\}$ and $b \notin Nb(b)$.

It is easy to verify that a *proper snarl comb* is a balanced recombination bundle. It is also easy to see that an improper snarl comb is, according to set inclusion of tuples, a proper

subset of a unique bundle.

Proposition 12 (Bundles do not cross snarl boundaries)**.** Let (S,s,t) be a snarl. Suppose that $B = (L,R)$ is a bundle. Then either all node-sides involved in B are members of S, or no node-side involved in B is a member of S.

Proof. Suppose that there exists a bundle $B = (L,R)$ with node-sides both within S and not within S. Let x,y be involved in B, with $x \in S$, $y \notin S$. W.l.o.g., suppose $x \in L$, $y \in R$. This implies that there exists an internal chain $p = x \to \cdots \to y$. But then this implies that there exists $a \in S, b \notin S$ such that $a \in Nb(b)$, which would allow us to use the edge $a \to b$ to create a path violating the 2-node separability of S □

1.6.4 Defining sites using bundles

Definition 10. An ordered pair (B_1,B_2) of balanced recombination bundles is *compatible* if either

1. $\forall x \in R_1, \hat{x} \in L_2$, and $\forall y \in L_2, \hat{y} \in R_1$
2. $\exists$ a bijection $f : L_1 \longrightarrow R_2$ such that $\forall x \in R_1$, there exists a unique path $p(x)$ from $x \to \cdots \to f(x)$, and all paths $p(x)$ are disjoint.

Definition 11. If two recombination bundles are compatible, we define the set of all paths $p(x)$ to be a *bundled simple site P*.

Claim 2. Consider a bundled simple site P in a graph G, lying between compatible balanced recombination bundles B_1,B_2. The set of all paths in G which contain paths $p \in P$ as contiguous subpaths is isomorphic to the set-theoretic product $P(L_1) \times P \times P(R_2)$ consisting of the three sets

1. $P(L_1) := \{\text{paths in } G \backslash S \text{ terminating in } x, \text{ for some } x \in L_1\}$

2. P

3. $P(R_2) := \{\text{paths in } G \backslash S \text{ beginning with } y, \text{ for some } y \in R_2\}$

under the function mapping $p_1 \in P(L_1), p \in P, p_2 \in P(R_2)$ to their concatenation.

We will call a balanced recombination bundle $B = (L,R)$ *trivial* if both L and R are singleton sets.

Definition 12. An ultrabubble (U,s,t) is *a generalized simple bubble* if

1. $(\{s\}, Nb(s))$ and $(Nb(t), \{t\})$ are balanced recombination bundles

2. The set of all non-trivial balanced recombination bundles admits a linear ordering $X \to B_1 \to \ldots B_k \to Y$ such that X and Y are either of $(\{s\}, Nb(s))$ and $(Nb(t), \{t\})$, X is compatible with B_1, every B_i is compatible with B_{i+1}, and B_k is compatible with Y

Note that if an ultrabubble U is generalized simple, all node-sides in U are elements of bundled simple sites.

Definition 13. An ultrabubble U is decomposable into nested generalized sites if either:

1. It is generalized simple

2. If all nested ultrabubbles which are generalized simple bubbles are replaced with a single edge spanning their boundaries, then U is a generalized simple bubble

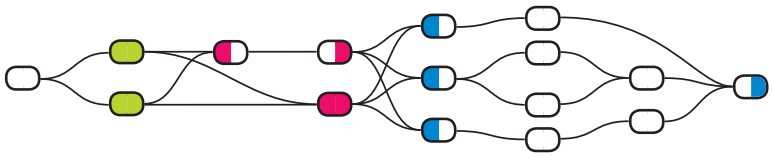


Figure 1.16: **An ultrabubble decomposable into nested generalized sites; some sites marked**

We sketch at linear-time method of building sites from a general graph. We can run this graph-wide or within individual snarls. The fact that bundles do not cross snarl boundaries allows us to use large snarls to "chunk" the work of finding sites.

Algorithm 2: Finding spans connecting bundles

Data: Ultrabubble U, and set $\mathbb{B}$ of balanced recombination bundles
Result: Set $\mathbb{P}$ of spans of unbranching sequence which potentially join compatible bundles
begin
  $\mathbb{T} \longleftarrow$ set of all trivial bundles
  $N_{\mathbb{B}} \longleftarrow$ map $(N \to \mathbb{B})$ of node-sides to bundles-sides of nontrivial R-bundles
  $N_{\mathbb{T}} \longleftarrow$ map $(N \to \mathbb{T})$ of node-sides to trivial bundles
  for *trivial bundle* $t = (\{t_l\}, \{t_r\}) \in \mathbb{T}$ **do**
    if $N_{\mathbb{T}}[\hat{t_r}]$ *is found* **then**
      $(\{u_l\}, \{u_r\}) \longleftarrow N_{\mathbb{T}}[\hat{t_r}]$
      replace $(\{t_l\}, \{t_r\})$ in $\mathbb{T}$ with $(\{t_l\}, \{u_r\})$
      flag $(\{t_l\}, \{t_r\})$ as having been right-extended
  $\mathbb{P} \longleftarrow \varnothing$
  for $t = (\{t_l\}, \{t_r\}) \in \mathbb{T}$ **do**
    if t *not flagged as having been right-extended* **then**
      $\mathbb{P} \longleftarrow \mathbb{P} \cup \{t\}$

1.7 Discussion

Graph formalism has the potential to revolutionize the discourse on genetic variations by creating a model and lexicon that more fully embraces the complexity of sequence change. This is vital: the current linear genome model of a reference sequence interval and alternates is insufficient. It fails to express nested variation and can not properly describe information about the breakpoints that comprise structural variations.

The introduction, in order, of bubbles, superbubbles, ultrabubbles and snarls progressively generalizes the concept of a genetic site to accommodate more general types of variation using progressively more general graph types. In this paper we both review and build on these developments, showing how the recently introduced ultrabubbles can be furthered sub-classified using concepts from circuit theory. This expands the simple notion of proper nesting described in the original ultrabubble paper. Furthermore, we describe how we can extend the theory of ultrabubbles by generalizing ultrabubble boundaries to another sort of boundary structure—the bundle—which allows us to describe regions where variants are packed too closely to be segregated into separate ultrabubbles.

Our methods are powerful in decomposing dense collections of nested or closely packed variation into meaningful genetic sites. We anticipate that these structures will become increasingly common in the analysis of variation using graph methods, as sequencing datasets containing variation from increasing numbers of individuals become available.

Algorithm 3: Finding compatible bundles

Data: Ultrabubble U, set $\mathbb{P}$ of node-side tuples containing endpoints of spanning segments, and set $\mathbb{B}$ of balanced recombination bundles

Result: Set $\mathbb{C}$ of all compatible pairs of bundle-sides

begin

- $N_{\mathbb{P}} \longleftarrow$ map $(N \to \mathbb{P})$ of node-sides of $p \in \mathbb{P}$ to elements of $\mathbb{P}$
- $N_{\mathbb{B}} \longleftarrow$ map $(N \to \mathbb{B})$ of node-sides to bundles-sides of nontrivial R-bundles
- $\mathbb{C} \longleftarrow \varnothing$
- **for** *R-bundle side* $X \in \mathbb{B}$ **do**
  - $x \longleftarrow X[0]$
  - R-bundle side $X_{opposite} \longleftarrow \varnothing$
  - $CompatibleNodesides \longleftarrow \varnothing$
  - **if** $\hat{x}$ *found in* $N_{\mathbb{B}}$ **then**
    - $X_{opposite} \longleftarrow N_{\mathbb{B}}[\hat{x}]$
    - $CompatibleNodesides \longleftarrow \hat{x}$
  - **else if** $\hat{x}$ *found in* $N_{\mathbb{P}}$ **then**
    - $p = (\{p_l\}, \{p_r\}) \longleftarrow N_{\mathbb{P}}[\hat{x}]$
    - $y \longleftarrow$ node-side of p which isn't $\hat{x}$
    - $X_{opposite} \longleftarrow N_{\mathbb{B}}[\hat{y}]$
    - $CompatibleNodesides \longleftarrow \hat{y}$
  - **if** $X_{opposite} \neq \varnothing$ *and* $|X_{opposite}| = |X|$ **then**
    - **for** $a \in X \backslash x$ **do**
      - **if** $\hat{x}$ *found in* $N_{\mathbb{B}}$ **then**
        - $CompatibleNodesides \longleftarrow \hat{x}$
      - **else if** $\hat{x}$ *found in* $N_{\mathbb{P}}$ **then**
        - $p = (\{p_l\}, \{p_r\}) \longleftarrow N_{\mathbb{P}}[\hat{x}]$
        - $y \longleftarrow$ node-side of p which isn't $\hat{x}$
        - $CompatibleNodesides \longleftarrow \hat{y}$
    - **if** $CompatibleNodesides = X_{opposite}$ **then**
      - $\mathbb{C} \longleftarrow \mathbb{C} \cup (X, X_{opposite})$
- return $\mathbb{C}$

Part II

Algorithms for Large Haplotype Panels

Chapter 2

An average-case sublinear forward algorithm for the haploid Li and Stephens Model

Preamble

This chapter is adapted from my paper "An average-case sublinear forward algorithm for the haploid Li and Stephens Model" for which I am first author [109]. The paper describes my theoretical work on applying sparse data structures and lazy evaluation to a sequence analysis algorithm in genomics. I contributed the mathematical concepts, algorithms, proofs and writing which follows.

2.1 Abstract

Hidden Markov models of haplotype inheritance such as the Li and Stephens model allow for computationally tractable probability calculations using the forward algorithm as long

as the representative reference panel used in the model is sufficiently small. Specifically, the monoploid Li and Stephens model and its variants are linear in reference panel size unless heuristic approximations are used. However, sequencing projects numbering in the thousands to hundreds of thousands of individuals are underway, and others numbering in the millions are anticipated.

To make the forward algorithm for the haploid Li and Stephens model computationally tractable for these datasets, we have created a numerically exact version of the algorithm with observed average case sublinear runtime with respect to reference panel size k when tested against the 1000 Genomes dataset.

We show a forward algorithm which avoids any tradeoff between runtime and model complexity. Our algorithm makes use of two general strategies which might be applicable to improving the time complexity of other future sequence analysis algorithms: sparse dynamic programming matrices and lazy evaluation.

Background

Probabilistic models of haplotypes describe how variation is shared in a population. One application of these models is to calculate the probability $P(o|H)$, defined as the probability of a haplotype o being observed, given the assumption that it is a member of a population represented by a *reference panel* of haplotypes H. This computation has been used in estimating recombination rates [66], a problem of interest in genetics and in medicine. It may also be used to detect errors in genotype calls.

Early approaches to haplotype modeling used coalescent [60] models which were accurate but computationally complex, especially when including recombination. Li and Stephens wrote the foundational computationally tractable haplotype model [66] with recombination. Under their model, the probability $P(o|H)$ can be calculated using the forward algorithm for hidden Markov models (HMMs) and posterior sampling of genotype probabilities can be achieved using the forward-backward algorithm. Generalizations of their model have been used for haplotype phasing and genotype imputation. [69, 15, 136, 25, 91]

The Li and Stephens model

Consider a *reference panel* H of k haplotypes sampled from some population. Each haplotype $h_j \in H$ is a sequence $(h_{j,1}, \ldots, h_{j,n})$ of alleles at a contiguous sequence $1, \ldots, n$ of genetic sites. Classically [66], the sites are biallelic, but the model extends to multiallelic sites. [72]

Consider an observed sequence of alleles $o = (o_1, \ldots, o_n)$ representing another haplotype. The monoploid Li and Stephens model (LS) [66] specifies a probability that o is descended from the population represented by H. LS can be written as a hidden Markov model wherein the haplotype o is assembled by copying (with possible error) consecutive contiguous subsequences of haplotypes $h_j \in H$.

Definition 14 (Li and Stephens HMM)**.** Define $x_{j,i}$ as the event that the allele o_i at site i of the

haplotype o was copied from the allele $h_{j,i}$ of haplotype $h_j \in H$. Take parameters

$$\rho^*_{i-1\to i} \qquad \text{probability of any recombination between sites } i-1 \text{ and } i \tag{2.1}$$

$$\mu_i \qquad \text{probability of a mutation from one allele to another at site } i \tag{2.2}$$

and from them define the transition and recombination probabilities

$$p(x_{j,i}|x_{j',i-1}) = \begin{cases} 1-(k-1)\rho_i & \text{if } j = j' \\ \rho_i & \text{if } j \neq j' \end{cases} \qquad \text{where } \rho_i = \frac{\rho^*_{i-1\to i}}{k-1} \tag{2.3}$$

$$p(o_i|x_{j,i}) = \begin{cases} 1-(A-1)\mu_i & \text{if } o_i = h_{j,i} \\ \mu_i & \text{if } o_i \neq h_{j,i} \end{cases} \qquad \text{where } A = \text{number of alleles} \tag{2.4}$$

We will write $\mu_i(j)$ as shorthand for $p(o_i|x_{j,i})$. We will also define the values of the initial probabilities $p(x_{j,1}, o_1|H) = \frac{\mu_1(j)}{k}$, which can be derived by noting that if all haplotypes have equal probabilities $\frac{1}{k}$ of randomly being selected, and that this probability is then modified by the appropriate emission probability.

Let $P(o|H)$ be the probability that haplotype o was produced from population H. The forward algorithm for hidden Markov models allows calculation of this probability in $O(nk^2)$ time using an $n \times k$ dynamic programming matrix of *forward states*

$$p_i[j] = P(x_{j,i}, o_1, \ldots, o_i|H) \tag{2.5}$$

The probability $\mathrm{P}(o|H)$ will be equal to the sum $\sum_j p_n[j]$ of all entries in the final column of the dynamic programming matrix. In practice, the Li and Stephens forward algorithm is $O(nk)$. (See §2.3)

Li and Stephens like algorithms for large populations

The $O(nk)$ time complexity of the forward algorithm is intractable for reference panels with large size k. The UK Biobank has amassed $k = 500,000$ array samples. Whole genome sequencing projects, with a denser distribution of sites, are catching up. Major sequencing projects with $k = 100,000$ or more samples are nearing completion. Others numbering k in the millions have been announced. These large population datasets have significant potential benefits: They are statistically likely to more accurately represent population frequencies and those employing genome sequencing can provide phasing information for rare variants.

In order to handle datasets with size k even fractions of these sizes, modern haplotype inference algorithms depend on models which are simpler than the Li and Stephens model or which sample subsets of the data. For example, the common tools Eagle-2, Beagle, HAPI-UR and Shapeit-2 and -3 [69, 15, 136, 25, 91] either restrict where recombination can occur, fail to model mutation, model long-range phasing approximately or sample subsets of the reference panel.

Lunter's "fastLS" algorithm [72] demonstrated that haplotypes models which include all k reference panel haplotype could find the Viterbi maximum likelihood path in time sublinear in k, using preprocessing to reduce redundant information in the algorithm's input. However, his techniques do not extend to the forward and forward-backward algorithms.

Our contributions

We have developed an arithmetically exact forward algorithm whose expected time complexity is a function of the expected allele distribution of the reference panel. This expected time complexity proves to be significantly sublinear in reference panel size. We have also developed a technique for succinctly representing large panels of haplotypes whose size also scales as a sublinear function of the expected allele distribution.

Our forward algorithm contains three optimizations, all of which might be generalized to other bioinformatics algorithms. In (§2.2), we rewrite the reference panel as a sparse matrix containing the minimum information necessary to directly infer all allele values. In (§2.3), we define recurrence relations which are numerically equivalent to the forward algorithm but use minimal arithmetic operations. In (§2.4), we delay computation of forward states using a lazy evaluation algorithm which benefits from blocks of common sequence composed of runs of major alleles. Our methods apply to other models which share certain redundancy properties with the monoploid Li and Stephens model.

2.2 Sparse representation of haplotypes

The forward algorithm to calculate the probability $P(o|H)$ takes as input a length n vector o and a $k \times n$ matrix of haplotypes H. In general, any algorithm which is sublinear in its input inherently requires some sort of preprocessing to identify and reduce redundancies in the data. However, the algorithm will truly become effectively sublinear if this preprocessing can be amortized over many iterations. In this case, we are able to preprocess H into a sparse

representation which will on average contain better than $O(nk)$ data points.

This is the first component of our strategy. We use a variant of column-sparse-row matrix encoding to allow fast traversal of our haplotype matrix H. This encoding has the dual benefit of also allowing reversible size compression of our data. We propose that this is one good general data representation on which to build other computational work using very large genotype or haplotype data. Indeed, extrapolating from our single-chromosome results, the 1000 Genomes Phase 3 haplotypes across all chromosomes should simultaneously fit uncompressed in 11 GB of memory.

We will show that we can evaluate the Li and Stephens forward algorithm without needing to uncompress this sparse matrix.

2.2.1 Sparse column representation of haplotype alleles

Consider a biallelic genetic site i with alleles $\{A, B\}$. Consider the vector $h_{1,i}, h_{2,i}, \ldots, h_{k,i} \in \{A, B\}^k$ of alleles of haplotypes j at site i. Label the allele A, B which occurs more frequently in this vector as the major allele 0, and the one which occurs less frequently as the minor allele 1. We then encode this vector by storing the value A or B of the major allele 0, and the indices $j_1, j_2, \ldots$ of the haplotypes which take on allele value 1 at this site.

We will write ϕ_i for the subvector $h_{j_1,i}, h_{j_2,i}, \ldots$ of alleles of haplotypes consisting of those haplotypes which possess the minor allele 1 at site i. We will write $|\phi_i|$ for the multiplicity of the minor allele. We call this vector ϕ_i the *information content* of the haplotype cohort H at the site i.

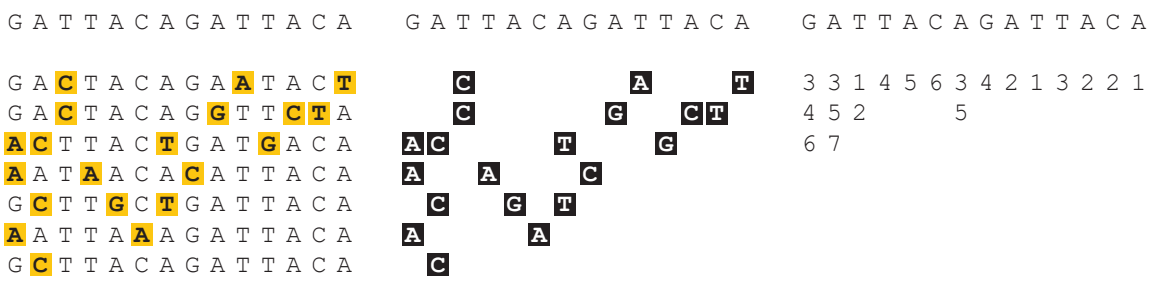


Figure 2.1: **Information content of array of template haplotypes.** **i)** Reference panel $\{h_1,\ldots,h_5\}$ with mismatches to haplotype o shown in yellow. **ii)** Alleles at site i of elements of $\phi_i(o_i)$ in black. **iii)** Vectors to encode $\phi_i(o_i)$ at each site.

2.2.2 Relation to the allele frequency spectrum

Our sparse representation of the haplotype reference panel benefits from the recent finding [58] that the distribution over sites of minor allele frequencies is biased towards low frequencies[1].

Clearly, the distribution of $|\phi_i|$ is precisely the allele frequency spectrum. More formally,

Lemma 6. Let $\mathbb{E}[\overline{f}](k)$ be the expected mean minor allele frequency for k genotypes. Then

$$\mathbb{E}\left[\frac{1}{n}\sum_{i=1}^{n}\left|\phi_i\right|\right] = \mathbb{E}[\overline{f}](k) \tag{2.6}$$

Corollary 2. If $O(\mathbb{E}[\overline{f}]) < O(k)$, then $O(\sum_i \left|\phi_i\right|) < O(nk)$ in expected value.

[1] We observe similar results in our own analyses in section 2.5.2

2.2.2.1 Dynamic reference panels

Adding or rewriting a haplotype is constant time per site per haplotype unless this edit changes which allele is the most frequent. It can be acheived by addition or removal or single entries from the row-sparse-column representation, wherein, since our implementation does not require that the column indices be stored in order, these operations can be made $O(1)$ This allows our algorithm to extend to uses of the Li and Stephens model where one might want to dynamically edit the reference panel. The exception occurs when $\phi_i = \frac{k}{2}$—here it is not absolutely necessary to keep the formalism that the indices stored actually be the minor allele.

2.2.3 Implementation

For biallelic sites, we store our ϕ_i's using a length-n vector of length $|\phi_i|$ vectors containing the indices j of the haplotypes $h_j \in \phi_i$ and a length-n vector listing the major allele at each site. (See Figure 2.1 panel iii) Random access by key i to iterators to the first elements of sets ϕ_i is $O(1)$ and iteration across these ϕ_i is linear in the size of ϕ_i. For multiallelic sites, the data structure uses slightly more space but has the same speed guarantees.

Generating these data structures takes $O(nk)$ time but is embarrassingly parallel in n. Our "*.slls" data structure doubles as a succinct haplotype index which could be distributed instead of a large VCF record (though genotype likelihood compression is not accounted for). A .vcf $\rightarrow$ .slls conversion tool is found in our github repository.

2.3 Efficient dynamic programming

We begin with the recurrence relation of the classic forward algorithm applied to the Li and Stephens model [66]. To establish our notation, recall that we write $p_i[j] = P(x_{j,i}, o_1, \ldots, o_i|H)$, that we write $\mu_i(j)$ as shorthand for $p(o_i|x_{j,i})$ and that we have initialized $p_1[j] = p(x_{j,1}, o_1|H) = \frac{\mu_1(j)}{k}$. For $i > 1$, we may then write:

$$p_i[j] = \mu_i(j)\big((1-k\rho_i)p_{i-1}[j] + \rho_i S_{i-1}\big) \tag{2.7}$$

$$S_i = \sum_{j=1}^{k} p_i[j] \tag{2.8}$$

We will reduce the number of summands in (2.8) and reduce the number indices j for which (2.7) is evaluated. This will use the **information content** defined in (§2.2.1).

Lemma 7. The summation (2.8) is calculable using strictly fewer than k summands.

Proof. Suppose first that $\mu_i(j) = \mu_i$ for all j. Then

$$S_i = \sum_{j=1}^{k} p_i[j] = \mu_i \sum_{j=1}^{k} \big((1-k\rho_i)p_{i-1}[j] + \rho_i S_{i-1}\big) \tag{2.9}$$

$$= \mu_i\big((1-k\rho_i)S_{i-1} + k\rho_i S_{i-1}\big) = \mu_i S_{i-1} \tag{2.10}$$

Now suppose that $\mu_i(j) = 1-\mu_i$ for some set of j. We must then correct for these j. This gives us

$$S_i = \mu_i S_{i-1} + \frac{1-\mu_i-\mu_i}{1-\mu_i} \sum_{j \text{ where } \mu_i(j)\neq\mu_i} p_i[j] \tag{2.11}$$

The same argument holds when we reverse the roles of μ_i and $1-\mu_i$. Therefore we can choose which calculation to perform based on which has fewer summands. This gives us the following formula:

$$S_i = \alpha S_{i-1} + \beta \sum_{j\in\phi_i} p_i[j] \tag{2.12}$$

where

$$\alpha = \quad \mu_i \qquad \beta = \quad \frac{1-2\mu_i}{1-\mu_i} \qquad \text{if } \phi_i \text{ have allele a} \tag{2.13}$$

$$\alpha = \quad 1-\mu_i \qquad \beta = \quad \frac{2\mu_i-1}{\mu_i} \qquad \text{if } \phi_i \text{ do not have allele a} \tag{2.14}$$

□

We note another redundancy in our calculations. For the proper choices of μ_i', μ_i'' among $\mu_i, 1-\mu_i$, the recurrence relations (2.7) are linear maps $\mathbb{R} \to \mathbb{R}$

$$f_i : x \longmapsto \mu_i'(1-k\rho)x + \mu_i'\rho S_{i-1} \tag{2.15}$$

$$F_i : x \longmapsto \mu_i''(1-k\rho)x + \mu_i''\rho S_{i-1} \tag{2.16}$$

of which there are precisely two unique maps, f_i corresponding to the recurrence relations for those x_j such that $j \in \phi_i$, and F_i to those such that $j \notin \phi_i$.

Lemma 8. If $j \notin \phi_i$ and $j \notin \phi_{i-1}$, then S_i can be calculated without knowing $p_{i-1}[j]$ and $p_i[j]$.

If $j \notin \phi_{i-1}$ and $j' \neq j$, then $p_i[j']$ can be calculated without knowing $p_{i-1}[j]$.

Proof. Equation (2.12) lets us calculate S_{i-1} without knowing any $p_{i-1}[j]$ for any $j \notin \phi_{i-1}$. From S_{i-1} we also have f_i and F_i. Therefore, we can calculate $p_i[j'] = f_i(p_{i-1}[j'])orF_i(p_{i-1}[j'])$ without knowing $p_{i-1}[j]$ provided that $j' \neq j$. This then shows us that we can calculate $p_i[j']$ for all $j' \in \phi_i$ without knowing any j such that $j \notin \phi_i$ and $j \notin \phi_{i-1}$. Finally, the first statement follows from another application of (2.12). □

Corollary 3. The recurrences (2.8) and the minimum set of recurrences (2.7) needed to compute (2.8) can be evaluated in $O(|\phi_i|)$ time, assuming that $p_{i-1}[j]$ have been computed $\forall j \in \phi_i$.

We address the assumption on prior calculation of the necessary $p_{i-1}[j]$'s in section 2.4.

2.3.1 Time complexity

Recall that we defined $\mathbb{E}[\overline{f}](k)$ as the expected mean minor allele frequency in a sample of size k. Suppose that it is comparatively trivial to calculate the missing $p_{i-1}[j]$ values. Then by Corollary 3 the procedure in eq. (2.12) has expected time complexity $O\left(\sum_i |\phi_i|\right) = O\left(n\mathbb{E}[\overline{f}](k)\right)$.

Figure 2.2: **Work done to calculate the sum of haplotype probabilities at a site for the conventional and our sublinear forward algorithm.** Using the example that at site i, $\phi_i(o_i) = \{h_3\}$, we illustrate the number of arithmetic operations used in **i)** the conventional $O(nk)$ Li and Stephens HMM recurrence relations **ii)** Our procedure specified in equation (2.12). Black lines correspond to arithmetic operations; operations which cannot be parallelized over j are colored yellow.

2.4 Lazy evaluation of dynamic programming rows

Corollary 3 was conditioned on the assumption that specific forward probabilities had already been evaluated. We will describe a second algorithm which performs this task efficiently by avoiding performing any arithmetic which will prove unnecessary at future steps.[2]

2.4.1 Equivalence classes of longest major allele suffixes

Lemma 9. Suppose that $h_j \notin \phi_\ell \cup \phi_{\ell+1} \cup \ldots \cup \phi_{i-1}$. Then the dynamic programming matrix entries $p_\ell[j], p_{\ell+1}[j], \ldots, p_{i-1}[j]$ need not be calculated in order to calculate $S_\ell, S_{\ell+1}, \ldots, S_{i-1}$.

Proof. By repeated application of Lemma (8). □

Corollary 4. Under the same assumption on j, $p_\ell[j], p_{\ell+1}[j], \ldots, p_{i-1}[j]$ need not be calculated in order to calculate $F_{\ell+1}, \ldots, F_i$. This is easily seen by definition of F_i.

Lemma 10. Suppose that $p_{\ell-1}[j]$ is known, and $x_j \notin \phi_\ell \cup \phi_{\ell+1} \cup \ldots \cup \phi_{i-1}$. Then $p_{i-1}[j]$ can be calculated in the time which it takes to calculate $F_{i-1} \circ \cdots \circ F_\ell$.

[2] This approach is known as *lazy evaluation*.

Proof. $p_{i-1}[j] = F_{i-1} \circ \cdots \circ F_\ell(p_{\ell-1}[j])$ □

It is immediately clear that calculating the $p_i[j]$ lends well to lazy evaluation. Specifically, the $x_j \notin \phi_i$ are data which need not be evaluated yet at step i. Therefore, if we can aggregate the work of calculating these data at a later iteration of the algorithm, and only if needed then, we can potentially save a considerable amount of computation.

Definition 15 (Longest major allele suffix classes)**.** Define $E_{\ell \to i-1} = \phi_{\ell-1} \cap \left[\bigcup_{\iota=\ell}^{i-1} \phi_\iota\right]^c$ That is, let $E_{\ell \to i-1}$ be the class of all haplotypes whose sequence up to site $i-1$ shares the suffix from ℓ to $i-1$ inclusive consisting only of major alleles, but lacks any longer suffix composed only of major alleles.

$E_{\ell \to i-1}$ is the set of all h_j where $p_{\ell-1}[j]$ was needed to calculate $S_{\ell-1}$ but no $p_{(\cdot)}[j]$ has been needed to calculate any $S_{(\cdot)}$ since.

Note that for each i, the equivalence classes $E_{\ell \to i-1}$ form a disjoint cover of the set of all haplotypes $h_j \in H$.

$\forall h_j \in E_{\ell \to i-1}$, $p_{i-1}[j] = F_{i-1} \circ \cdots \circ F_\ell(p_{\ell-1}[j])$

Definition 16. Write $F_{a \to b}$ as shorthand for $F_b \circ \cdots \circ F_a$.

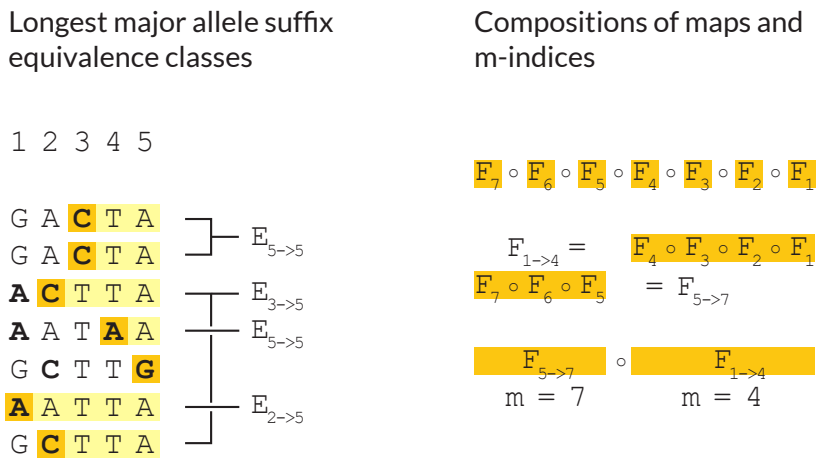


Figure 2.3: **Longest major allele suffix classes, linear map compositions.** Illustrations clarifying the meanings of the equivalence classes $E_{\ell\to i-1}$ (left) and the maps $F_{a\to b}$. Indices m are sites whose indices are b's in stored maps of the form $F_{a\to b}$.

2.4.2 The lazy evaluation algorithm

Our algorithm will aim to:

1. Never evaluate $p_i[j]$ explicitly unless $h_j \in \phi_i$.

2. Amortize the calculations $p_i[j] = f_i \circ F_{i-1} \circ \cdots \circ F_\ell(p_{\ell-1}[j])$ over all $h_j \in E_{\ell\to i-1}$.

3. Share the work of calculating subsequences of compositions of maps $F_{i-1} \circ \cdots \circ F_\ell$ with other compositions of maps $F_{i'-1} \circ \cdots \circ F_{\ell'}$ where $\ell' \leq \ell$ and $i' \geq i$.

To accomplish these goals, at each iteration i, we maintain the following auxiliary data. The meaning of these are clarified by reference to figures (2.3), (2.4) and (2.5).:

1. The partition of all haplotypes $h_j \in H$ into equivalence classes $E_{\ell \to i-1}$ according to longest major allele suffix of the truncated haplotype at $i-1$. See Definition 2 and figure (2.3).

2. The tuples $T_\ell = (E_{\ell \to i-1}, F_{\ell \to m}, m)$ of equivalence classes $E_{\ell \to i-1}$ stored with linear map prefixes $F_{\ell \to m} = F_m \circ \cdots \circ F_\ell$ of the map $F_{\ell \to i-1}$ which would be necessary to fully calculate $p_i[j]$ for the j they contain, and the index m of the largest index in this prefix. See figure (2.5).

3. The ordered sequence $m_1 > m_2 > \ldots$, in reverse order, of all distinct $1 \leq m \leq i-1$ such that m is contained in some tuple. See figures (2.3), (2.5).

4. The maps $F_{min\{\ell\} \to m_{min}}, \ldots, F_{m_2+1 \to m_1}, F_{m_1+1 \to i-1}$ which partition the longest prefix $F_{i-1} \circ \cdots \circ F_{min\{\ell\}}$ into disjoint submaps at the indices m. See figure (2.3). These are used to rapidly extend prefixes $F_{\ell \to m}$ into prefixes $F_{\ell \to i-1}$.

Finally, we will need the following ordering on tuples T_ℓ to describe our algorithm:

Definition 17. Impose a partial ordering $<$ on the $T_\ell = (E_{\ell \to i-1}, F_{\ell \to m}, m)$ by $T_\ell < T_{\ell'}$ iff $m < m'$. See figure (2.4).

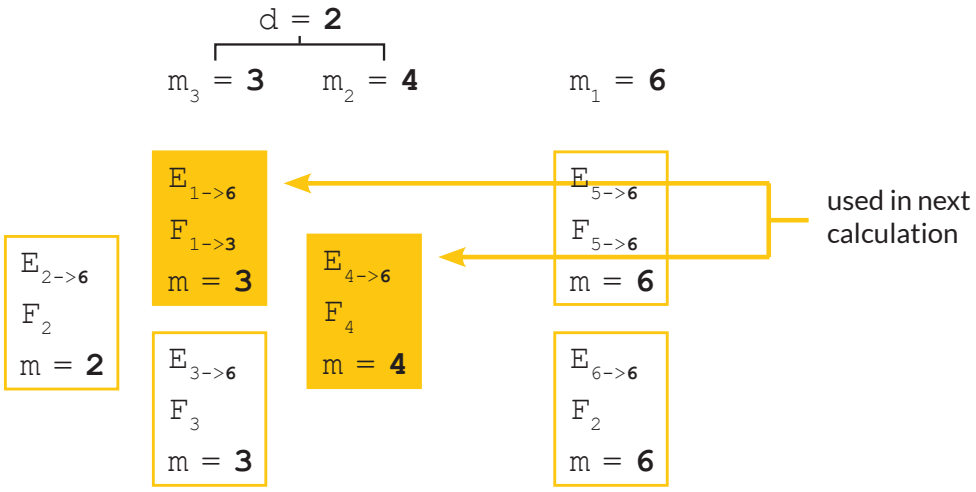


Figure 2.4: **Partial ordering of tuples of (equivalence class, linear map, index) used as state information in our algorithm. The ordering of the tuples $T_\ell = (E_{\ell\to i-1}, F_{\ell\to m}, m)$. Calculation of the depth d of an update which requires haplotypes contained in the equivalence classes defining the two tuples shown in solid yellow.**

We are now ready to describe our lazy evaluation algorithm which evaluates $p_i[j] = f_i \circ F_{\ell\to i-1}(p_{\ell-1}[j])$ just-in-time while fulfilling the aims listed at the top of this section, by using the auxiliary state data specified above.

The algorithm is simple but requires keeping track of a number of intermediate indices. We suggest referring to the figures (2.3 - 2.5) as a visual aid. We state it in 6 steps as follows:

Step 1: ***Identifying the tuples containing*** ϕ **—** $O(\phi_i)$ **time complexity:**

Identify the subset $U(\phi)$ of the tuples T_ℓ for which there exists some $h_j \in \phi_i$ such that $h_j \in E_{\ell \to i-1}$.

Step 2: ***Identifying the preparatory map suffix calculations to be performed —*** $O(\phi_i)$ **time complexity:**

Find the maximum depth d of any $T_\ell \in U(\phi)$ with respect to the partial ordering above. Equivalently, find the minimum m such that $T_\ell = (E_{\ell \to i-1}, F_{\ell \to m}, m) \in U(\phi)$. See figure (2.4).

Step 3: ***Performing preparatory map suffix calculations —*** $O(d)$ **time complexity:**

1. $O(d)$: Let $m_1, \ldots, m_d$ be the last d indices m in the reverse ordered list of indices $m_1, m_2, \ldots$. By iteratively composing the maps $F_{m_1+1 \to i-1}$,$F_{m_2+1 \to m_1}$ which we have already stored, construct the telescoping suffixes $F_{m_1+1 \to i-1}$, $F_{m_2+1 \to i-1}, \ldots, F_{m_d+1 \to i-1}$ needed to update the tuples $(E_{\ell \to i-1}, F_{\ell \to m}, m)$ to $(E_{\ell \to i-1}, F_{\ell \to i-1}, i-1)$.

2. $O(d)$: For each $m_1 \leq m_i \leq m_d$, choose an arbitrary $(E_{\ell \to i-1}, F_{\ell \to m_i}, m_i)$ and update it to $(E_{\ell \to i-1}, F_{\ell \to i-1}, i-1)$.

Step 4: ***Performing the deferred calculations for the tuples containing*** $h_j \in \phi_i$ **—** $O(\phi_i)$ **time complexity:**

If not already done in Step 3.2, for every $T_\ell \in U(\phi)$, extend its map element from $(E_{\ell \to i-1}, F_{\ell \to m}, m)$ to $(E_{\ell \to i-1}, F_{\ell \to i-1}, i-1)$ in $O(1)$ time using the maps calculated in Step 3.1. See figure (2.5).

Step 5: *Calculating* $p_i[j]$ *just-in-time* — $O(\phi_i)$ time complexity:

Note: The calculation of interest is performed here.

Using the maps $F_{\ell\to i-1}$ calculated in Step 3.2 or 4, finally evaluate the value $p_i[j] = f_i \circ F_{\ell\to i-1}(p_{\ell-1}[j])$. See figure (2.5).

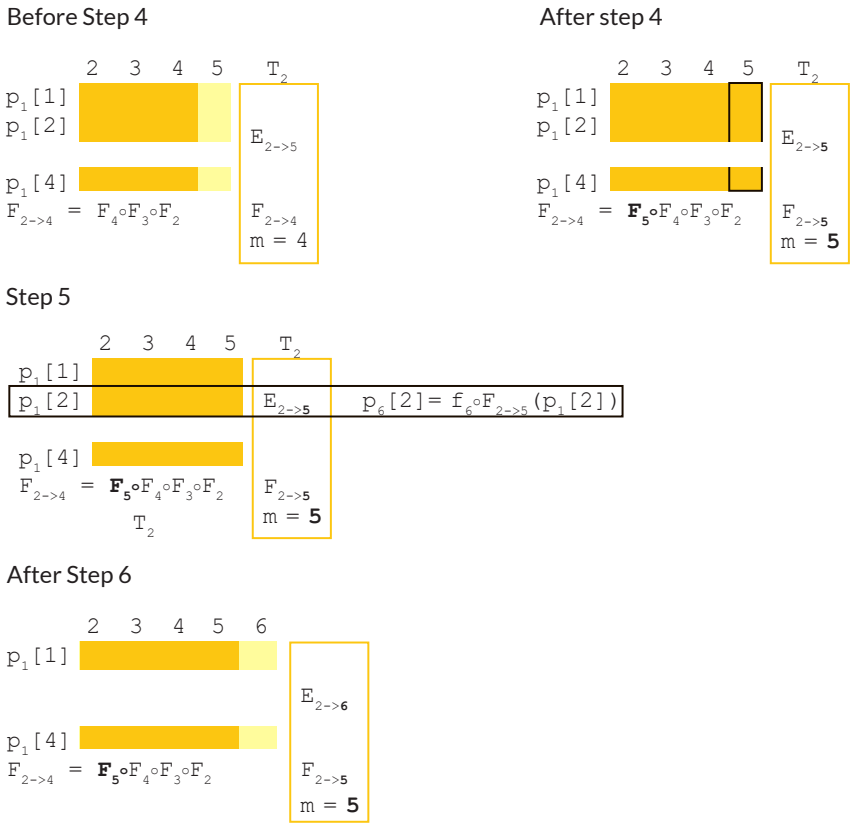


Figure 2.5: **Key steps involved in calculating $p_i[j]$ by delayed evaluation.** An illustration of the manipulation of the tuple $T_2 = (E_{\ell \to i-1}, F_{\ell \to m}, m)$ by the lazy evaluation algorithm, and how it is used to calculate $p_i[j]$ from $p_{\ell-1}[j]$ just-in-time. In this case, we wish to calculate $p_6[2]$. This is a member of the equivalence class $E_{2\to 5}$, since it hasn't needed to be calculated since time 1. In step 4 of the algorithm, we therefore must update the whole tuple T_2 by post-composing the partially completed prefix $F_{2\to 4}$ of the map $F_{2\to 5}$ which we need using our already-calculated suffix map F_5. In step 5, we use $F_{2\to 5}$ to compute $p_6[2] = f_6 \circ F_{2\to 5}(p_1[j])$. In step 6, we update the tuple T_2 to reflect its loss of h_2, which is now a member of $E_{6\to 6}$

Step 6: ***Updating our equivalence class/update map prefix tuple auxiliary data structures*** **— $O(\phi_i + d)$ time complexity:**

1. Create the new tuple $(E_{i\to i}, F_{i\to i} =$ identity map $, i)$

2. Remove the $h_j \in \phi_i$ from their equivalence classes $E_{\ell \to i-1}$ and place them in the new equivalence class $E_{i\to i}$. If this empties the equivalence class in question, delete its tuple. *To maintain memory use bounded by number of haplotypes, our implementation uses an object pool to store these tuples*

3. If an index m_i no longer has any corresponding tuple, delete it, and furthermore replace the stored maps $F_{m_{i-1}+1 \to m_i}$ and $F_{m_i+1} \to m_{i+1}$ with a single map $F_{m_{i-1}+1 \to m_{i+1}}$. *This step is added to reduce the upper bound on the maximum possible number of compositions of maps which are performed in any given step.*

The following two trivial lemmas allow us to bound d by k such that the aggregate

time complexity of the lazy evaluation algorithm cannot exceed $O(nk)$. Due to the irregularity of the recursion pattern used by the algorithm, is likely not possible to calculate a closed-form tight bound on $\sum_i d$, however, empirically it is asymptotically dominated by $\sum_i \phi_i$ as shown in the results which follow.

Lemma 11. The number of nonempty equivalence classes $E_{\ell \to i-1}$ in existence at any iteration i of the algorithm is bounded by the number of haplotypes k.

Proof. Trivial but worth noting. □

Lemma 12. The number of unique indices m in existence at any iteration i of the algorithm is bounded by the number of nonempty equivalence classes $E_{\ell \to i-1}$.

2.5 Results

2.5.1 Implementation

Our algorithm was implemented as a C++ library located at `https://github.com/yoheirosen/sublinear-Li-Stephens` Details of the lazy evaluation algorithm will be found there.

We also implemented the linear time forward algorithm for the haploid Li and Stephens model in C++ as to evaluate it on identical footing. Profiling was performed using a single Intel Xeon X7560 core running at 2.3 GHz on a shared memory machine. Our reference panels H were the phased haplotypes from the 1000 Genomes [23] phase 3 VCF records for chromosome 22 and subsamples thereof. Haplotypes o were randomly generated simulated descendants.

2.5.2 Minor allele frequency distribution for the 1000 Genomes dataset

We found it informative to determine the allele frequency spectrum for the 1000 Genomes dataset which we will use in our performance analyses. We simulated haplotypes o of $1,000,000$ bp length on chromosome 22 and recorded the sizes of the sets $\phi_i(o_i)$ for $k = 5008$. These data produced a mean $|\phi_i(o_i)|$ of 59.9, which is 1.2% of the size of k. We have plotted the distribution of $|\phi_i(o_i)|$ which we observed from this experiment in (Fig. 2.6). It is skewed toward low frequencies; the minor allele is unique at 71% of sites, and it is below 1% frequency at 92% of sites.

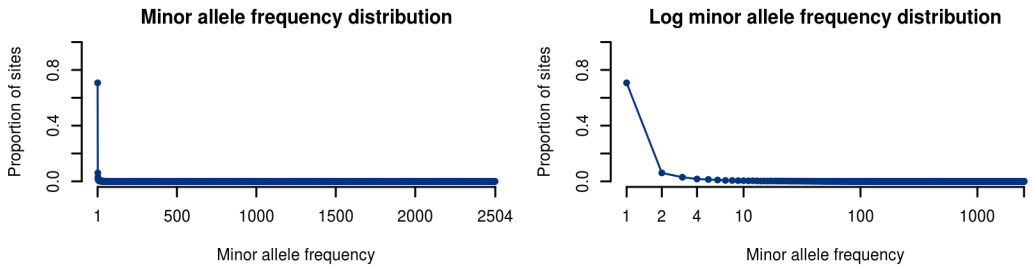


Figure 2.6: **Biallelic site minor allele frequency distribution from 1000 Genomes chromosome 22.** Note that the distribution is skewed away from the $\frac{1}{f}$ distribution classically theorized. The data used are the genotypes of the 1000 Genomes Phase 3 VCF, with minor alleles at multiallelic sites combined.

2.5.3 Comparison of our algorithm with the linear time forward algorithm

In order to compare the dependence of our algorithm's runtime on haplotype panel size k against that of the standard linear LS forward algorithm, we measured the CPU time per genetic site of both across a range of haplotype panel sizes from 30 to 5008. This analysis was achieved as briefly described above. Haplotype panels spanning the range of sizes from 30 to

5008 haplotypes were subsampled from the 1000 Genomes phase 3 VCF records and loaded into memory in both uncompressed and our column-sparse-row format. Random sequences were sampled using a copying model with mutation and recombination, and the performance of the classical forward algorithm was run back to back with our algorithm for the same random sequence and same subsampled haplotype panel. Each set of runs was performed in triplicate to reduce stochastic error.

Figure 2.7 shows this comparison. Observed time complexity of our algorithm was $O(k^{0.35})$ as calculated from the slope of the line of best fit to a log-log plot of time per site versus haplotype panel size.

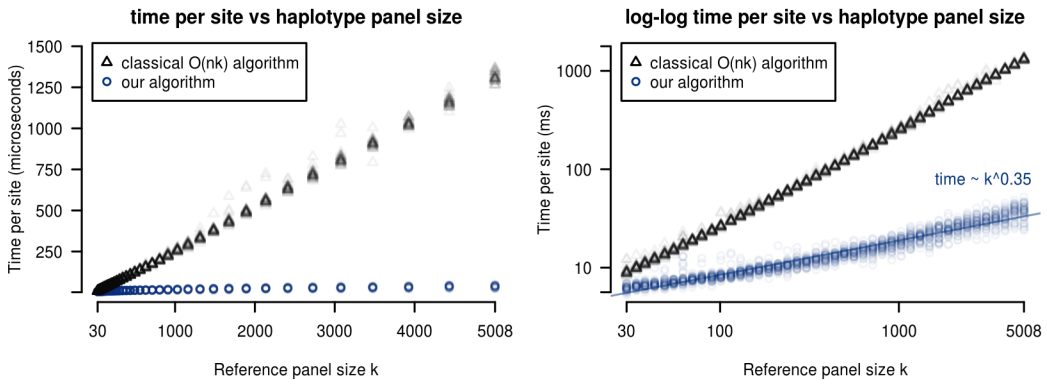


Figure 2.7: **Runtime per site for conventional linear algorithm vs our sparse-lazy algorithm.** Runtime per site as a function of haplotype reference panel size k for our algorithm (blue) as compared to the classical linear time algorithm (black). Both were implemented in C++ and benchmarked using datasets preloaded into memory. Forward probabilities are calculated for randomly generated haplotypes simulated by a recombination-mutation process, against random subsets of the 1000 genomes dataset.

For data points where we used all 1000 Genomes project haplotypes ($k = 5008$), on

average, time per site is 37 μs for our algorithm and 1308 μs for the linear LS algorithm. For the forthcoming 100,000 Genomes Project, these numbers can be extrapolated to 251 μs for our algorithm and 260,760 μs for the linear LS algorithm.

2.5.3.1 Lazy evaluation of dynamic programming rows

We also measured the time which our algorithm spent within the d-dependent portion of the lazy evaluation subalgorithm. In the average case, the time complexity of our lazy evaluation subalgorithm does not contribute to the overall algebraic time complexity of the algorithm. (Fig. 2.8, right) The lazy evaluation runtime also contributes minimally to the total actual runtime of our algorithm. (Fig. 2.8, left)

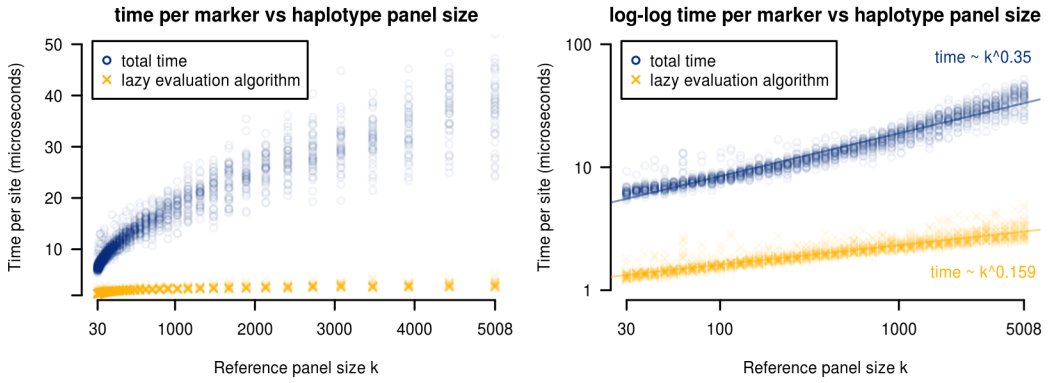


Figure 2.8: **Runtime per site for the overall algorithm and for the recursion-depth dependent portion.** Time per site for the lazy evaluation subalgorithm (yellow) vs. the full algorithm (blue). The experimental setup is the same as previously described, with the subalgorithm time determined by internally timing the recursion-depth d dependent portions of the lazy evaluation subalgorithm.

2.5.4 Sparse haplotype encoding

2.5.4.1 Generating our sparse vectors

We generated the haplotype panel data structures from (§2.2) using the VCF-encoding tool `vcf2slls` which we provide. We built indices with multiallelic sites, which increases their time and memory profile relative to the results in (§2.5.2) but allows direct comparison to VCF records. Encoding of chromosome 22 was completed in 38 minutes on a single CPU core. Use of M CPU cores will reduce runtime proportional to M.

2.5.4.2 Size of sparse haplotype index

In uncompressed form, our whole genome `*.slls` index for chromosome 22 of the 1000 genomes dataset was 285 MB in size versus 11 GB for the VCF record using `uint16_t`'s to encode haplotype ranks. When compressed with gzip, the same index was 67 MB in size versus 205 MB for the VCF record.

In the interest of speed (both for our algorithm and the $O(nk)$ algorithm) our experiments loaded entire chromosome sparse matrices into memory and stored haplotype indices as `uint64_t`'s. This requires on the order of 1 GB memory for chromosome 22. For long chromosomes or larger reference panels on low memory machines, the algorithm can operate by streaming sequential chunks of the reference panel.

2.6 Conclusion

To the best of our knowledge, ours is the first forward algorithm for any haplotype model to attain sublinear time complexity with respect to reference panel size. Our algorithms could be incorporated into haplotype inference strategies by interfacing with our C++ library. This opens the potential for tools which are tractable on haplotype reference panels at the scale of current 100,000 to 1,000,000+ sample sequencing projects.

2.6.1 Applications which use individual forward probabilities

Our algorithm attains its runtime specifically for the problem of calculating the single overall probability $P(o|H,\rho,\mu)$ and does not compute all nk forward probabilities. We can prove that if m many specific forward probabilities are also required as output, and if the time complexity of our algorithm is $O(\sum_i |\phi_i|)$, then the time complexity of the algorithm which also returns the m forward probabilities is $O(\sum_i |\phi_i| + m)$.

In general, haplotype phasing or genotype imputation tools use stochastic traceback or other similar sampling algorithms. The standard algorithm for stochastic traceback samples states from the full posterior distribution and therefore requires all forward probabilities. The algorithm output and lower bound of its speed is therefore $O(nk)$. The same is true for many applications of the forward-backward algorithm.

There are two possible approaches which might allow runtime sublinear in k for these applications. Using stochastic traceback as an example, first is to devise an $O(f(m))$ sampling algorithm which uses $m = g(k)$ forward probabilities such that $O(f \circ g(k)) < O(k)$. The second

is to succinctly represent forward probabilities such that nested sums of the nk forward probabilities can be queried from $O(\phi) < O(nk)$ data. This should be possible, perhaps using the positional Burrows-Wheeler transform [29] as in [72], since we have already devised a forward algorithm with this property for a different model in [108].

2.6.2 Generalizability of algorithm

The optimizations which we have made are not strictly specific to the monoploid Li and Stephens algorithm. Necessary conditions for our reduction in the time complexity of the recurrence relations are The number of distinct transition probabilities is constant with respect to number of states k. The number of distinct emission probabilities is constant with respect to number of states k. Favourable conditions for efficient time complexity of the lazy evaluation algorithm are The number of unique update maps added per step is constant with respect to number of states k. The update map extension operation is composition of functions of a class where composition is constant-time with respect to number of states k. The reduction in time complexity of the recurrence relations depends on the Markov property, however we hypothesize that the delayed evaluation needs only the semi-Markov property.

2.6.2.1 Other haplotype forward algorithms

Our optimizations are of immediate interest for other haplotype copying models. The following related algorithms have been explored without implementation.

Example 1 (Diploid Li and Stephens)**.** We have yet to implement this model but expect average runtime at least subquadratic in reference panel size k. We build on the statement of the model

and its optimizations in [67]. We have found the following recurrences which we believe will work when combined with a system of lazy evaluation algorithms:

Lemma 13. The diploid Li and Stephens HMM may be expressed using recurrences of the form

$$p_i[j_1, j_2] = \alpha_p p_{i-1}[j_1, j_2] + \beta_p(S_{i-1}(j_1) + S_{i-1}(j_2)) + \gamma_p S_{i-1} \tag{2.17}$$

which use on the intermediate sums defined as

$$S_i := \alpha_c S_{i-1} + \beta_c \sum_{j \in \phi_i} S_{i-1}(j) + \gamma_c \sum_{(j_1, j_2) \in \phi_i^2} p_{i-1}[j_1, j_2] \qquad O(|\phi_i|^2) \tag{2.18}$$

$$S_i(j) := \alpha_c S_{i-1} + \beta_c S_{i-1}(j) + \gamma_c \sum_{j_2 \in \phi_i} p_{i-1}[j, j_2] \qquad \text{for } O(k|\phi_i|) \text{ many } j \tag{2.19}$$

where $\alpha_{(\cdot)}, \beta_{(\cdot)}, \gamma_{(\cdot)}$ depend only on the diploid genotype o_i.

Implementing and verifying the runtime of this extension of our algorithm will be among our next steps.

Example 2 (Multipopulation Li and Stephens)**.** [26] We maintain separate sparse haplotype panel representations $\phi_i^A(o_i)$ and $\phi_i^B(o_i)$ and separate lazy evaluation mechanisms for the two populations A and B. Expected runtime guarantees are similar.

This model, and versions for > 2 populations, will be important in large sequencing cohorts (such as NHLBI TOPMed) where assuming a single related population is unrealistic.

Example 3 (More detailed mutation model)**.** It may also be desirable to model distinct mutation probabilities for different pairs of alleles at multiallelic sites. Runtime is worse than the biallelic

model but remains average case sublinear.

Example 4 (Sequence graph Li and Stephens analogue)**.** In [108] we described a hidden Markov model for a haplotype-copying with recombination but not mutation in the context of sequence graphs. Assuming we can decompose our graph into nested sites then we can achieve a fast forward algorithm with mutation. An analogue of our row-sparse-column matrix compression for sequence graphs is being actively developed within our research group.

While a haplotype HMM forward algorithm alone might have niche applications in bioinformatics, we expect that our techniques are generalizable to speeding up other forward algorithm-type sequence analysis algorithms.

Chapter 3

Modelling haplotypes with respect to reference cohort variation graphs

Preamble

This chapter is adapted from my paper "Modelling haplotypes with respect to reference cohort variation graphs" for which I am first author [108]. The paper describes my theoretical work on adapting an algorithm in genomics from the linear reference paradigm to graph representations of reference cohorts of haplotypes. I contributed the mathematical concepts, algorithms, proofs and writing which follows. Jordan Eizenga helped with initial brainstorming as well as checking and proofreading of the mathematical work herein.

3.1 Abstract

We present a model for an observed haplotype having arisen through recombinations from a population whose sequences are represented in a variation graph. A variation graph is a structure for representing genetic sequences in which variation between haplotypes can be expressed through divergence of sequences of nodes. It is important to extend haplotype inference to variation graphs because these structures allow representation of complex relationships between haplotypes including arbitrary structural variation. They furthermore admit succinct representation of relatedness across very large reference cohorts. Our model generalizes the recombination component of the classical Li and Stephens model and introduces the concept of the "rectangular decomposition" of a haplotype with respect to a reference cohort embedded in a variation graph. We describe efficient algorithms to construct this decomposition and to use it to compute haplotype likelihoods. As demonstrations, we use this model to compute the likelihoods of short reads and of chromosome-length haplotypes given a reference panel. Our results suggest that haplotype inference using variation graphs is feasible.

3.2 Background

Statistical modelling of individual haplotypes within population distributions of genetic variation dates back to Kingman's *n-coalescent* [60]. In general, the coalescent and other models describe haplotype sequences as generated from some structured state space via recombination and mutation events.

Whereas coalescent models are powerful generative tools, their computational com-

plexity is unsuited to inference on observed haplotype data. Therefore, the dominant haplotype likelihood model used for statistical inference is Li and Stephens' 2003 model (LS) [66] and its various modifications. LS closely approximates to the more exact coalescent models but admits implementations with rapid runtime. A recent Lunter paper [72] refactors LS to run independent of reference cohort size by using Durbin's positional Burrows Wheeler transform [29], a data representation designed for rapid haplotype matching. However, these use relatively simplistic representations of variation.

Orthogonal to statistical models, another important frontier in genomics is the development of variation graphs [97]. In brief, these use mathematical graphs to encode the wide variety of polymorphism and structural variation found in the population. Variation graphs are a natural structure to represent reference cohorts for haplotype inference since they encode haplotypes in a canonical manner: as node sequences compatible with the edge structure of the graph [89].

In this paper, we present the first statistical model for haplotype inference on variation graphs. We also describe an efficient algorithm for calculating haplotype likelihoods with large reference panels. The algorithm makes significant use of Novak's graph positional Burrows-Wheeler transform (gPBWT) [89], a graph analogue to Durbin's PBWT [29] which underlies Lunter's "fast Li-Stephens" algorithm [72].

3.3 Embedding haplotypes in a variation graph

Classically, haplotypes have been represented by values at a set of linearly ordered sites [66, 72]. A variation graph is an alternative structure for storing genetic information which is composed of nodes labelled with DNA sequences [97]. Edges are placed between two nodes if the cohort contains individuals in whose genomes these sequences are adjacent. Haplotypes are naturally represented in these graphs as sequences of nodes. The node labels of these graph-embedded haplotypes spell out their DNA sequence.

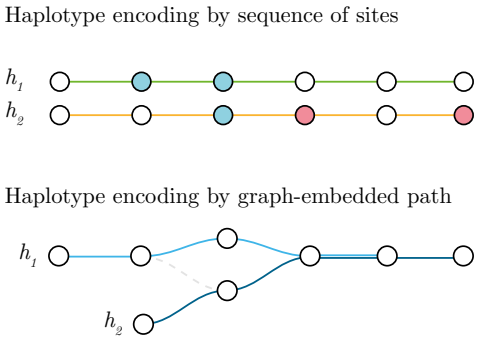


Figure 3.1: **LS model**: haplotypes are fixed-length sequences of alleles (colored). **Graph embedded paths**: haplotypes are paths in a graph.

Variation produces characteristic structures in these graphs. For example, simple polymorphisms among a cohort are represented as "bubbles," repeats as loops in the graph, and deletions or more complex structural rearrangements are represented by addition of new edges between nodes. Note that the latter motifs are not representable in the linearly ordered structures used in classical haplotype inference. A principal advantage of variation graphs is the ability to represent all types of variation. However, LS and related haplotype models benefit from their limited representation of variation. In particular, they admit simpler models of

mutation than exist for graphs.

3.4 A recombination-only model for variation graphs

The Li and Stephens model (LS) [66, 72] can be described by an HMM with a state space consisting of a haplotype cohort and observations consisting of alleles at sites. Transitions between states correspond to recombinations, and mutations are modelled within the emission probabilities. Our model mimics LS's recombination model, but we omit mutations since the mutation model of LS does not directly translate to graphs. Since variation graphs can embed full DNA sequences rather than lists of sites, we extend the model to allow recombinations at base-pair resolution rather than just between sites.

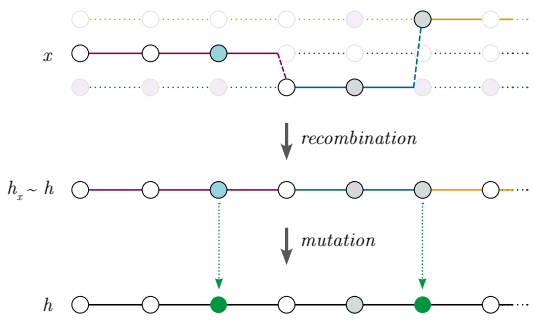


Figure 3.2: **The Li and Stephens model: a (Markovian) random walk x on the set H of haplotypes generates an allele sequence h** Mutation is contained completely in the emmission probabilities of the underlying HMM

Given a variation graph G, and a set of embedded haplotypes H, let x be a DNA sequence spelled out by a concatenation of subhaplotypes of haplotypes in H. Between each subhaplotype is a recombination. We will call this x a mosaic of subhaplotypes. Let $|x|$ be the

length (in base pairs) of x, and let $R(x)$ be the number of recombinations in x. We can assign x a probability by representing it as a Markov chain $x_0 \ldots x_i x_{i+1} \ldots x_{|x|}$ of length $|x|$ on the set H of haplotypes. Given transition probabilities

$$p_{recomb} \text{ if } x_i \neq x_{i+1}$$

$$p_{continue} = (1 - |H| p_{recomb}) \text{ if } x_i = x_{i+1}$$

then this model gives us $P(x|G,H) = p_{recomb}^{R(x)}(1 - |H| p_{recomb})^{|h|-R(x)}$.[1]

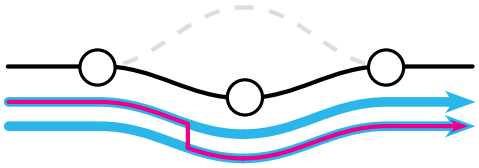


Figure 3.3: **Black path:** an observed haplotype h. Blue arrows: two subhaplotypes with the same node sequence as h. **Red path:** random walk $x \in \chi(h)$ which produces haplotype h.

Seeing as inference is concerned with the observed haplotype sequences rather than the mosaics which produced them, given query haplotype h, graph G, multiset of embedded haplotypes H, we seek a model for the likelihood of h with respect to the graph-embedded reference cohort H. This is achieved by applying of the law of total probability to likelihoods of all mosaics among haplotypes in the graph which produce the same node sequence as h. We

[1] We make the simplifying assumption that from a specific position on a specific haplotype, there exits only one site per haplotype to which to recombine. Our model admits extension to a more permissive class of recombinations.

call this set $\chi(h)$. We have

$$P(h|G,H) = \sum_{x \in \chi(h)} P(x|G,H)$$

with $P(x|G,H) = p_{recomb}^{R(x)} p_{continue}^{|h|-R(x)}$. This can be simplified letting $\rho = \frac{p_{continue}}{p_{recomb}}$ such that

$$P(h|G,H) = p_{continue}^{|h|} \sum_{x \in \chi(h)} \rho^{-R(x)}$$

and therefore $P(h|G,H)$ is proportional to a $\rho^{-R(x)}$-weighted enumeration of all $x \in \chi(h)$. This statement as an enumeration problem suggests that a division of $\chi(h)$ into equivalence classes may allow efficient computation.

The number of terms in this sum is at worst proportional to the product of the number of haplotypes passing through each node of h, therefore computational complexity for naive summation is $O(|H|^{|h|})$ in the worst case. We will demonstrate that with careful factoring of the paths in H, we can make this sum $O(mn)$ for $n \leq |h|$ and $m \leq n \leq |h|$. We will show preliminary results which suggest that m may be bounded with respect to $|h|$.

3.5 Efficient subwalk counting using the *rectangular decomposition*

Before describing our algorithm to compute the likelihood of a haplotype, we must establish a preliminary result. In this section, we show that the problem of enumerating consistent subhaplotypes can be reduced to operations supported by the gPBWT. By a *consistent subhaplotype*, we mean a subsequence of nodes over which two or more haplotypes are identical. This concept is important for the likelihood function because consistent subhaplotypes represent the

set of possible subsegments for constructing a haplotype as a patchwork of previously observed haplotypes.

Note that two haplotypes contribute an equivalent set of possible subsegments to walks in $\chi(h)$ if and only if they they are consistent with h on the same set of subintervals. Thus, we can divide $\chi(h)$ into equivalence classes based this criterion. It is also evident that two haplotypes are consistent on some interval if and only if they are consistent on all of its subintervals. Therefore, if we can construct the maximal intervals on which $g \in H$ are consistent with h, we can construct all mosaics of containing subhaplotypes of g which have the same node sequence as h. Novak et al.'s gPBWT allows us to find these intervals in linear time with respect to the length of the interval in nodes. The following concepts are necessary to explain how this is achieved:

Definition 18. Given nodes $a, b \in h$ such that a precedes b over the sequence of h, define

$$S_b^a := \{g \in H \,|\, g \text{ is consistent with } h \text{ over the interval from } a \text{ to } b, \text{ but is not consistent at the node before a}\}$$

We call $\{S_b^a \,|\, a \leq b; a, b \in h\}$ the *rectangular decomposition* of H with respect to h.

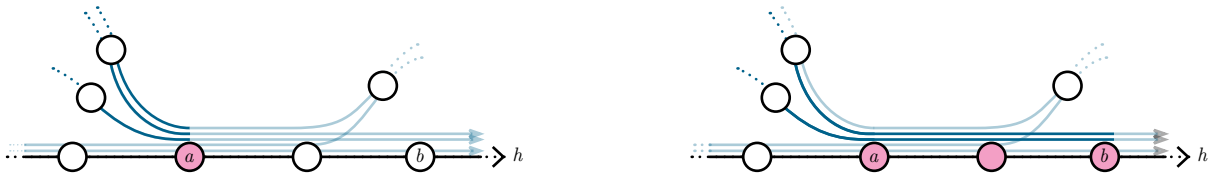


Figure 3.4: In dark blue, top: S_a^a; bottom: S_b^a

Since we are simply performing a weighted enumeration of the sequences, it suffices

to determine the sizes of each equivalence class without explicitly determining membership. The sizes of these equivalence classes can be directly calculated trivially from the sizes of the intervals used to encode haplotypes in Novak's graph positional Burrows-Wheeler transform index (gPBWT).

Definition 19. Define $I_b^a := |S_b^a|$

Definition 20. $A := \{a \in h \,|\, \text{haplotypes leave before } a \text{ or join at } a\}$ and $A_{curr}^b := \{a \in A \,|\, I_b^a \neq 0\}$

A comprises the nodes which are boundaries of maximal subhaplotypes consistent with h. A can be constructed in $O(|h|)$ time using the gPBWT. This is detailed in the appendix. To calculate the I_b^a's we define

Definition 21. $J_b^a := |\{g \in H \,|\, g \text{ is consistent with } h \text{ over the interval from } a \text{ to } b\}|$

and note that the sequence $J_a^a, J_{a+1}^a, \ldots, J_{b-1}^a, J_b^a$ admits $O(b-a)$ query from the gPBWT. The I_b^a's may then be computed by $I_b^a = J_b^a - J_b^{a-1}$. The full algorithms may be found in the appendix.

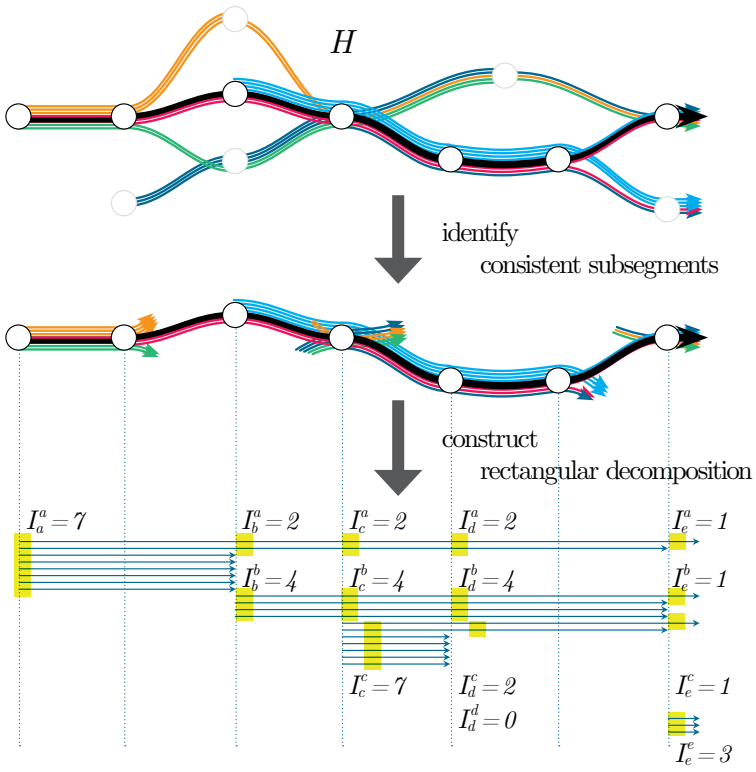


Figure 3.5: In order to characterize $\chi(x)$ we only need to consider subsegments of $g \in H$ which are consistent with the path of h (thick black arrow); these can be divided into equivalence classes for rapid enumeration.

3.6 An algorithm for likelihood computation

Now that we have divided H into equivalence classes which contribute the same subintervals to mosaics in $\chi(x)$ it remains to use representatives of each equivalence class to compute

$$\sum_{x \in \chi(h)} \rho^{-R(x)} \propto P(h|G,H)$$

We achieve this by iterating along the nodes of A (the boundaries of maximal consistent sub-haplotypes) and calculating the likelihoods of increasing prefixes of the query haplotype until the full sequence is reached, at which point ths process yields the desired likelihood. At each successive node b in A, we produce an A^b_{curr}-indexed vector $\overrightarrow{R_b}$ of likelihood contributions at b of mosaics ending with subintervals of haplotypes which started at each $a \in A^b_{curr}$. We claim the following properties of this likelihood contribution vector $\overrightarrow{R_b}$:

Claim 3.

1. $\overrightarrow{R_b}$ is calculable from $3|A^{b-1}_{curr}|+2$ or fewer data points, namely:
   (a) $\overrightarrow{R_{b-1}}$ ($|A^{b-1}_{curr}|$ data points)
   (b) rectangle width (1 data point)
   (c) $I^0_{b-1},\ldots,I^{b-1}_{b-1}$ ($|A^{b-1}_{curr}|$ data points)
   (d) $I^0_b,\ldots,I^b_b$ ($\leq |A^{b-1}_{curr}|$ data points)
2. Calculation of $\overrightarrow{R_b}$ is $O(|A^{b-1}_{curr}|)$
3. $\sum_{a \in A_{curr}} I^a_{end} R_{end}(a) = \sum_{x \in \chi(h)} \rho^{-R(x)} \propto P(h|G,H)$

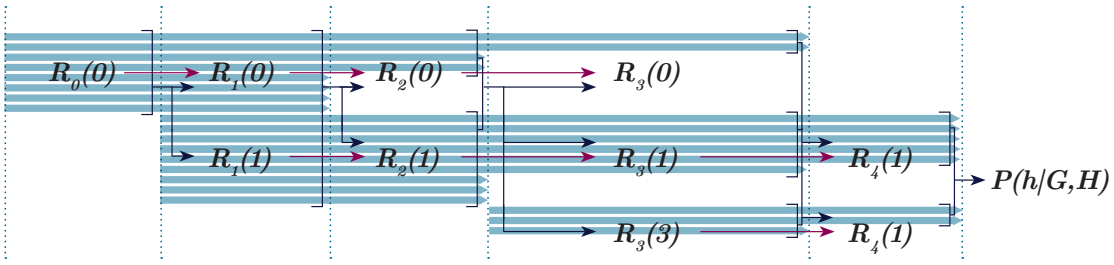


Figure 3.6: **A sketch of flow of information in the calculation described** Blue arrows are the *rectangular decomposition*, $R.(\cdot)$ are prefix likelihoods

This relies on equivalence of likelihood contributions from the sets of mosaics ending in subintervals from haplotypes if these haplotypes belong to the same S_b^a. The following lemma, whose proof is deferred to the appendix, validates this assumption:

Lemma 14. For any $b \in A, a \leq b$ and $\forall g_1, g_2 \in S_b^a$, the sets of partial walks through haplotypes ending at b which end on g_1 and g_2 respectively are in bijective correspondence.

To determine the exact formula for $\overrightarrow{R_b}$ given $\overrightarrow{R_{b-1}}$ we apply this lemma and the single path recombination probability $P(x|G,H) = p_{recomb}^{R(x)} p_{continue}^{|h|-R(x)}$. A series of calculatiosn shows that for any $g \in S_{b-1}^{b-1}$

$$R_b(g) = \mathbf{1}(g \notin S_b^b) R_{b-1}(g)(1-\rho)F_t(\ell) + \mathbf{1}(g \notin S_b^b)\Big[(1-\rho)\frac{F_s(h,\ell)-F_t(\ell)}{h}\Big]S_1 + \rho F_s(h,\ell)(S_1+S_2) \tag{3.1}$$

where F_i's are the polynomials $F_s[h,\ell] := (1+(h-1)\rho)^{\ell-1}$ and $F_t[\ell] := (1-\rho)^{\ell-1}$ (whose logarithms take an especially simple form and are calculable in constant time) and S_1, S_2 are the

$|A^{b-1}_{curr}|$ element sums

$$S_1 := \sum_{a<b} I^a_b R_{b-1}(a), \qquad S_2 := \sum_{a<b} [I^a_{b-1} - I^a_b] R_{b-1}(a)$$

Proof that this satisfies our desired criteria and derivation of this formula for R_b is arithmetically dense but straightforward and combinatorially elegant. It may be found in the supplement but has been ommitted here due the density of the arithmetic.

Inspection of the formula above will convince the reader that calculating $P(h|G,H)$ using this algorithm admits a $O(n \cdot m)$ implementation, for $n = |A|$ and $m = average(|A|_{curr})$.

3.7 Implementation

The methods described above are implemented in the "haplotypes" branch of the variation graph genomics suite *vg* [35] and may be found at https://github.com/yoheirosen/vg. This implementation depends on xg, [89] the gPBWT-based succinct graph indexing component of this suite.

In this implementation, a population reference cohort graph is constructed from sequence data, and the reference cohort haplotypes are embedded and indexed within its gPBWT. At this point, our implementation can calculate the likelihood of any arbitrary sequence so long as it can be embedded in the graph as well. We present results using this implementation in the demonstrations below.

3.8 Experiment: The distribution of $average(|A|_{curr})$

Seeing as our algorithm is $O(n \cdot m)$ for $n = |A|$ and $m = average(|A|_{curr})$, it is important to determine the dependence of this quantity both on sequence length and reference cohort size in real sequence data.

To evaluate dependence of $average(|A|_{curr})$ on haplotype length, we generated random 10–1000000 node subintervals of haplotypes embedded a 2504 individual variation graph from *1000 Genomes* [23] data.

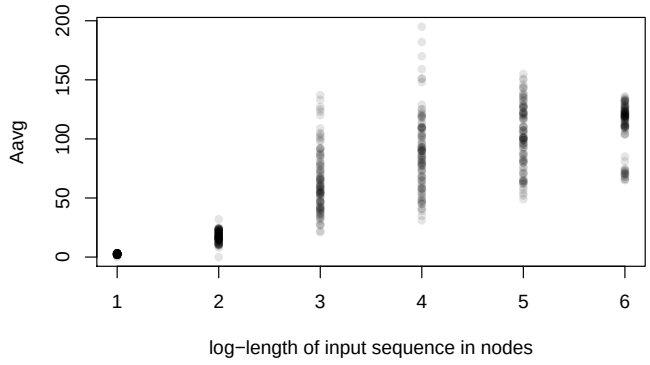


Figure 3.7: $average(|A|_{curr})$ **vs.** $|h|$ Note sublogarithmic growth of $average(|A|_{curr})$ and decrease in variance for large $|h|$

To determine the effect of reference cohort size on $m = average(|A|_{curr})$, we used *vcftools* [24] to generate different sized random subpopulations of the whole *1000 Genomes* dataset. Variation graphs were constructed as previously described. This was done in five replicates at each of four population sizes, the data from the previous experiment provided a fifth population size. The same length analysis was done, using a total of 100 replicates per

length per population size.

$Average(|A|_{curr})$ shows an approximately logarithmic relationship between both the mean and variance of this value as sequence length increases from 100 to 10,000 nodes. After this point, as sequence length increases to 1,000,000, the growth of $average(|A|_{curr})$ with respect to sequence length slows, and its variance decreases. Note that in our graphs, chromosome 22 measures approximately 1,600,000 nodes in length.

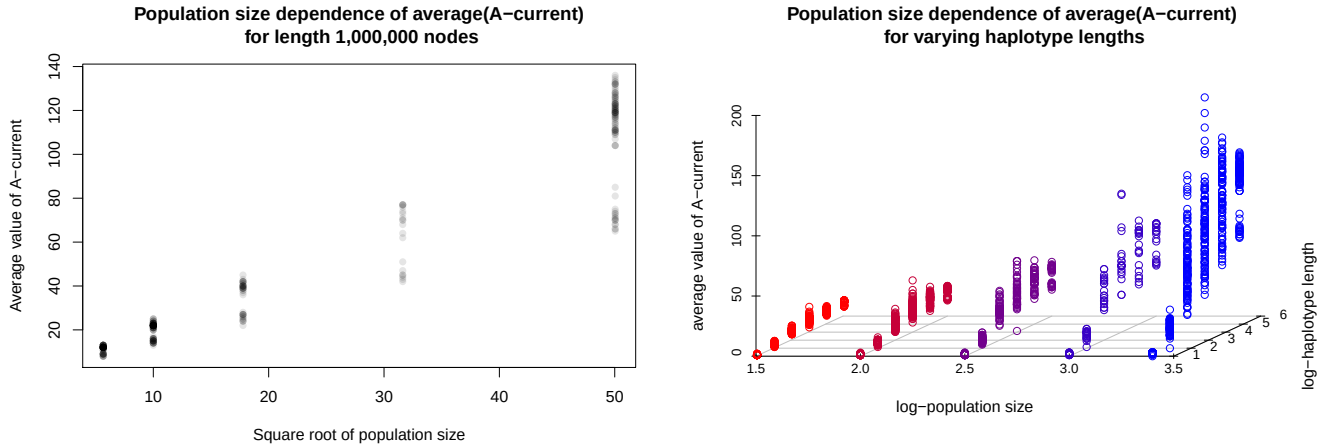


Figure 3.8: (left): $average(|A|_{curr})$ shows growth approximately proportional to $|H|^{\frac{1}{2}}$ (right): Relationship between $average(|A|_{curr})$ and $|h|$ is consistent over varying population size.

$Average(|A|_{curr})$ shows a sublinear, supralogarithmic dependence of on reference cohort size, a relationship which appears consistent across sequence length. The growth of $average(|A|_{curr})$ appears to be approximately $O(n^{-\frac{1}{2}})$ in reference cohort size.

These results are for Chromosome 22 in the 1000 Genomes dataset only and we expect variation according to region and indeed organism. However, extrapolation of these results is promising in that $average(|A|_{curr})$ may be in general bounded with respect to sequence length

up to full chromosome length.

3.9 Experiment: read likelihood with respect a reference cohort

This model was in part motivated by a forthcoming haplotype aware graph-based variant caller. To evaluate its use in such a role, we used our model to assess the likelihood of mapped reads, considering each as a query path equivalent to a short haplotype.

Very short haplotypes are also a useful indicator of the model's sensitivity to individual recombination events or rare variants, since these should be few in such short sequences. Sets of two recombinations relative to the reference cohort should be rare in short reads, however the current model (without mutations) may employ a double recombination to explain a substitution.

We used the 1000 Genomes low coverage reads for individual NA12878 on chromosome 22 and mapped them against a variation graph built from the full 1000 Genomes [23] cohort built as described in section 6. Reads were mapped with vg. Read likelihoods were then computed with the implementation of our model in *vg*. An arbitrary value of 10^{-9} was used for p_{recomb}. All reported relative log-likelihood scores are the log-ratio of $P(h|G,H)$ the likelihood of a theoretical sequence of the same length which is everywhere consistent with all individuals.

We simulated reads by randomly generating paths through the graph without regard to haplotype distributions. These were used as a control against which to compare the distribution of likelihood generated from actual reads.

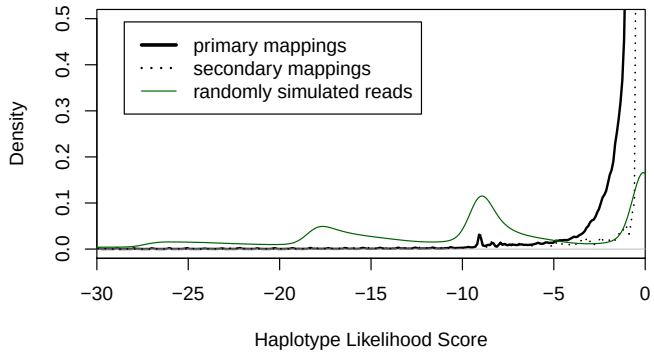


Figure 3.9: **Density plot of relative log-likelihood of mapped reads versus randomly generated simulated haplotypes**

We also used heuristics to determine subsets of reads which involved one or more likely recombination relative to the majority of the reference cohort, or one or more variant which was poorly represented in the reference cohort.

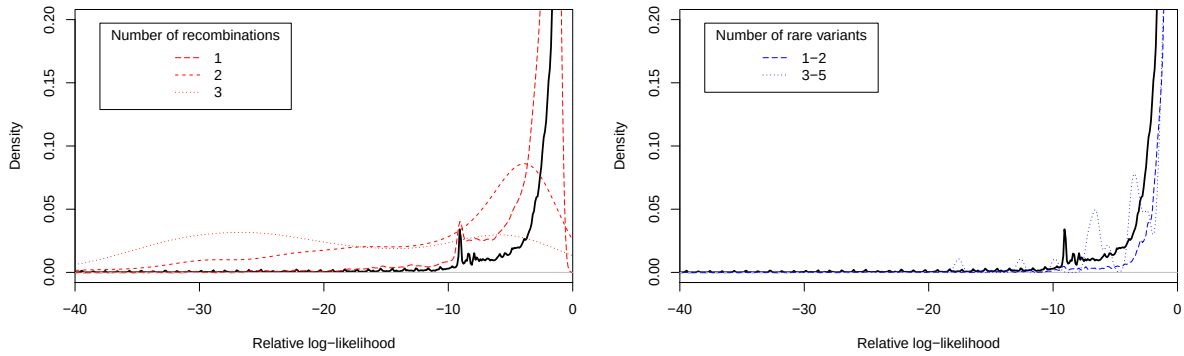


Figure 3.10: Left: density plot of relative log-likelihood of all mapped reads (black line) vs reads containing recombinations. Right: density plot of relative log-likelihood of all mapped reads (black line) vs reads containing variants present at under 5% prevalence.

In the likelihood vs number of structural events assessment, we note that increasing number of recombinations is associated with lower likelihoods. We see a similar trend with increasing numbers of "rare" or variants. For one, two, and three recombinations we see density peaks centered near $n \cdot p_{recomb}$ (ie, $10^{-9}, 10^{-18}, 10^{-27}$). However, we observe dominating peaks at higher log-likelihoods. Reads containing 3 or more variants poorly represented in the reference cohort are also associated with lower likelihoods. These findings suggest that poorly represented structural motifs or variants do produce lower likelihoods as one might expect from our model.

In the mapped read vs control assessment, we see that the density of read likelihoods from NA12878 is distributed towards higher likelihoods than the simulated control. This is the result expected, since they should generally more closely reflect the limited set of possible paths through the graph followed by haplotypes. We also see clustering of likelihoods of the randomly generated control "reads" around factors of $10^{-9} = p_{recomb}$ as would be expected: essentially, at each node, the simulated paths are as likely to take one haplotype as any other, which should produce a high recombination density.

3.10 Demonstration: Comparing chromosome-length haplotypes from a population against a reference

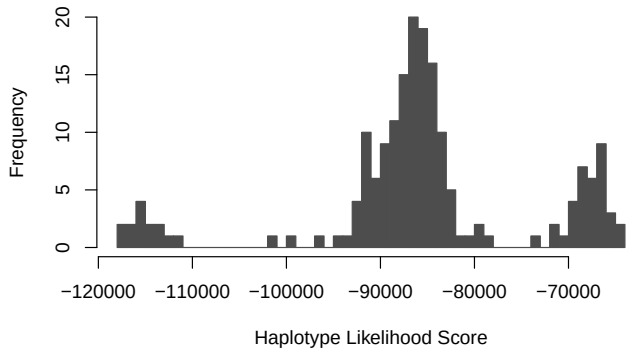


Figure 3.11: Log-relative likelihods of 201 randomly selected individuals with respect to the maximum possible likelihood for a sequence of the same length, over the length of Chromosome 22

Complementary to our short-read assessment, the structural flexibility of our model also allows us to calculate meaningful likelihoods for arbitrarily long query haplotypes. To demonstrate this, we calculated the likelihoods of individual members of the *1000 Genomes* dataset with respect to the others. We observe multiple clusterings of likelihoods as one might expect from a cohort wherein individuals diverge in sequence to varying degrees.

3.11 Conclusions

We have introduced a method of representing the shared sequence relationship of an individual haplotype to a reference cohort represented in a variation graph. We also demon-

strated an efficient algorithm based on the gPBWT. We used 5008 haplotypes from the 1000 Genomes data set to show that this algorithm can efficiently model recombination with respect to both long sequences and large reference cohorts.

We do not model mutation. This would depend on our ability to efficiently calculate, for a given observed sequence, the subgraph of our variation graph containing all haplotypes which could produce this sequence via a reasonable set of mutations. If this is made possible, there exist mathematical generalizations of our algorithm which allow us to simultaneously compute the likelihood of haplotypes which are similar with respect to their embedding in the graph. Graph decomposition algorithms which allow us to efficiently identify these related haplotypes are a subject of our continuing research. [98]

3.12 Appendix A: An $O(n \cdot m)$ implementation of the rectangular decomposition construction

In brief, in the gPBWT, indexing information for haplotypes is stored in such a way that finding subhaplotypes consistent with a query sequence h of nodes involves starting a search in the gPBWT index using operation STARTSEARCHATNODE($Node$) on the first node of h, which returns a search interval $gPBWTInt$ of a form analogous to the search interval of a Burrows Wheeler Transform based index of sequences. This search interval is extended by calling an operation STARTSEARCHATNODE($gPBWTInt, Node$) to extend this search with each additional node in h. Finally, this search interval can be converted into a count of matching subhaplotypes using a function COUNT(gPBWTInt). It is shown in (Novak 2016) that START-

SEARCHATNODE, EXTEND and COUNT all admit $O(1)$ implementations.

It is evident that this search process yields a function COUNTHAPLOTYPEMATCHES(h) which is $O(n)$ in the length $|h|$ of h in nodes. Let $h_1h_2h_3\ldots h_{|h|-1}h_{|h|}$ denote the node sequence of h. Using COUNTHAPLOTYPEMATCHES we can identify the set A in $O(n)$ independent length-2 subhaplotype count queries:

Algorithm 4: Identifying A, the set of "relevant" nodes

[1] BuildA$h, B[]$ $A \leftarrow [1]$ $ht_{prev} \leftarrow |B[h_1]|$ **for** $i = 2,\ldots,|h|$ **do**

$ht_{new} \leftarrow |B[h_i]|$ $J_i^{i-1} \leftarrow$ CountHaplotypeMatches$h_{i-1}h_i$ **if** $J_i^{i-1} < ht_{new}$ *or* $ht_{prev} > ht_{new}$ **then**

AppendA, i $ht_{prev} \leftarrow ht_{new}$

Given construction of A we can determine the rest of the rectangular decomposition and all of the J-values according to the following algorithm:

Algorithm 5: Building the J's and A_{curr}'s

[1] BuildJs$h, B[]$ $J_1^1 \leftarrow |B[h_1]|$ $A_{curr}^1 \leftarrow 1$ **for** $i \in A$ **do**

$A_{curr}^i \leftarrow []$ **if** $|B[h_i]| > J_i^i$ **then**

AppendA_{curr}^i, i $S_i \leftarrow$ StartSearchh_i **for** $j \in A_{curr}^{i-1}$ **do**

$S_j \leftarrow$ ExtendS_j, h_i **if** $CountS_j \neq 0$ **then**

$J_i^j \leftarrow$ CountS_j AppendA_{curr}^i, j **else**

break

3.13 Appendix B: Arithmetic for derivation of Equation 3.1

Here we lay out the arithmetic to derive Equation 3.1, which is used in our iterative computation of likelihood of a haplotype h with respect to a population reference cohort H

embedded in a variation graph G. The reasoning is straighforward but involves many subcases which require care.

3.13.1 Notation

Definition 22. A haplotype is a sequence of nodes $n_1 \to \cdots \to n_{|h|}$ in a variation graph. The base sequence of a haplotype is the sequence of DNA bases spelled by its node labels. A haplotype subinterval is a contiguous subsequence of a haplotype. A haplotype base sequence subinterval is analogously defined. Denote by $|h|$ the length of a haplotype base sequence in base pairs.

Definition 23. Haplotypes h, h' are consistent if $|h| = |h'|$ and $n_i = n'_i \, \forall i$.

Definition 24. A mosaic of haplotypes x consistent with h is a vector $\langle x_{(i)} \rangle$ of subintervals of base sequences of haplotypes in H whose concatenation is consistent with the base sequence of h. The recombination count $R(x)$ is one less than the number of elements in $\langle x_{(i)} \rangle$. NB: defining these in terms of base sequence rather than node subintervals permits recombination within nodes. Recall Figure 3 from the main text.

Definition 25. $\chi(h)$ is the set of all mosaics x consistent with h. $\chi(h)^R$ is the subset with $R(x) = R$. $\chi(h)[,g]$ is the subset whose final subinterval is a subinterval of g. $\chi(h)[g,]$ is that with initial subinterval a subinterval of g. $|\chi(h)|$ is the number of elements in $\chi(h)$.

3.13.2 Arithmetic shortcuts

Lemma 15. There exists a partition of h into subintervals $h_1, h_2, \ldots, h_n$ such that if a haplotype $g \in H$ has a subinterval consistent with a subinterval of h_i then it has a subinterval consistent with all of h_i.

Proof. It is straightforward to verify that the intervals between successive nodes in the set A described in the main text produce such a partition of h. □

This is important because we will show that it is simple to calculate $|\chi(h_i)|$ within any interval with this property.

The following is a more notationally precise statement of Lemma 1 from the main text:

Lemma 16. For any $b \in A, a \leq b$, given that f and g are members of the same equivalence class S_b^a of haplotypes, the haplotype mosaics $\chi^R(h_{[0,b]})[,f]$ and $\chi^R(h_{[0,b]})[,g]$ consistent with the subinterval $h_{[0,b]}$ and ending with subintervals of f and g are in bijective correspondence.

Proof. We assume that $g \neq f$ else this is trivial. Consider any mosaic x in $\chi^R(h_{[0,b]})[,f]$. Given $x = \langle x_1, x_2, \ldots, x_{R+1} \rangle$, let $j = \max\{i \in 1, \ldots, R | x_i \text{ is not a subinterval of } g \text{ or } f\}$. We will construct a mosaic $y = \langle y_1, y_2, \ldots, y_{R+1} \rangle$ such that for all $i \leq j$, $y_i = x_i$, and for all $i > j$, y_i is the subinterval of the same length as x_i but derived from the opposite haplotype of the pair f, g.

The concatenation $y_1 y_2 \cdots y_{R+1}$ is consistent with $h_{[0,b]}$ since given that both $f, g \in S_b^a$, the first node of y_{j+1} must be at or after a. Therefore clearly $y_i \in \chi^R(h_{[0,b]})[,g]$ since its final subinterval corresponds to g. The inherent invertibility of this transformation proves that it is a bijection. □

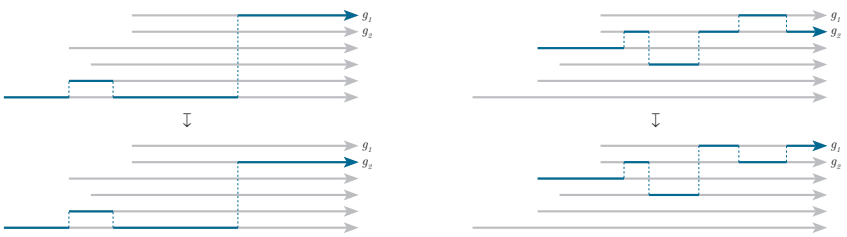


Figure 3.12: Visual proof of the above lemma by explicit construction of the bijection involved

Lemma 17. Suppose that h_i is a subinterval of h such that if a haplotype $g \in H$ has a subinterval consistent with a subinterval of h_i then it has a subinterval consistent with all of h_i. Then suppose that $f_1, f_2, g \in H$, and all have subintervals consistent with all of h_i. Then for all $R < |h_i|$ there is a bijection between $\chi(h_i)[f_1, g]$ and $\chi(h_i)[f_2, g]$.

Proof. The proof imitates that of the previous lemma. □

3.13.3 The case of a single simple interval h_i

Suppose that h_i is an interval of the form in Lemma 2, ℓ base pairs in in length, and has subintervals of ht haplotypes of H consistent with it. Consider $f, g \in H$ such that both have subintervals consistent with h_i. Suppose that we wish to calculate, for some $R < \ell$, the number $|\chi^R(h)[, g]|$ of mosaics consistent with h_i having R recombinations and ending with haplotype g. To calculate $|\chi(h_i)|$ within an interval of the form above, we need only calculate

1. $|\chi^R(h_i)[g, g]|$, the number of paths both beginning on and ending on g and

2. $|\chi^R(h_i)[f, g]|$, the number of paths beginning on $f \neq g$ and ending on g, which, by lemma

4, is the same for all such f.

Consider $R = \ell - 1$. It is clear that

$$\sum_j |\chi^R(h_i)[j,g]| = (ht-1)^R \tag{3.2}$$

Lemma 4 tells us that all haplotypes $f \neq g$ are equivalent for the purposes of enumeration, therefore we write $\neg g$ to denote any arbitrary representative $f \neq g$. There are $ht-1$ such haplotypes.

$$|\chi^R(h_i)[g,g]| + (ht-1)|\chi^R(h_i)[\neg g,g]| = (ht-1)^R \tag{3.3}$$

We begin by calculating $|\chi^R(h_i)[\neg g,g]|$. Consider first $\ell = R+1 = 1$, for which, given the lack of possible recombinations, $|\chi^R(h_i)[\neg g,g]| = 0$. For $\ell = R+1 = 2$, any $x \in \chi^R(h_i)[\neg g,g]$ must at its second node visit a haplotype which is neither g nor the $\neg g$ under consideration, therefore $|\chi^R(h_i)[\neg g,g]| = (ht-2)$. Suppose now that, for arbitrary $\ell = R$, we know $|\chi^R(h_i)[\neg g,g]|$. Then, counting the $(ht-1)$ possible haplotypes before finally recombining to g shows us that

$$|\chi^{R+1}(h_i)[g,g]| = (ht-1)|\chi^R(h_i)[\neg g,g]| \tag{3.4}$$

By (3.3), we know that

$$|\chi^{R+1}(h_i)[\neg g,g]| = \frac{(ht-1)^R - |\chi^{R+1}(h_i)[g,g]|}{(ht-1)}$$

Which by (3.4) implies

$$|\chi^{R+1}(h_i)[\neg g,g]| = \frac{(ht-1)^R - (ht-1)|\chi^R(h_i)[\neg g,g]|}{(ht-1)}$$

$$\implies |\chi^{R+1}(h_i)[\neg g,g]| = (ht-1)^{R-1} - |\chi^R(h_i)[\neg g,g]| \tag{3.5}$$

Using (3.5) as the induction step with base case $\ell = R+1 = 2$ we find that $\forall \ell = R+1 \geq 2$

$$|\chi^R(h_i)[\neg g,g]| = \frac{(ht-1)^{R-1} + (-1)^R}{ht}$$

We now relax the restriction that $R = \ell$. For given $R < \ell$ each subset of nodes at which recombinations happen will define an additional set of possible recombinations. Counting all possible such subsets

$$|\chi^R(h_i)[\neg g,g]| = \binom{\ell-1}{R} \frac{(ht-1)^{R-1} + (-1)^R}{ht}$$

and

$$|\chi^R(h_i)[g,g]| = (ht-1)|\chi^R(h_i)[\neg g,g]|$$

.

3.13.4 Extending a computation for a prefix by a simple subinterval h_i

To extend our ability to calculate $|\chi(h)|$ beyond the single interval h_i, suppose we have a partition $\{h_1, h_2, \ldots, h_n\}$ of h into subintervals of the form in Lemma 2. Let $b \in A$ such that b is a node on the boundary of such an interval, let $h_{[0,b-1]}$ be the prefix of h formed by concatenation of the subintervals preceding node b, and let $h_{[b-1,b]}$ be the subinterval beginning with node b.

Suppose now that we have calculated each $|\chi^R(h_{[0,b-1]})[,f]|$ and now wish to calculate these values up to b, the node in A succeeding $b-1$. By Lemma 2, the intervening sequence $h_{[b-1,b]}$ is of the form for which we have just calculated $|\chi^R(h)[g,g]|$ and $\chi^R(h)[\neg g,g]|$. We divide this into cases.

Case 1: Suppose that f has no subinterval consistent with $h_{[b,b+1]}$, that is, $f \in S^a_{b-1}$ for some a but $f \notin S^a_b$. Then any mosaic extending any mosaic in $\chi^R(h_{[0,b]})[,f]$ must recombine. Since $f \notin S^a_b$, there are $ht := J^b_b$ possible haplotypes to which this recombination at $b-1 \to b$ may occur. Let ℓ be the length (in base pairs) of the interval $b-1$ to b, then $\forall R' < \ell(b)$ we have previously calculated in (3.3) that

$$|\chi^{R'}(h_{[b-1,b]})[,g]| = \binom{\ell-1}{R'}(ht-1)^{R'-1}$$

and therefore, where we write $\chi^R(h_{[0,b-1]})[,f] \oslash \chi^{R'}(h_{[b-1,b]})[,g]$ for the set of mosaics formed by continuing mosaics in $\chi^R(h_{[0,b-1]})[,f]$ such that they recombine between $h_{[0,b-1]}$ and $h_{[b-1,b]}$ and end with a subinterval of g,

$$|\chi^R(h_{[0,b-1]})[,f] \oslash \chi^{R'}(h_{[b-1,b]})[,g]| = |\chi^R(h_{[0,b-1]})[,f]| \binom{\ell-1}{R'} (ht-1)^{R'-1} \tag{3.6}$$

Case 2: Suppose now that we know $|\chi^R(h_{[0,b-1]})[,f]|$, and $f \in S_b^a$ for some a, that is, f does have a subinterval consistent with $h_{[b-1,b]}$

There are two subcases: either there is, or there is not a recombination between the last base in $h_{[0,b-1]}$ and the subsequent base at the beginning of $h_{[b-1,b]}$. Suppose that there is not. In this case, where we write $\chi^R(h_{[0,b-1]})[,f] \ominus \chi^{R'}(h_{[b-1,b]})[,g]$ for the set of mosaics formed by continuing mosaics in $\chi^R(h_{[0,b-1]})[,f]$ such that they do not recombine between $h_{[0,b-1]}$ and $h_{[b-1,b]}$ and such that they do end with a subinterval of g,

$$|\chi^R(h_{[0,b-1]})[,f] \ominus \chi^{R'}(h_{[b-1,b]})[,g]| = |\chi^R(h_{[0,b-1]})[,f]||\chi^{R'}(h_{[b-1,b]})[f,g]| \tag{3.7}$$

such that if $f \neq g$

$$|\chi^R(h_{[0,b-1]})[,f] \ominus \chi^{R'}(h_{[b-1,b]})[,g]| = |\chi^R(h_{[0,b-1]})[,f]||\chi^{R'}(h_{[b-1,b]})[\neg g,g]| \tag{3.8}$$

else

$$|\chi^R(h_{[0,b-1]})[,f] \ominus \chi^{R'}(h_{[b-1,b]})[,g]| = |\chi^R(h_{[0,b-1]})[,f]||\chi^{R'}(h_{[b-1,b]})[g,g]| \tag{3.9}$$

The other subcase is that there is a recombination between the last base in $h_{[0,b-1]}$ and the subsequent base at the beginning of $h_{[b-1,b]}$. In this case if $f \neq g$,

$$|\chi^R(h_{[0,b-1]})[,f] \oslash \chi^{R'}(h_{[b-1,b]})[,g]| = |\chi^R(h_{[0,b-1]})[,f] \oslash \chi^{R'}(h_{[b-1,b]})[g,g]| \quad + \sum_{f' \neq f,g} |\chi^R(h_{[0,b-1]})[,f] \oslash \chi^{R'}(h_{[b-1,b]})[f,g]| \tag{3.10}$$

$$= |\chi^R(h_{[0,b-1]})[,f]||\chi^{R'}(h_{[b-1,b]})[g,g]|| \quad + \underbrace{(ht-2)|\chi^R(h_{[0,b-1]})[,f]||\chi^{R'}(h_{[b-1,b]})[\neg g,g]|}_{\text{by Lemma 4}} \tag{3.11}$$

else

$$|\chi^R(h_{[0,b-1]})[,f] \oslash \chi^{R'}(h_{[b-1,b]})[,g]| = (ht-1)|\chi^R(h_{[0,b-1]})[,f]||\chi^{R'}(h_{[b-1,b]})[\neg g,g]| \tag{3.12}$$

3.13.5 Deriving the Formula for $P(h|G,H)$

Suppose that we have calculated $|\chi(h_{[0,b-1]})[,f]|$ for all f and now wish to calculate $|\chi(h_{[0,b]})[,g]|$ for some $g \in S_b^a$, for some $a \leq b$.

Note that as defined in the main text, $R_{b-1}(a) = |\chi(h_{[0,b-1]})[,f]|$ for the a such that $f \in S_b^a$. This means this calculatiuon will in fact give us the formula with which to calculate $\overrightarrow{R_b}$ given $\overrightarrow{R_{b-1}}$. Let us write $R_b(f)$ for $R_b(a)$ such that $f \in S_b^a$.

Accounting for all prefixes in $\chi(h_{[0,b-1]})$ which can produce mosaics in $\chi(h_{[0,b]})[,g]$, then

$$|\chi(h_{[0,b]})[,g]| = \sum_{\substack{R_1<|h_{[0,b-1]}| \\ R_2<|h_{[b-1,b]}|}} \rho^{(R_1+R_2)} |\chi^{R_1}(h_{[0,b-1]}) \ominus \chi^{R_2}(h_{[b-1,b]})[,g]| \; + \sum_{\substack{R_1<|h_{[0,b-1]}| \\ R_2<|h_{[b-1,b]}|}} \rho^{(R_1+R_2+1)} |\chi^{R_1}(h_{[0,b-1]}) \oslash \chi^{R_2}(h_{[b-1,b]})[,g]|$$

$$= \sum_{\substack{R_1<|h_{[0,b-1]}| \\ R_2<|h_{[b-1,b]}|}} \Big(\rho^{(R_1+R_2)} |\chi^{R_1}(h_{[0,b-1]}) \ominus \chi^{R_2}(h_{[b-1,b]})[,g]| \; + \rho^{(R_1+R_2+1)} |\chi^{R_1}(h_{[0,b-1]}) \oslash \chi^{R_2}(h_{[b-1,b]})[,g]| \Big)$$

$$= \sum_{\substack{R_1<|h_{[0,b-1]}| \\ R_2<|h_{[b-1,b]}|}} \rho^{(R_1+R_2)} \Bigg(\sum_{a<b} \sum_{\substack{f\in S_b^a \\ f\neq g}} |\chi^{R_1}(h_{[0,b-1]})[,f] \ominus \chi^{R_2}(h_{[b-1,b]})[f,g]|$$
$$+ |\chi^{R_1}(h_{[0,b-1]})[,g] \ominus \chi^{R_2}(h_{[b-1,b]})[g,g]|$$
$$+ \rho \Big(|\chi^{R_1}(h_{[0,b-1]})[,g] \oslash \chi^{R_2}(h_{[b-1,b]})[,g]|$$
$$+ \sum_{a<b} \sum_{\substack{f\in S_b^a \\ f\neq g}} |\chi^{R_1}(h_{[0,b-1]})[,f] \oslash \chi^{R_2}(h_{[b-1,b]})[,g]|$$
$$+ \sum_{a<b} \sum_{f\notin S_b^a} |\chi^{R_1}(h_{[0,b-1]})[,f] \oslash \chi^{R_2}(h_{[b-1,b]})[,g]| \Big) \Bigg)$$

$$= \sum_{\substack{R_1<|h_{[0,b-1]}| \\ R_2<|h_{[b-1,b]}|}} \rho^{(R_1+R_2)} \Bigg(\sum_{a<b} \sum_{\substack{f\in S_b^a \\ f\neq g}} \underbrace{|\chi^{R_1}(h_{[0,b-1]})[,f]||\chi^{R_2}(h_{[b-1,b]})[\neg g,g]|}_{\text{by (3.8)}}$$
$$+ \underbrace{|\chi^{R_1}(h_{[0,b-1]})[,g]||\chi^{R_2}(h_{[b-1,b]})[g,g]|}_{\text{by (3.9)}}$$
$$+ \rho \Big(\underbrace{|\chi^{R_1}(h_{[0,b-1]})[,g]|(ht-1)|\chi^{R_2}(h_{[b-1,b]})[\neg g,g]|}_{by(3.12)}$$
$$+ \sum_{a<b} \sum_{\substack{f\in S_b^a \\ f\neq g}} \underbrace{|\chi^{R_1}(h_{[0,b-1]})[,f]|\big(|\chi^{R_2}(h_{[b-1,b]})[g,g]| + (ht-2)|\chi^{R_2}(h_{[b-1,b]})[\neg g,g]|\big)}_{\text{by (3.11)}}$$
$$+ \sum_{a<b} \sum_{f\notin S_b^a} |\chi^{R_1}(h_{[0,b-1]})[,f]||\chi^{R_2}(h_{[b-1,b]})[,g]| \Big) \Bigg)$$

$$= \sum_{\substack{R_1<|h_{[0,b-1]}| \\ R_2<|h_{[b-1,b]}|}} \rho^{(R_1+R_2)} \Bigg(\sum_{a<b} \sum_{f\in S_b^a} |\chi^{R_1}(h_{[0,b-1]})[,f]||\chi^{R_2}(h_{[b-1,b]})[\neg g,g]| - |\chi^{R_1}(h_{[0,b-1]})[,g]||\chi^{R_2}(h_{[b-1,b]})[\neg g,g]|$$

$$+ |\chi^{R_1}(h_{[0,b-1]})[,g]||\chi^{R_2}(h_{[b-1,b]})[g,g]|$$

$$+ \rho \Big(|\chi^{R_1}(h_{[0,b-1]})[,g]| \big(|\chi^{R_2}(h_{[b-1,b]})[,g]| - |\chi^{R_2}(h_{[b-1,b]})[g,g]| \big)$$

$$+ \sum_{a<b} \sum_{\substack{f\in S_b^a \\ f\neq g}} |\chi^{R_1}(h_{[0,b-1]})[,f]| \big(|\chi^{R_2}(h_{[b-1,b]})[,g]| - |\chi^{R_2}(h_{[b-1,b]})[\neg g,g]| \big)$$

$$+ \sum_{a<b} \sum_{f\notin S_b^a} |\chi^{R_1}(h_{[0,b-1]})[,f]||\chi^{R_2}(h_{[b-1,b]})[,g]| \Big) \Bigg)$$

$$= \sum_{\substack{R_1<|h_{[0,b-1]}| \\ R_2<|h_{[b-1,b]}|}} \rho^{(R_1+R_2)} \Bigg(\sum_{a<b} \sum_{f\in S_b^a} |\chi^{R_1}(h_{[0,b-1]})[,f]||\chi^{R_2}(h_{[b-1,b]})[\neg g,g]| - |\chi^{R_1}(h_{[0,b-1]})[,g]||\chi^{R_2}(h_{[b-1,b]})[\neg g,g]|$$

$$+ |\chi^{R_1}(h_{[0,b-1]})[,g]||\chi^{R_2}(h_{[b-1,b]})[g,g]|$$

$$+ \rho \Big(\sum_{a<b} \sum_{f\in S_b^a} |\chi^{R_1}(h_{[0,b-1]})[,f]| \big(|\chi^{R_2}(h_{[b-1,b]})[,g]| - |\chi^{R_2}(h_{[b-1,b]})[\neg g,g]| \big)$$

$$+ |\chi^{R_1}(h_{[0,b-1]})[,g]||\chi^{R_2}(h_{[b-1,b]})[\neg g,g]| - |\chi^{R_1}(h_{[0,b-1]})[,g]||\chi^{R_2}(h_{[b-1,b]})[g,g]|$$

$$+ \sum_{a<b} \sum_{f\notin S_b^a} |\chi^{R_1}(h_{[0,b-1]})[,f]||\chi^{R_2}(h_{[b-1,b]})[,g]| \Big) \Bigg)$$

$$= \sum_{\substack{R_1<|h_{[0,b-1]}| \\ R_2<|h_{[b-1,b]}|}} \rho^{(R_1+R_2)} \Bigg(\Big(1-\rho\Big) \Bigg(\sum_{a<b} \sum_{f \in S_b^a} |\chi^{R_1}(h_{[0,b-1]})[,f]||\chi^{R_2}(h_{[b-1,b]})[\neg g,g]| \\ - |\chi^{R_1}(h_{[0,b-1]})[,g]||\chi^{R_2}(h_{[b-1,b]})[\neg g,g]| + |\chi^{R_1}(h_{[0,b-1]})[,g]||\chi^{R_2}(h_{[b-1,b]})[g,g]| \Bigg) \\ + \rho \sum_{a<b} \sum_{f \in S_{b-1}^a} |\chi^{R_1}(h_{[0,b-1]})[,f]||\chi^{R_2}(h_{[b-1,b]})[,g]| \Bigg)$$

Letting

$$RRSame = \sum_{R_2<|h_{[b-1,b]}|} \rho^{R_2} |\chi^{R_2}(h_{[b-1,b]})[g,g]|,$$

$$RRDiff = \sum_{R_2<|h_{[b-1,b]}|} \rho^{R_2} |\chi^{R_2}(h_{[b-1,b]})[\neg g,g]|$$

(And we note that *RRSame* and $RRDiff$ do not actually depend on choice of g)

$$= \sum_{R_1<|h_{[0,b-1]}|} \rho^{R_1} \Bigg(\Big(1-\rho\Big) \Bigg(\sum_{a<b} \sum_{f \in S_b^a} |\chi^{R_1}(h_{[0,b-1]})[,f]| RRDiff \\ - |\chi^{R_1}(h_{[0,b-1]})[,g]| RRDiff + |\chi^{R_1}(h_{[0,b-1]})[,g]| RRSame \Bigg) \\ + \sum_{R_2<|h_{[b-1,b]}|} \rho^{(R_2+1)} \sum_{a<b} \sum_{f \in S_{b-1}^a} |\chi^{R_1}(h_{[0,b-1]})[,f]| \underbrace{\binom{|h_{[b-1,b]}|-1}{R_2} (ht-1)^{R_2}}_{\text{by (3.2)}} \Bigg)$$

Noting that

$$\sum_{R_1<|h_{[0,b-1]}|} \rho^{R_1} |\chi^{R_1}(h_{[0,b-1]})[,f]| = R_{b-1}(f)$$

Letting:

$$S_1 := \sum_{a<b} \sum_{f \in S_b^a} R_{b-1}(f)$$

$$S_2 := \sum_{a<b} \sum_{f \notin S_b^a} R_{b-1}(f)$$

then the above is equal to

$$\Big(1-\rho\Big)\Big(S_1 RRDiff - R_b(g)\big(RRDiff - RRSame\big)\Big) + \big(S_1+S_2\big) \sum_{R_2<|h_{[b-1,b]}|} \rho^{(R_2+1)} \binom{\ell(b)-1}{R_2} (ht-1)^{R_2}$$

For $g \in S_b^b$, the calculation is similar:

$$\begin{aligned}
|\chi(h_{[0,b]})[,g]| &= \sum_{\substack{R_1<|h_{[0,b-1]}| \\ R_2<|h_{[b-1,b]}|}} \rho^{(R_1+R_2+1)} |\chi^{R_1}(h_{[0,b-1]}) \oslash \chi^{R_2}(h_{[b-1,b]})[,g]| \\
&= \sum_{\substack{R_1<|h_{[0,b-1]}| \\ R_2<|h_{[b-1,b]}|}} \rho^{(R_1+R_2+1)} \left(\sum_{f \in S_b^a} |\chi^{R_1}(h_{[0,b-1]})[,f]||\chi^{R_2}(h_{[b-1,b]})[,g]| \right) \\
&= \big(S_1+S_2\big) \sum_{R_2<|h_{[b-1,b]}|} \rho^{(R_2+1)} \left(\binom{|h_{[b-1,b]}|-1}{R_2} (ht-1)^{R_2} \right)
\end{aligned}$$

We can simplify the sums above by writing

$$RRS(ht,\ell) := \sum_{R_2<\ell} \rho^{R_2}\left(\binom{\ell-1}{R_2}(ht-1)^{R_2}\right)$$
$$= \left(1+(ht-1)\rho\right)^{\ell-1} \tag{3.13}$$

Given a second definition

$$RRT(\ell) := (1-\rho)^{\ell-1} \tag{3.14}$$

we can actually write

$$RRSame - RRDiff = RRT(|h_{[b-1,b]}|)$$
$$RRDiff = \frac{RRS(ht,|h_{[b-1,b]}|) - RRT(|h_{[b-1,b]}|)}{ht}$$

and so finally, we can write our formula for $R_b(g)$ in a compact form as

$$R_b(g) = \begin{cases} (1-\rho)\left(S_1 \frac{RRS(ht,|h_{[b-1,b]}|)-RRT(|h_{[b-1,b]}|)}{ht} + R_{b-1}(g)RRT(|h_{[b-1,b]}|)\right) & \\ \qquad + \rho\left(S_1+S_2\right)RRS(ht,|h_{[b-1,b]}|) & \text{if } g \notin S_b^b \\ \rho\left(S_1+S_2\right)RRS(ht,|h_{[b-1,b]}|) & \text{if } g \in S_b^b \end{cases}$$

[77]

Part III

Methods for automated, complex, scalable tissue culture with continuous multimodality observation

Preamble

The following is currently unpublished work which will eventually be divided into several engineering and biological methods papers for which I will be first author. I designed and manufactured all devices except where otherwise noted, did all experiments except where otherwise noted and did all writing and figure-making. Cell culture for 2D culture experiments, fluidic device maintenance and bioanalyte characterization were performed by Kivilcim Doganyigit. Kivilcim Doganyigit and Jake Revino helped with early prototypes of motor driver and pH sensor electronics. Drew Ehrlich assisted in assembling later prototypes of microscopes, by designing and manufacturing sttructural components and optical components based on modifications to my designs and by performing microscopy experiments shown. Valeska Victoria wrote software for still photo acquisition via USB video protocol for the microscopy device. Drew Ehrlich wrote software for video aggregation and streaming via USB video protocol for the microscopy device.

Chapter 4

Cell culture and Observation Tools for Data-rich Biology

4.0.1 Data-rich biology: bioinformatics beyond sequence analysis

Next-generation sequencing technologies [135, 86] for nucleic acids have produced dramatic increases in the sophistication of biological experiments. The throughput and ease-of-use of these technologies has allowed the breadth of biomolecular information which can be captured from an experiment to expand orders of magnitude beyond early sequencing methods [83]. These technologies and their corresponding bioinformatics tools have transformed how laboratories perform biological experiments.

Despite these advances with respect to sequencing data, bioinformatics has yet to similarly transform the role of other types of laboratory data. Microscopy and biochemical assays are also key to biological research yet, outside of specialized high content screening laboratories, both techniques are still manual processes which are poorly integrated with bioinformatics

tools. We hypothesize that the next major advance in bioinformatics will be to augment sequencing data with other data types. Methods which allow for frequent, repeated, minimally invasive measurements to create long-term time-series data will be especially valuable.

4.0.2 Limitations of sequence analysis

It is important to expand the scope of bioinformatics beyond its current primary focus on sequencing because even the most detailed sequencing data is not enough to fully describe phenotypic information. For example, in single cell RNA sequencing, while the tissue subtypes and functional states of the cells are known, the actual spatial organization of these tissue subtypes is not known without corresponding microscopy data. Other attributes such as cell to cell interaction (such as in excitable tissues or immune tissues, among others), input-output patterns (such as in excitable tissues), morphology of intracellular or multicellular structures (such as in functional organs such as the kidney or liver, among others) or tissue growth and migration generally cannot be determined from any sort of biomolecular sequence data.

In addition, many sequencing methods are also lethal to the tissues: tissues must generally be dissociated and cells must be lysed for the technologies to work. This prevents true time-series experiments, since for each experimental replicate, only a single point measurement can be taken. Even in experimental designs wherein parallel replicates are sacrificed sequentially at different time points generate a pseudo-time series dataset, the data is clouded by unobserved inter-replicate differences. This is especially true in experimental models such as organoids where the overall composition of each replicate is a product of a lengthy stochastic [50] process.

4.0.3 Microscopy and bioimage informatics

Bioimage informatics is an emerging subfield of bioinformatics [100, 123] which seeks to interpret high dimensional data light or fluorescence microscopy to understand the outcomes of experiments. Light microscopy is among the longest-standing modalities of observation [76] in biology, first used in the 1600s to prove hypotheses relating microscopic structure and function in the lung. It allows us to observe the morphology of individual cells, which can give us important functional or developmental information. It allows us to observe interactions between cells such as formation of multicellular functional structures, interactions between interdependent cells or immune responses to foreign entities among other phenomena. It also allows us to observe more macroscopic patterns in cell density distribution, cell recruitment and cell cycle which are important in the study of phenomena such as wound healing or neoplasia.

Modern forms of microscopy also allow optical imaging devices to gather biological information beyond just shape and color. Fluorescent or luminescent probes can augment microscopy by tracking such things as intercellular connectivity [79], transmembrane potential [141] or intracellular [42, 14] or local chemical environment [47]. There are also emerging techniques such as live cell spatial transcriptomics [36] which directly marry genomic data and live cell microscopy. These build on previous spatial transcriptomics [16] techniques which allow microscopy-based spatial localization of transcriptomic data but suffer the same issues as many other transcriptomic techniques in that they require sacrificing the biological entity being studied.

4.0.4 Electrophysiology

Electrophysiology can also provide rich, spatially and temporally resolved information regarding the state of electrically active neuronal, sensory or myocardial tissue. Recent advances in microelectrode fabrication [87] have produced complementary metal-oxide semiconductor electrode arrays whose density and channel count resemble primitive camera sensors. This scale of sensing allows for simultaneous subcellular scale and network scale recording of electrical impulses, providing a rich source of time-series data which reveals functional connectivity and signal processing behavior in experimental tissues. This is a potentially transformative technology for scientific discovery because it allows routine in vitro recording at a density and parallelism which is unprecedented in the study of either in vivo or in vitro electrically active cells.

4.0.5 Biochemical assays via optical and electrochemical readouts

Another core observation modality in experimental biology is through biochemical assays which quantitatively interrogate samples for the presence of certain metabolites, biomarkers, proteins or other molecules which reflect the health or function of cells under study. Like microscopy, biochemical assays have been a cornerstone of biological experimentation for centuries.

Of particular interest are assays with optical [71, 55, 95, 63, 33] and electrochemical readouts [55], since they can be coupled to an automated measurement apparatus, enabling their use in time-series measurements.

For example, immunoassays such as enzyme-linked immunosorbent assays are widely

used to measure secreted factors, metabolites or other biomarkers, and colorimetric or luminescent outputs are commonly used. Concentrations of general classes of molecules, such as proteins or glycans can be quantitated using nonspecific dye binding and/or UV/visible light absorption assays. "Optode" fluorescent probes are a recent development in biochemical sensing: species of interest are detected directly by a fluorophore, whose structure, conformation or electrical environment is directly changed by physical interaction with the species in question. These are of particular interest for routine, high-frequency measurements because they are generally cheaper and longer-lasting than immunoassays.

4.0.6 Opportunities for modern statistics

Modern statistical methods, including but not limited to unsupervised machine learning techniques and computer vision are maturing rapidly both in theory and in implementation. However, on a mathematical level, any high-dimensional statistical method requires high replication, either via many parallel observations or via time-series observations. This is completely at odds with the current experimental practices of the field of biology where experiments generally have on the order of three biological replicates per experimental condition and only a handful of time point analyses, if any. This is not just a cultural phenomenon, it is also limited by the cost and practicality of performing high replication sequencing, microscopy or even manual biochemical assay measurements.

In order to apply the gifts of modern statistics to biology, biology must make it practical to collect adequate quantities of data on a routine basis. This will involve some combination of engineering new data acquisition methods for biology and in using automation to

vastly expand experimental replication. Standardized, automated, highly parallel microscopy is one opportunity for biology to benefit from modern statistics: computer vision is an area of immense research in statistics and computer science and microscopy, capturing the behavior of many cells in detail and simultaneously also provides a sort of internal replication in each microscope image or series of images. Time series automated and computer-acquired, biochemical assay measurements are also promising since in many cases, the actual underlying biochemical processes contributing partially or wholly to the analyte of interest have well validated mathematical models to describe the behavior such that the data can be more easily and accurately interpreted without the computational work and considerable uncertainty involved in also learning a model for the data from scratch.

4.1 Highly instrumented cell culture

We propose the idea of "highly instrumented cell culture": cell culture performed within a system specifically designed to gather time series data, ideally from multiple observation modalities simultaneously. We specifically mean for this technology to be accessible enough that it would be adopted by labs at a similar prevalence to sequencing based assays. It is important that the use of these additional data become standard not just in pharmaceutical drug discovery but also in ordinary academic research laboratory biology.

As a basic example, even automated imaging of all replicates of an experiment every few hours would be a meaningful transformation in the nature of data available in biological experiments. Instead of getting intermittent single images of cell density and morphology, it

would be possible to observe the actual dynamics of these parameters at intervals which are set by experimental needs rather than also influenced by the schedule of human technicians.

More sophisticated examples might include experiments with simultaneous, time-series electrophysiological and microscopy measurement. A system might also measure environmental/metabolic parameters on a per well basis using optochemical or electrochemical sensors coupled to a computer or measure the concentration of some by molecule produced by the cells using a similar technique. While we are taking the concept of highly instrumented cell culture to concern data collection and analysis there is also an obvious opportunity to integrate a highly parallel automated observation platform with one which also performs experimental manipulations such as delivery of chemical factors on an automated basis as well. Measurement of environmental conditions is of particular interest because it would potentially allow more accurate statistical analysis of experiments: major variations in cell growth conditions would be observed rather than unobserved variables. With an automated experimental manipulation system, this could be further built on, using feedback control to actually eliminate these variations.

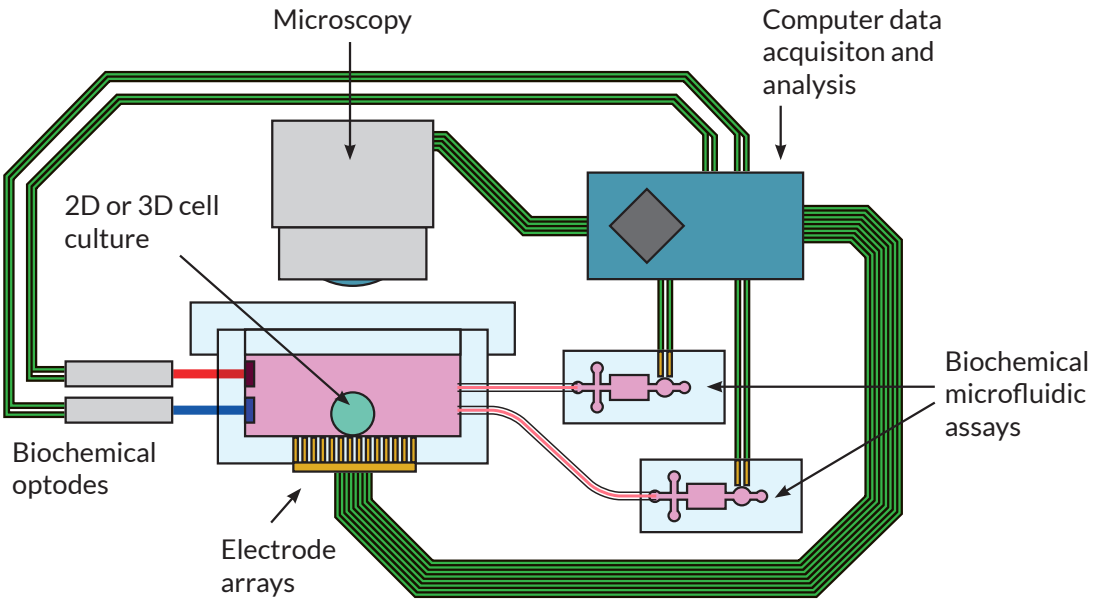


Figure 4.1: **Diagram of a single replicate of a highly instrumented cell culture** This example combines simultaneous electrophysiology and microscopy as well as in-well and fluidically integrated biochemical assays.

4.2 Current laboratory data acquisition technologies

Many of the concepts and approaches which we describe are not entirely new in the literature. However, the various methods described are generally reported as separate, standalone devices rather than an integrated cell culture and analysis platform. What we propose is simply a more unified and in some cases sophisticated approach to methods of experimentation which have been reported in the literature of some specialty biology subfields.

As discussed above, microscopy is among the most data-rich modalities for highly

instrumented cell culture. Commercially, the Curiosis Cell-Logger Mini [116] and the Zen-Cell Owl [51] are miniaturized in-incubator multiwell microscopes designed for long-term, time-series microscopy. However, these are not at present compact or low-cost enough for widespread deployment to observe a large number of experiments in a standard laboratory. The Picroscope [73] is a lower-cost, laboratory built system achieving the same goals however it suffers similar issues with in-incubator bulk.

There currently also exist parallel electrophysiology platforms for low-density multi-electrode arrays at up to 96 replicates per device [10]. For CMOS electrode arrays alone as data collection modality, there exist early commercial products for parallel measurement [12] at up to 24 replicates per device. Current research in our own laboratories is working toward simultaneous, intermittent, in-incubator electrophysiology and imaging using high density CMOS arrays [11].

There is a rich precedent for microfluidic automation of standard cell culture assays [19], though this work has often not been integrated with cell culture devices in any automated fashion. For example, standard, technician-performed benchtop molecular biology techniques such as PCR assays [62], biochemical functional assays [41], cell-counting [45], electrophoresis [59] and immunoassays [88] have been demonstrated in automated microfluidic systems. It is straightforward to imagine and integrating these with computer data acquisition using optical or electrochemical readouts with sample extraction and assay device loading actuated using general purpose cell culture assay automation systems [112, 143].

More recently, there is emerging work on devices which gather time series data from intermittent or continuous measurements obtained directly from cell cultures. Zirath et al, for

example, have demonstrated the use of low-density oxygen optode arrays for time-series sensing of oxygen in hydrogel cultures [142]. Commercially, the Lucid Scientific Resipher [114] and the Presens OxoDish OD24 [38] are similar commercial products for long-term continuous time series measurement of oxygen concentration in microwell format cell culture. Other assays are less well developed commercially, but being pioneered in academic research labs: For example, the Hierlemann lab has reported continuous cell culture monitoring incorporating electronic sensing of cell growth and division [18], glucose concentration [110] and dopamine level [117]. Whole microfluidic systems incorporating multiple readouts have also been developed [132, 143].

Chapter 5

Devices for highly parallel longitudinal microscopy

5.1 Introduction

As described previously, highly parallel microscopy and timeseries microscopy are straightforward choices of modality for increasing the information garnered from and the statistical sophistication of biological experiments. The technology has shown promising results in the academic space, demonstrating potential in genetics [5] and drug screening [127]. Cell tracking via longitudinal time series microscopy has also shown promise in analyzing dynamic phenotypes [40].

Nevertheless, these approaches to microscopy data remain relatively inaccessible to the academic biological research community for routine or large-scale use. Standard inverted microscopes, commonly used for cell culture, cost 10's of thousands of dollars as of the time of

writing. This price does not include the environmental control apparatus necessary to perform these longitudinal experiments on mammalian cells, and furthermore commercial environmental control apparatus suffers problems such as media evaporation. As an alternative, there also exist specialized in-incubator microscopes which allow for longitudinal imaging without needing to maintain a liveable environment outside of the cell culture incubator. Some of these devices allow imaging of multiple wells on the same cell culture plate. However, like conventional inverted microscopes, the cost of these devices makes their widespread adoption for routine use improbable.

To enable longitudinal microscopy, we need devices which can serve the full number of biological replicates that would normally be studied in a single laboratory, but without adding considerable expense to the laboratory's budget. Therefore, we must create devices that can provide microscope imaging at low monetary, labor and space cost. We will describe our building of such a device. This device is compact, low-cost and scales to 24 parallel imaging streams within a single device. It is also rapid to build and durable and designed to survive elevated temperatures such as in an incubator.

This device also served as an exploration of alternative "maker" technologies [46] with which laboratories can build their own prototype instruments for their research. We identified drawbacks to the standard "maker culture" approach of using 3D printing indiscriminately as the major prototyping modality for mechanical parts of devices to be used in the cell culture laboratory.

5.2 Design priorities

5.2.1 Scalability to high replication

There are three principal considerations when designing a microscopy device with the intention of performing hundreds or more simultaneous time series microscopy experiments. The first is the monetary and labor costs of the device: even relatively cheap microscopy devices in the hundreds to low thousands of dollars per microscopy replicate would cost tens to hundreds of thousands of dollars to purchase sufficient devices to perform 100 simultaneous microscopy experiments. This is outside of the budget of most biology laboratories, making highly parallel time series microscopy impractical for widespread adoption. It is similarly important that, should the microscopy devices be manufactured in-house, the labor costs are impractical for laboratories which generally cannot afford to divert significant labor away from their experiments. Second, the devices should be compact enough that it is reasonable to use them in an average laboratory. For example, this means that requiring a cell culture incubator per 24 replicates, as would be required by many in incubator microscopy platforms, is also impractical. It is not realistic with respect to either budget or space that an average biological laboratory would possess 4 or more incubators dedicated to microscopy on top of the incubators they already possess for cell culture purposes. The third is that it must be designed such that the data acquisition platform for the microscopes also scales to hundreds of replicates or beyond. This requires consideration of the networking, processing and storage requirements of the data acquisition platform. Considerations include the scalability and bandwidth of the local microscope to computer networking, the necessary hardware parallelism and associated costs in the data handling

platform, the available bandwidth, endurance and cost of storage hardware and if applicable, the available bandwidth of the Internet uplink.

5.2.2 Robustness to environmental conditions

Since we are designing the device to operate both outside and inside humidified incubator conditions, it is essential that the structural and electronic components are tolerant to the environmental conditions of a standard humidified incubator. As well as total failure of the structure or electronics of the device, we also consider any deformation of the device such that it cannot take repeatable images to be a device failure.

One of the major concerns is deformation of structural components via creep or thermal deformation at the elevated incubator temperature. This outright eliminates the common plastics polylactic acid, polyvinyl chloride, polyvinylidene fluoride, polytetrafluoroethylene, polyethylene terephthalate, polystyrene and acetylene butadiene styrene, among others from consideration for structural components, since these are some of the common plastics with the lowest reported thermal deformation temperatures. [92] While it would be possible to manufacture components from higher operating temperature and/or creep resistant thermoplastics such as polycarbonate or polysulfone, this is not necessarily practical and these plastics will still experience creep in the long term under sufficient load. Therefore, we propose that our parts be designed such that all load bearing structural components are made from metal, ceramic or some other material which is not susceptible to creep at incubator temperatures.

Among metals, either used as structural frame components or as fasteners, it is important that we use metals which are resistant to a humid environment as well as resistant to high

salt concentrations which would be experienced in the event of a media spill and subsequent evaporation. Aluminum and titanium are suitable for this purpose due to their oxide layer. Steel is a versatile material, but in general only stainless steel is resistant to humidity for extended periods of time, and only 316 or more resistant grade stainless steel is resistant to corrosion in contact with humidity and concentrated sodium chloride.

Temperature is also a major concern for the electrical or electronic components of the system. Many electrical components generate heat, which can cause very elevated local temperatures in the actively heated, insulated environment of an incubator. The heat from electrical components can also cause local or global overheating of the incubator. Many electrical components are also sensitive to heat and may cease to function if they are not able to dissipate. It is generally unwise to employ fans to cool electrical components inside of an incubator because this will disrupt air currents in the incubator, compromising the temperature, gas and humidity homogeneity of the incubator, acting as a sweep gas system to hasten evaporation and potentially causing false readings of thermostats and/or gas sensors, compromising the feedback control of those parameters in the incubator. Therefore, it is beneficial to develop designs where as many of the heat generating and/or sensitive components such as microprocessors are situated in a separate module which can be placed outside of the incubator while still enabling the actual imaging apparatus to be located in the incubator. Some components such as motors cannot be feasibly removed from the immediate location where their function is needed. Since electromagnetic motors are both highly heat generating and heat sensitive devices, it is important that the design of any actuators be such that the microscope can hold position with the motors demagnetized.

5.2.3 Open hardware manufacturability and modifiability

5.2.3.1 Maker hardware ecosystems

Many homebrewed laboratory equipment projects rely heavily on open hardware community-adjacent closed hardware ecosystems aimed at makers (such Raspberry Pi [56] and Adafruit Feather [39].) These are attractive, especially to non-specialist labs (as almost all biology labs are) because they provide an ecosystem of microcontrollers, sensors and actuators which are interoperable and have well documented software packages and example code freely provided.

The disadvantage of these devices is that they generally lock the design of the system in to using hardware designed and sold by a single vendor. On a philosophical level, this is contrary to the principle of open hardware [75] unless the hardware ecosystem is in fact completely built on open hardware and software such that it is completely interoperable. On a practical level, this presents challenges with actually sourcing the components to build the system. First, only sourcing complex parts from a single vendor puts fabrication at the mercy of possible supply chain disruptions [130]. Second, it also exposes the design to the possibility of obsolescence at the mercy of the vendor. Third, these parts are generally intended to be sold to consumers rather than hardware manufacturers such that sourcing large quantities of them can be difficult and will generally fail to yield an economy of scale. Fourth, the interoperability of these devices others outside of their ecosystems may be limited. Lastly, these devices are generally designed for ease of typing rather than compactness or efficiency such that they may not be appropriate for construction of devices such as parallel sensor or actuator arrays.

It is therefore important to use maker hardware ecosystems judiciously. For example, despite the caveats listed above, the Arduino [30] class of microcontrollers is often a good choice for prototyping and fabricating devices since Arduino prototyping is essentially a branded and well documented encapsulation of widely available microcontrollers (generally Atmel 8-bit AVR microcontrollers) and low level input output protocols (generally Serial Peripheral Interface (SPI) or Inter-Integrated Circuit (I2C) protocols [65] as well as simple digital, pulse-width modulation and analog input-ouput.)

5.2.3.2 Maker devices and hardware and software standards

Instead of relying on branded ecosystems to guarantee interoperability, we suggest that design of open hardware or other lab-built hardware should be based around industry-standard hardware and software protocols. While some of these standards are not in fact open, this is a philosophical nuance since due to the wide adoption of the standard and low licensing fees it is cheaper and more practical to include the closed but widely accepted standard over something which is truly free hardware/software but which is more expensive or more difficult to implement into a design.

Examples of these standards include the universal serial bus hardware and communication protocol standard, digital signal encoding protocols such as SPI and I2C, and high bandwidth internal connector standards such as Mobile Industry Processor Interface Alliance Camera Serial Interface 2 (MIPI CSI-2) [1] and Peripheral Component Interconnect Express (PCIe) [80]. Non-electronic examples include the National Electrical Manufacturers Association (NEMA) geometry standards [3] for stepper motors, metric thread standards for threaded

lens assemblies, the Deutsches Institut für Normung rail (DIN Rail) mounting format, the Arca-Swiss dovetail joint, the Unified Screw Thread thread standards for connectors (co-opted for fluidic connectors, ie $\frac{1}{4}$"-28 UNF or 10-32 UNF connectors), the Luer taper fluidic connector and the national pipe thread fluidic connector. We also recognize the existence of less formal standards which are nonetheless important to design around. For example, while there is no such formal standard, stepper motors are often available from multiple vendors all conforming to a common informally agreed upon standard of half-moon driveshaft geometry.

The advantage of doing this over just confining designs to a specific maker ecosystem is to better guarantee that hardware will be available from a wide variety of vendors, including the wholesale vendors which would be used if a product were to be adapted to large-scale manufacturing, the existence of backup manufacturers if a source should be impacted either by product discontinuation or by supply-chain issues and also the higher likelihood of a wider ecosystem of interchangeable parts existing, which would make it much easier to adopt a particular design to some other purpose just by changing one or two components without having to make any major modifications to the rest of the system.

5.2.4 Manufacturing techniques

In general, we also adopt the spirit but not necessarily the common practice [4] of the maker community by judiciously choosing if and how to manufacture hardware components in house. While in-house rapid prototyping technologies such as 3D printing have definite advantages to mass manufacturing technologies such as injection molding, they often do not pass scrutiny with respect to layperson accessibility, cost and performance [133] when we are

allowed to consider using services such as fabrication bureaus [57, 2] to make small-run fabrication of components available at low prices within 1-2 week time frames.

Considering using fabrication bureaus is important not just because they make methods requiring large capital investments for equipment, such as metal laser cutting, waterjet cutting and metal CNC milling available to laboratories without these technologies in-house or on-campus, they in fact also make acquiring custom subcomponents even more accessible to most biology labs than it would be to 3D print or otherwise manufacture parts in-house. Most biology labs do not have the expertise, personnel, budget or time to invest in 3D printing, and even if they did, most 3D printers which are within budget of most biology labs produce products which are vastly inferior, in materials, in print quality or in both to what could be produced on industrial machines as service bureau. These concepts are familiar to engineering laboratories which use custom printed circuit boards: it is routine to order small production runs of printed circuit boards from service bureaus whereas it is rare for a research laboratory to own its own circuit board printing apparatus and pick-and-place machines of sufficient quality to make circuit boards for routine use.

For example, one might design a component of an open labware device using either in-house 3D printing or using flat pack assembly of a small number of stainless steel components ordered from a service bureau. The 3D printed device would be dimensionally inconsistent device to device, not robust to high temperature or chemicals, not fire-resistant and not sterilizable by standard laboratory methods, even if printed using polyethylene terephthalate-glycol, the most performant material which is accessible on budget consumer machines [96]. Furthermore it would incur the capital investment of purchasing a 3D printer, the one-off time

investment of having a technician learning to use and troubleshoot a 3D printer and the repeated time investment of manually operating a 3D printer. In comparison, ordering the parts from a service bureau in a material such as aluminum would yield extremely durable, nonflammable, autoclavable, sterilizable parts with essentially zero time investment other than screwing parts together according to an instruction manual.

Therefore, aside from prototyping, my preferred manufacturing tool for our home-brew devices is service bureau ordering of laser cut and CNC press brake bent sheet-metal. We also favor service bureau ordering of CNC milled metal parts over 3D printing plastic alternatives. These technologies and ordering platform allow for inexpensive, durable, environmentally robust components. Furthermore, using modular bolt-together designs manufactured via service bureaus also allows for easy distribution of the parts: parts can be drop shipped to an interested lab or other user and then assembled on site in a straightforward manner akin to building flatpack furniture.

In my analyses of costs and benefits, the main two technologies where in-house manufacturing using maker-space equipment is preferable to use of a service bureau are 3D printing of fluidic devices and CNC milling of plastic parts. 3D printing of microfluidics devices allows the manufacturing of experiment-specific biocompatible devices in the laboratory [106], and it has a key advantage over more sophisticated technologies in that it can produce complex internal negative features in a single operation, something that is topologically impossible with any subtractive manufacturing method. My work on this technology is outside the scope of this thesis. CNC milling of plastic parts will be discussed in the fabrication of optical components as well as elsewhere in this thesis where it is used for fluidic parts.

5.3 Device design

5.3.1 Multi-camera design

We opted for a multi-camera design with the goal of being able to capture images in parallel at high frame rates. The 24-camera design also serves as a testbed for the data acquisition challenges that would eventually arise when even single-camera devices are scaled up to the point of high frequency image acquisition across hundreds to thousands of replicates.

The main disadvantages to a multi-camera design with fixed camera positions is that it is inflexible with respect to different physical layouts of cell culture devices. In our 24-camera device, the camera array is specifically arranged to align with the wells of a standard 24-well plate. As such, it is not compatible with other labware, and a separate device would need to be built.

5.3.2 Camera communication protocol

In general, the options for communication protocol linking cameras to microcontrollers are: direct MIPI CSI-2 [1] connection and USB video protocol [22, 34]. Within fully integrated, consumer products, MIPI CSI-2 is generally the interface of choice, being purpose-built for high bandwidth imaging applications. Provided that the right drivers exist for the sensor and the right interface exists for the data acquisition computer, it is both simple and high bandwidth: up to 10 Gb/s, rivalling the theoretical maximum bandwidth of USB 3 Gen 2. The disadvantages of MIPI CSI-2 is that its use limits the choice of image acquisition hardware. In the realm of consumer microcomputers, the Jetson Nano and a variety of Rockchip systems are

the only options which have a MIPI CSI-2 interface which is fully compatible with a variety of sensors. The interface is usually instead integrated into products by designing custom hardware and firmware to interface it with the processing unit: this is seen in mass production phones, virtual reality headsets, cameras and the like.

For ease of prototyping and modification as well as for component price for small manufacturing runs, we opted for USB video protocol (UVC) as our image capture hardware/firmware protocol. USB video protocol is a driver and interface standard which allows standardized communication between a camera and any microcomputer possessing a USB interface (or able to add one via a PCIe lane.) [34] While the use of USB hardware and an intermediate UVC driver adds complexity to the system, the platform independence of UVC is extremely attractive for open hardware and/or modifiable imaging devices. Systems built around UVC can have the choice of camera swapped without this being apparent to the functioning of the computer and data acquisition software, or can have the computer implementation swapped without having to redesign the interface of the camera. This makes the platform flexible as well as rapid and cheap for iterating through prototypes.

5.3.3 Data aggregation and storage strategy

We opted not to use a cloud platform for initial storage and processing of data due to lack of scalability to hundreds of parallel microscope cameras. This problem has been analyzed in the literature as a criticism of cloud architecture in general [53]. An uncompressed image acquired from a single OV5640 (2592 x 1944 pixels at 24 bits per pixel) sensor totals 120 megabits based on a 24-bit data format [93]; transmitting a conservative image every minute requires 2

Mbps of bandwidth per camera. While this might seem small, this will quickly overwhelm the internet connections currently available to the city of Santa Cruz. Transmitting data from 100 cameras would require a costly enterprise internet connection separate from the university's network; transmitting data from 500 cameras would exceed the bandwidth commercially available to any single enterprise client [52], likely requiring that new physical fiber optic cable be installed through the Santa Cruz mountains connecting the county of Santa Cruz to the internet gateway required to connect to cloud providers. Higher bandwidths, as would be required for video microscopy, would overwhelm the network at a similar rate.

Instead, our system uses a local image acquisition server, which aggregates images from dozens or more camera sensors, stores the images locally and can perform tasks such as image processing, compression or computer vision analysis of the images (to annotate, perform time series analyses or identify important images) before transmitting data to a cloud platform. This avoids the bandwidth demands of a cloud-first system.

We also opted away from the use of 802.11a–802.11ac "WiFi" wireless networking for data transmission to the image acquisition server. This was similarly on account of the bandwidth bottleneck. Our previous work on multi-camera microscopy, which did use 802.11n wireless networking, showed that the wireless network was a bottleneck. Despite its 300 Mbps nominal bandwidth, its real world performance, especially when transmitting through the metal walls of an incubator, was insufficient to transmit images simultaneously from 24 cameras. We attribute this to first, the known problem of scheduling in high data rate, low latency use cases for wireless networks. [74] Second, it is known that the metal walls of the cell culture incubator interfere with electromagnetic waves (including wireless internet signals.) [128] Third,

it is known that the presence of low quality 802.11 wireless connections cause degradation of whole network performance; this increases with additional poor quality connections. [113] It is important to note that this issue is not solved by using additional wireless networking channels: there currently exist only three nonoverlapping full bandwidth wireless networking channels available for consumer wireless networking use in the United States. This bandwidth issue necessitated a move to using USB as our camera module to server data transmission modality and gigabit ethernet as our upstream data transmission modality.

This was an additional factor necessitating our move away from the Raspberry Pi infrastructure (aside from the closed hardware infrastructure and unreliable supply chain) since Raspberry Pi Zero devices do not provide wired network connectivity, an additional factor making them unscalable at the parallelism demanded by our vision for biological experiments.

5.3.4 Camera module

Our goal was to choose a camera module which was extremely cheap and compact and therefore could scale to manufactured devices containing hundreds or thousands of such sensors without incurring unacceptable physical size or component costs.

The Omnivision OV5640 is a 5 megapixel, 2592 x 1944 pixel complementary metal-oxide semiconductor camera sensor with a 1/4" diagonal dimension and a 1.4 micron square pixel size. [93] It is available commercially in multiple packages from multiple vendors in a USB video camera (UVC) driver format. Unfortunately it is only widely available in USB 2.0 interface. This is a disadvantage for multiplexing of camera streams, since using USB 2.0 (which has a maximum bitrate of 640 Mbps) causes a bottleneck which limits the use of multiple

cameras on a single USB controller. However, we find that this disadvantage is offset by the extremely low cost of the camera package. At the time of writing, the camera module (with CMOS sensor, lens, UVC driver PCB and cables) could be obtained for $19.80 USD and had wide stock availability.

5.3.5 Approach to magnification

In earlier prototypes of parallel microscopy devices [73], our laboratory found that, even when using basic lenses (ie without aberration correction) the lenses were a major expense among components of the entire system. We propose a microscopy system which uses the standard lens which is sold with the OV5640 sensor, with only a modification to the optical path in order to achieve higher magnification.

Extension tubes [140] are devices used in macro photography to increase the magnification of a lens at the expense of shortening its working distance. They add no lens elements to the optical path but instead move the lens farther away from the image sensor. We hypothesized that, since the actual size of a pixel on a standard CMOS sensor is smaller than the size of a mammalian cell, not much better than 1:1 magnification would be necessary in order to achieve microscopy which can visualize features on the cellular level. Ultimately we detremined that magnification was 3.5x in terms of ratio between actual size and size projected on the CMOS image sensor.

This principle can be applied to camera modules such as the OV5640 module described above because they possess a removable lens. Our design involves taking that lens which is supplied with the camera and constructing extension tubes in order to increase the

camera's magnification so that it can act as a microscope.

5.3.6 Structural elements

As previously described, we avoided using 3D printed plastic, laser cut acrylic or other commonly used "maker" strategies for manufacturing the structural components of our microscopy system. This avoids the disadvantages which are usually associated with "maker" technologies: that the items produced are usually built to poor tolerances, not durable, not heat resistant and often flammable.

Instead we explored technologies to allow us to prototype in aluminum and stainless steel, materials which are not susceptible to creep at incubator temperatures. Stainless steel is not susceptible to meaningful deformation at loads and temperatures which would be experienced by subcomponents of a microscope at temperatures in the range of 40 °C. This allowed us to also build microscopes using cantilevered structures rather than needing to support the cameras or the stage with some multi-sided actuator as would generally be necessary to compensate temporarily for (ultimately irrecoverable) deformation of a plastic component. This is the strategy which is used in professional microscopes, which are also generally made from metal and have vertical motion components made of metal and cantilevered from a metal column.

Using metal for structural components also has the advantage of adding weight and rigidity to the frame of the device. This is important for the possible application of video microscopy during which the focal plane is moving. If the frame of the device were to be light compared to the part under motion and/or not rigid, the whole device would be susceptible to vibrations which would hamper its ability to gather high quality image sequences.

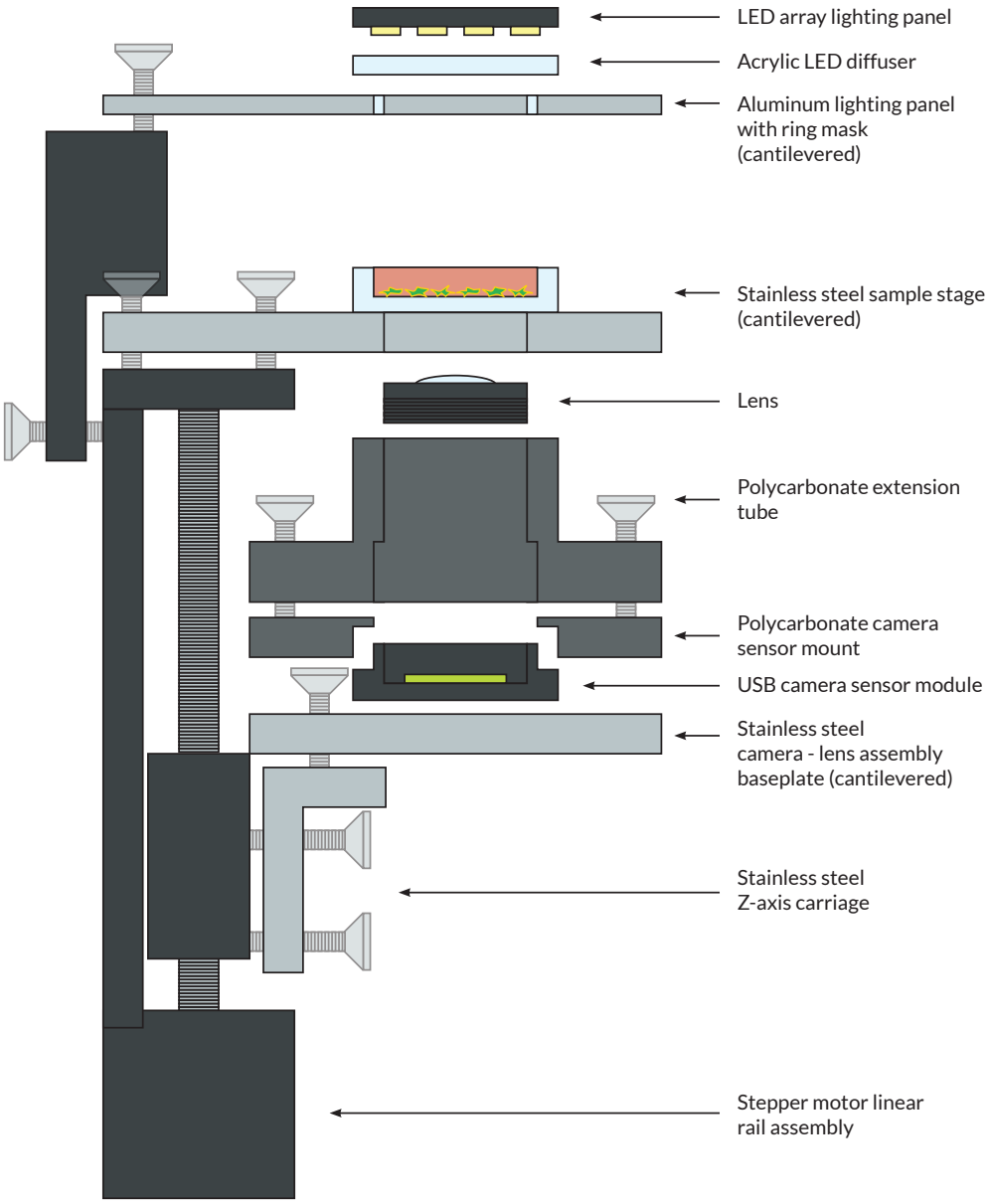

LED array lighting panel
Acrylic LED diffuser
Aluminum lighting panel with ring mask (cantilevered)
Stainless steel sample stage (cantilevered)
Lens
Polycarbonate extension tube
Polycarbonate camera sensor mount
USB camera sensor module
Stainless steel camera - lens assembly baseplate (cantilevered)
Stainless steel Z-axis carriage
Stepper motor linear rail assembly

Figure 5.1: **A schematic of the overall physical design of the imaging system, using a single camera version as an example** The body is built around a stepper motor linear rail assembly and its integrated aluminium frame. CNC laser cut and bent walls (not pictured) provide the rest of the frame. A right angle Z-axis carriage is CNC laser cut and bent or CNC milled from stainless steel. To it is mounted a CNC laser cut camera-lens assembly baseplate. The USB camera sensor module is separated from its lens. The camera module is mounted into a polycarbonate receiving mount, which has been milled to hold it by friction fit. The lens is mounted to a polycarbonate extension tube which has similarly been milled to hold it by friction fit. The camera module mount is fastened along with a polycarbonate lens extension tube to the baseplate. The stainless steel sample stage and lighting apparatus are attached directly to the linear rail frame, with the lighting ring coaxial with the camera(s).

5.4 Device materials and fabrication

5.4.1 Camera sensors

HBVCam 5MP FPC camera modules based on the Omnivision OV5640 CMOS sensor with fixed focus 2.8 mm focal length, f/2.8 aperture M7x0.35mm format lenses and USB Video Camera protocol driver boards were obtained from AliExpress vendors. The camera modules were used as-is with the exception of removing the lens and mounting it on an extension tube as described below.

5.4.2 Lens extension tubes

Black impact-resistant polycarbonate sheet was obtained McMaster-Carr and used as stock for the production of the lens extension barrel assembly via two-sided CNC milling. Milling cutters used were the Harvey Tool Nº 52808 $\frac{1}{8}$" diameter, $\frac{3}{8}$" length of cut square upcut single flute plastic cutting endmill, 52462 $\frac{1}{16}$" diameter, $\frac{5}{16}$" length of cut square upcut single flute plastic cutting endmill and Nº 902208 $\frac{1}{8}$" diameter, $\frac{3}{8}$" length of cut square upcut 3 flute high helix plastic finishing endmills. Milled parts were deburred using an S-blade deburring tool and a razor scraper. Parts were cleaned with compressed air.

All milling steps were performed on a Carbide3D Nomad 3 desktop CNC milling machine. Toolpaths were generated in Carbide Create version 6. XY-axis zeroing was performed using a LCA mini LC/EF spindle mounted laser centering-edge finding tool and Z-axis zeroing was performed manually by eye. Fixturing for basic cutting operations was achieved using Shurtape DS 154 double sided tape on an medium-density fiberboard (MDF) wasteboard which was faced using a fly cutter prior to use.

For fluorescence experiments, an additional opaque black polycarbonate part was CNC milled which held the fluorescence filter and protected against stray light from the fluorescence light source.

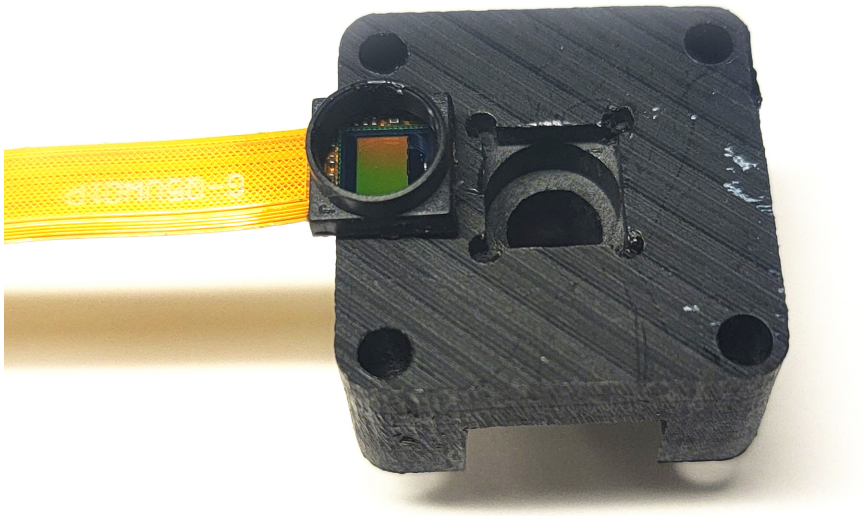


Figure 5.2: **Detail of OV5640 CMOS camera module and polycarbonate CNC milled camera module mount** A three-dimensional cutout in the base of the camera mount provides a friction fit with consistent axial plane positioning as well as height alignment. Dogbones allow for a precise fit along the module edges. The version pictured, designed for a single camera device, is integrated with the extension tube in a single part.

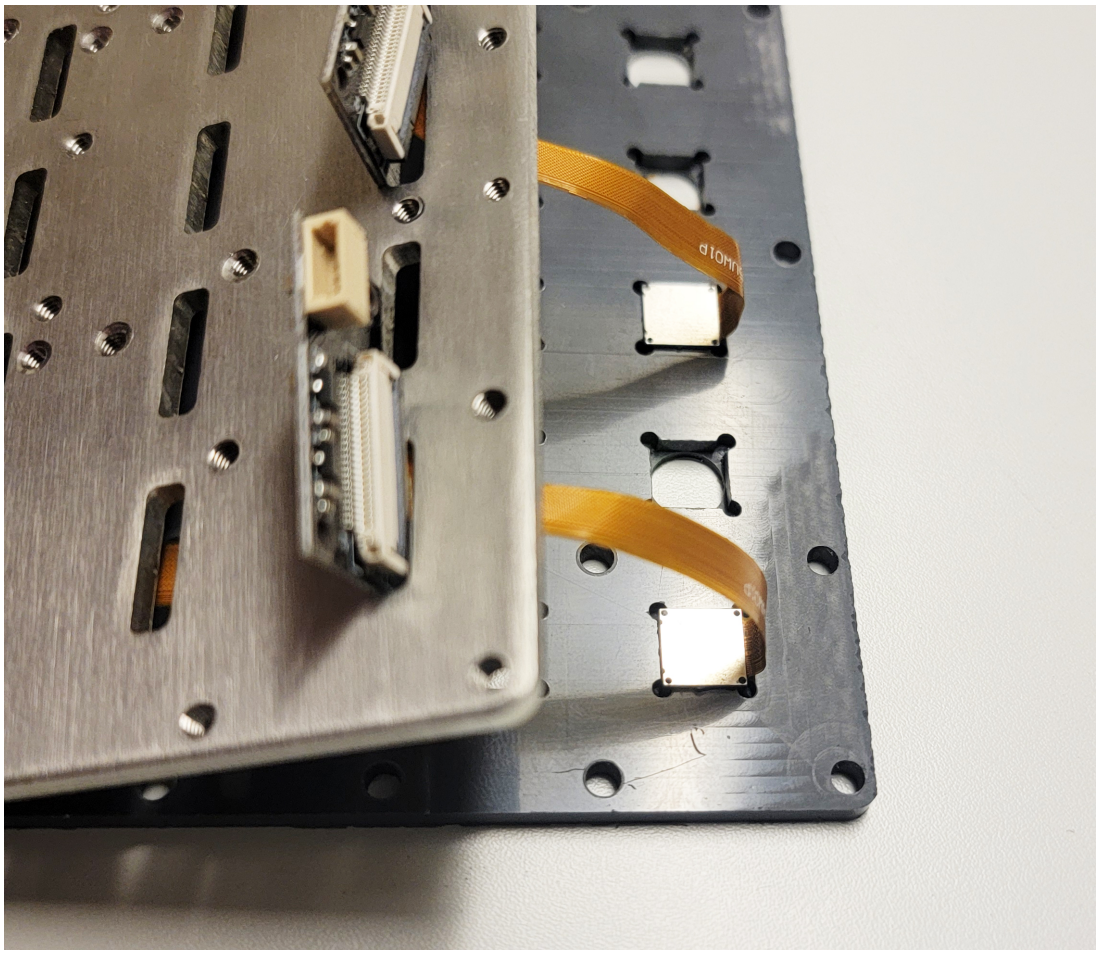


Figure 5.3: **Detail of camera module mounting plate and camera module baseplate for 24 well device** Analogous to the single camera device, a CNC milled camera module mounting plate holds the camera modules in 3D pockets which hold them in place by friction fit as well as aligning them in all axes. This camera module mounting plate is attached to the baseplate, which is shown overlaying it in this image. Note cutouts in the base plate for ribbon cable routing from the cameras as well as threaded mounting holes for the camera module holding plate, the extension tubes and the PCBs of the camera module USB video drivers.

Figure 5.4: **Detail of lens extension tube mounted to baseplate** A polycarbonate CNC milled extension tube, designed for a single camera device, holds a lens taken from an OV5460 camera module. A friction fit is used to hold the lens. The extension tube shown has four corner posts to hold an LED illumination ring.

Figure 5.5: **Detail of lens extension tube array used for 24 camera device** An array of extension tubes is CNC milled from polycarbonate. CNC milling allows precise, consistent positioning of the lenses throughout the array.

Figure 5.6: **Detail of lens extension tube array and camera module mounting block mounted to baseplate, mounted to Z-axis carriage** The extension tube array is mounted directly above the camera module mounting block, which in turn are both mounted directly to the laser cut stainless steel baseplate. This in turn is attached to the Z-axis via a CNC milled Z-axis carriage.

5.4.3 Z-axis

ZeberoXYZ Rail-based Mini Stepper Motor Linear Actuators with T6x1mm lead screws were obtained from Amazon and used for the Z-axis. To mount the plate holding the camera sensors and lens barrels to the Z-axis, a bracket was designed for CNC milling from 6061 Aluminum and ordered milled at an external bureau (PCBWay.)

Pololu Tic Nº T825 USB motor driver controllers were used to control and drive the motors of the ZeberoXYZ actuators. Alitove 24V 4A power supplies were used as independent power supplies for the motors. Electronic leads were terminated with hand-crimped ferrules for all connections between electrical components. The motor drivers were controlled either by the Pololu Tic Control Center or Pololu Tic command line utility.

The linear actuators are driven by a 1.8 degree per step stepper motor using a lead-screw with a 1mm pitch. Without microstepping, this gives a 5 micron precision in Z-axis positioning. Allowing for conservative $\frac{1}{4}$ step microstepping, this gives a theoretical 1.25 micron precision in Z-axis positioning. While higher precision is theoretically possible using finer microstepping, by nature of stepper motors' diminishing holding torque with finer microstepping, this would result in a possible loss of repeatability.

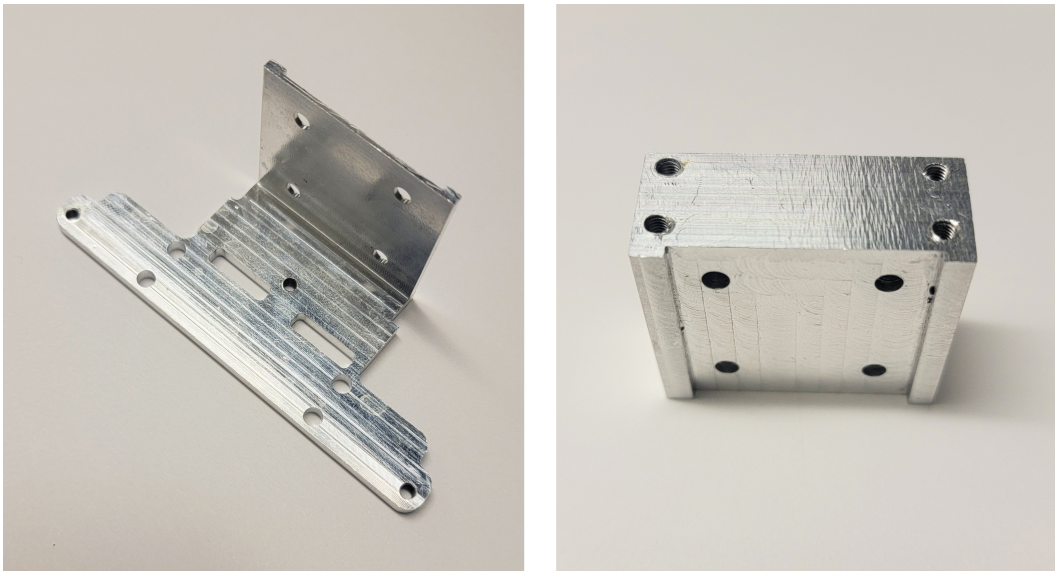


Figure 5.7: **Detail of milled Z-axis carriages** These parts hold a cantilevered baseplate at a right angle to the carriage block of the linear actuator used for Z-axis positioning. Left: minimal carriage, designed for supporting a single-camera device. Right: Version of carriage designed to support the baseplate of a 24 well device. A pocket on the mounting surface provides alignment to the carriage block.

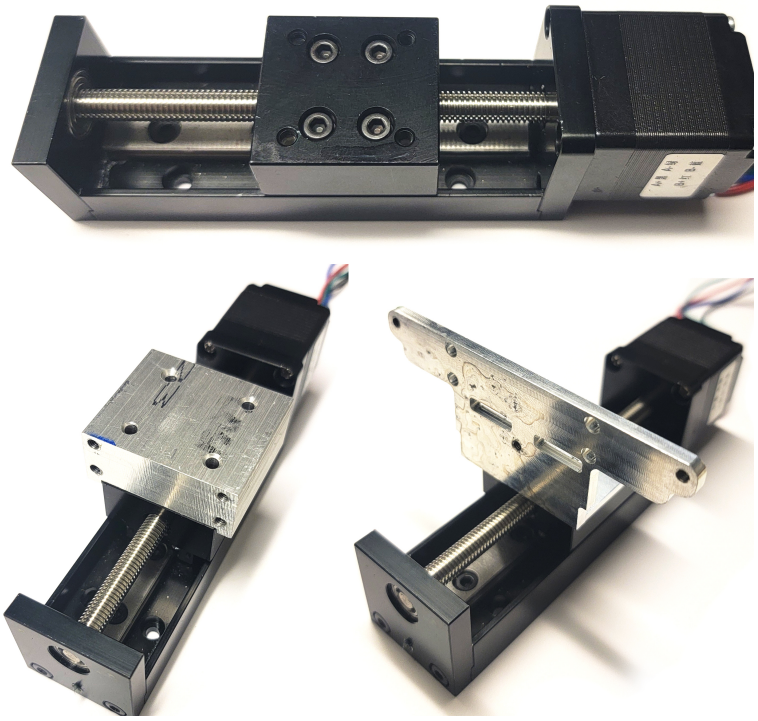


Figure 5.8: **Detail of Z-axis linear actuator and mounting of Z-axis carriages** Top: compact stepper motor linear actuator used in our designs is shown. It comprises a stepper motor driving a leadscrew, driving a carriage block mounted on a linear rail. The carriage block has four M3 threaded holes for mounting. Bottom left: The minimal Z-axis carriage for the single camera design, positioned precisely for mounting to the linear rail. Bottom right: The same mounting scheme, for the Z-axis carriage designed for the 24 camera design.

5.4.4 Camera-lens assembly baseplate

Since plastics are generally susceptible to creep under load at incubator temperatures, or even at room temperature, it was necessary to use a metal baseplate supporting the CNC milled optical apparatus in order to prevent deformation and loss of alignment. This part was

laser cut from 304 stainless steel by a service bureau (SendCutSend) according to .dwg files we provided.

These parts were cut from 304 stainless steel rather than aluminum for thread durability. Aluminum threads on $\frac{1}{8}$" or thinner parts were found to strip when tightened beyond finger-tight. Thread tapping was performed manually in a single downstroke using a Tapmatic Nº 13033 RX 30 self-reversing tapping head and self-clearing black oxide cobalt steel through-hole threading machine taps obtained from McMaster Carr. Gino Development TruePower Drill Press Locking Clamps were used for workholding. Anchorlube All-Purpose Metalworking Compound was used as tapping fluid. This method was found to give good thread surface finish and holding capacity with rapid cycle time.

5.4.5 Microscope body

Sheet metal parts were laser-cut by SendCutSend from either 5052 grade aluminum according to .dwg files we prepared. Where necessary, parts were bent by CNC press brake according to our specifications. The only threads tapped were the mounting holes to mount the rear sidewalls to the linear actuator frame. These screws are for alignment rather than being load-bearing therefore it was acceptable to use tapped 5052 aluminum without risking easily stripping the threads.

Zinc-plated alloy steel Torx M3 and Nº 4-40 flat head screws obtained from McMaster Carr were used as fasteners. Keystone Electronics Nº 4-40 thread zinc plate right angle mounting brackets were obtained from Digikey Electronics and used to fasten body pieces at right angles.

For the 24-well microscope, the specific parts comprising the microscope body were:

1. A set of four 5052 aluminium sheet-metal parts comprising the outer frame. These were a symmetrical pair of wall components, each of which was bent into a U-shape and provided the rigid structure of the microscope as well as its protective walls. A base plate was used as a lower brace for the scope body, and a top plate was used both as an upper brace and as a platform on which samples could be rested.

2. A 90 degree rigid adapter bracket previously described as part of the z-axis which couples the z-axis linear actuator to the stage described below.

3. A camera stage baseplate which provides a rigid and precise attachment point between the z-axis carriage and the moving imaging assembly, which consists of image sensors, lens barrels, PCB holders, lenses and illumination hardware.

5.4.6 Light masking ring

Light masking rings for ring illumination were prepared from 40 mil aluminum sheet which was obtained from McMaster Carr. CNC routing operations were performed on a Carbide3D Nomad 3 desktop CNC milling machine. Toolpaths were generated in Carbide Create version 6. Tooling used was a Harvey Tool Nº 968780-C8 titanium diboride ceramic coated 0.08" diameter, 0.12" length of cut 3 flute variable helix aluminum cutting endmill.

5.4.7 Lighting and diffuser

For both uniform and ring backlight illumination, a plastic diffuser was used to diffuse the otherwise point sources of the LED arrays used for white light illumination. A number of materials were tried; the best performing material was Acrylite® Satinice Nº 0D010DF light diffusing polymethyl methacrylate, which was obtained from TAP Plastics. This material was routed, drilled and countersunk for mounting to the lighting plate using the Carbide3D Nomad 3 CNC mill. The acrylic was then mounted using M3 screws. For illumination, commercial LED panels were used. The images shown in this dissertation used a large photography LED panel purchased from Neewer.

5.4.8 Image acquisition computer

In order to maintain maximum available USB bandwidth per camera sensor, our design used one dedicated USB channel per camera module. This was achieved using Startech Nº PEXUSB3S44V USB 3.1 Gen 1 four-channel USB controller cards connected via PCIe x4 as well as the dedicated USB controllers on the computer's motherboard. An ASRock X399 Taichi sTR4 AMD X399 Motherboard was used to support sufficient USB controllers to provide one per camera module. The CPU used was an AMD Ryzen Threadripper 1900X. 32 GB of RAM was provided. Samsung 970 EVO M.2 solid state drives were used for initial image storage to reduce the potential for the hard drive to act as a bottleneck to image acquisition. The computer was connected to the internet via a 1 Gbps ethernet connection, and could separately be connected to a local storage or processing server via an independent network on a second 1 Gbps ethernet connection.

5.5 Results

5.5.1 Field of view

Field of view and thereby magnification was estimated by measuring the pitch of markings on microscope ruler test slide (Muhwa C4 type 0.05 mm/foot eyepiece ruler) using Adobe Photoshop. With the longest extension barrel, the field of view was measured at 1.039 mm x 0.779 mm. This corresponds to a roughly 3.5x magnification given the dimensions of 3.629 mm x 2.722 mm of the camera sensor.

5.5.2 Camera resolution

An Edmund Optics Nº 58-198, positive high resolution USAF 1951 standard target was used for characterization of resolution. Images were converted to grayscale and the brightness differences between lines were measured using Adobe Photoshop. Using a criterion of observing a 90% or better difference in brightness between lines, the smallest slide element resolved was Group 7 Element 2, corresponding to a measurement of 6.96 microns. Using a criterion of observing a 50% difference in brightness between lines, the smallest slide element resolved was Group 8 Element 2, corresponding to a measurement of 3.48 microns. Using a criterion of observing any discernable (ie 10% or better) difference in brightness between lines, the smallest slide element resolved was Group 8 Element 4, corresponding to a measurement of 2.76 microns.

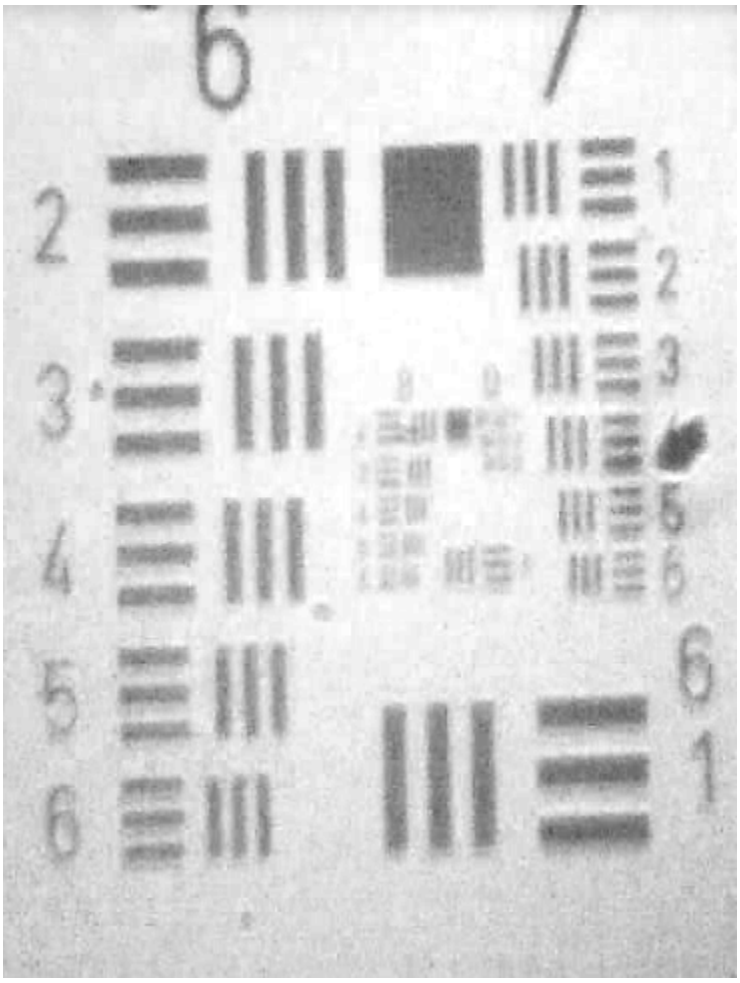


Figure 5.9: **Resolution measured on USAF 1951 slide for our microscopy device** The light and dark tones of the image have been adjusted linearly for better qualitative human viewing; brightness differences were measured on the unadjusted image.

Given the f-number of f/2.8 the lens, corresponding to a numerical aperture of 0.18, using the Ernst Abbe formula for a wavelength of 550 nm to approximate wide spectrum white light, the calculated minimum resolved feature size is 1.57 microns.

5.5.3 Imaging of 3D organoid cultures

Mouse cortical organoids were obtained from the Mostajo lab. Imaging was performed under diffuse uniform backlight illumination. A representative image is shown below, captured in a focal plane which intersects the body of the organoid. The microscope is able to capture individual cells in focus at the periphery of the organoid as well as differences in organoid translucency as a proxy for cell density throughout the body of the organoid. The former is useful for observing and characterizing cell populations or cellular processes projecting from the organoid surface. The latter is useful for identifying putative neural rosettes or other macroscopic internal features of the organoid. Overall, we characterize these results as adequate for tracking the overall morphology and internal features of brain cortical organoids. In particular, the magnification is more than adequate and the resolution of microscopic features is adequate for observing the distribution of cells at the periphery.

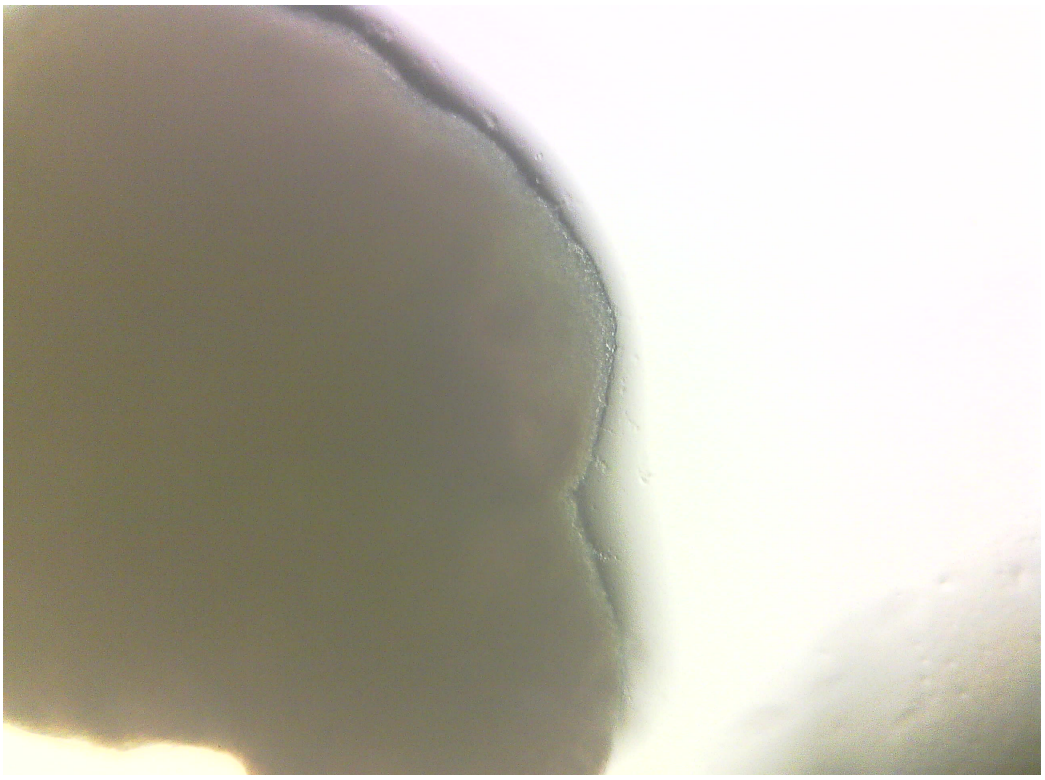


Figure 5.10: **Image of organoid taken with our device** A murine cell derived cortical organoid was imaged under diffuse background illumination using our device. An image showing the full field of view is shown above. A subregion of the same image, demonstrating resolution at the organoid periphery, is shown below.

5.5.4 Imaging of 2D cell cultures

NIH/3T3 murine fibroblast cells were cultured under standard conditions and plated in a T25 filter flask. After one day of incubation, they were imaged. Imaging was performed under the ring backlight illumination previously described as well as without, in order to demonstrate the difference between the two modes. Images were compared against images acquired on a standard inverted phase contrast microscope as would normally be used for basic imaging of

2D cell culture experiments in a standard laboratory. The specific microscope used was a Zeiss Axiovert 25, using a Zeiss Nº 440149 CP Achromat 10x magnification, 0.25 numerical aperture phase contrast objective with the 5x phase ring of the microscope. Representative images from this microscope and from our device are shown below.

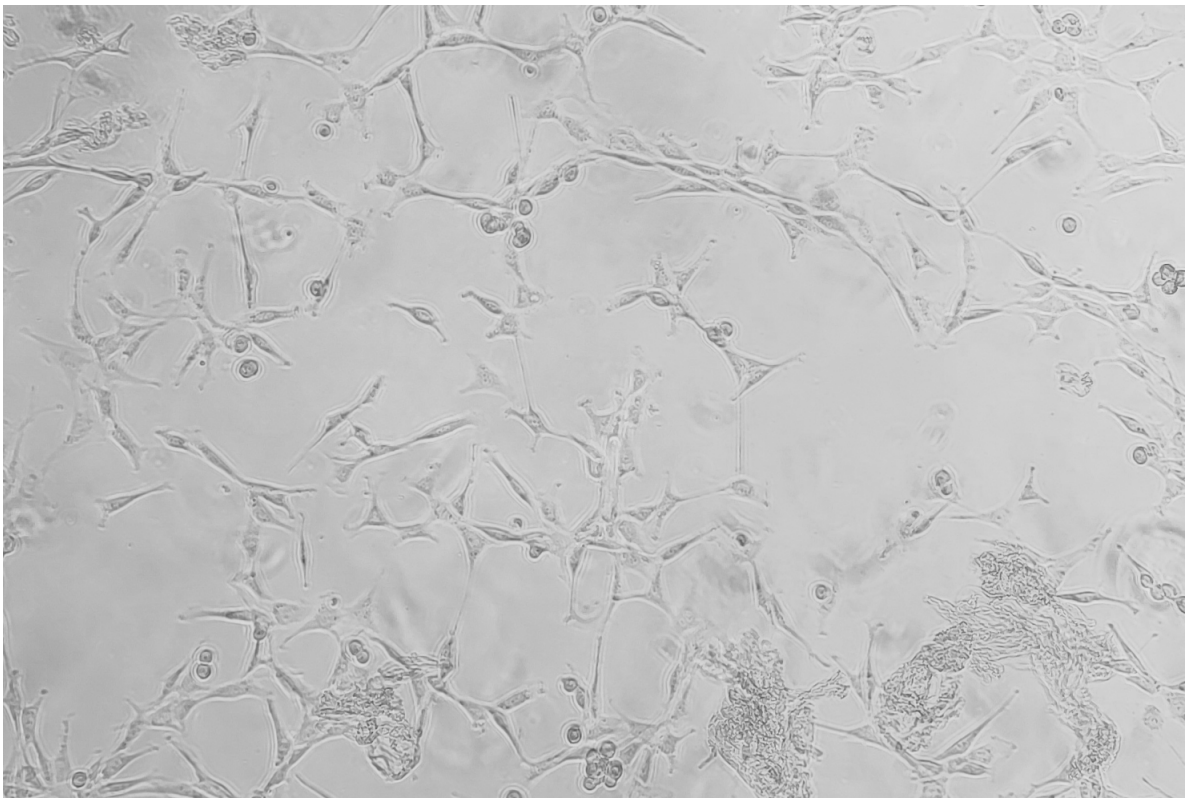


Figure 5.11: **Commercial microscope image of murine fibroblast cells** A gold standard image was taken using a Zeiss Axiovert 25 inverted phase contrast microscope using a Zeiss Nº 440149 CP Achromat 10x magnification, 0.25 numerical aperture phase contrast objective. The image was acquired with a mobile phone camera through the eyepiece of the microscope. Image was desaturated to provide a stricter comparison between imaging modes.

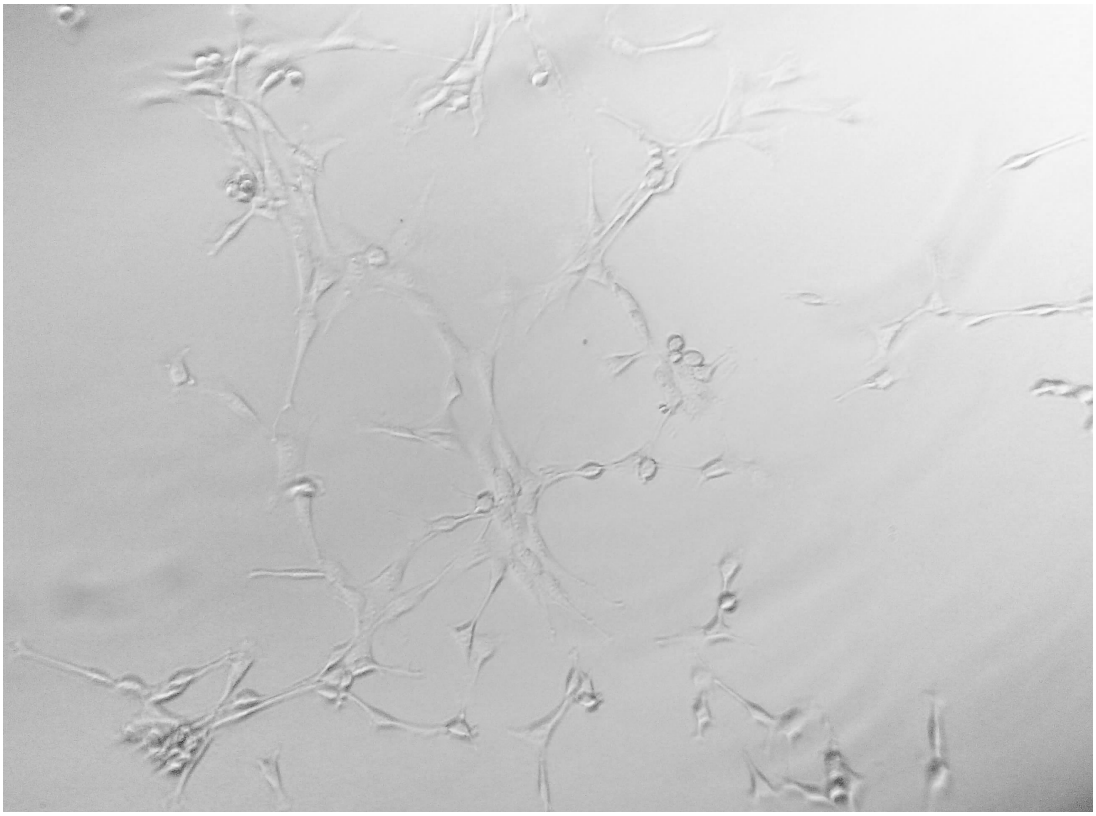


Figure 5.12: **Image of murine NIH/3T3 fibroblasts taken with our device, using tangential lighting ring** Image was desaturated to provide a stricter comparison between imaging modes.

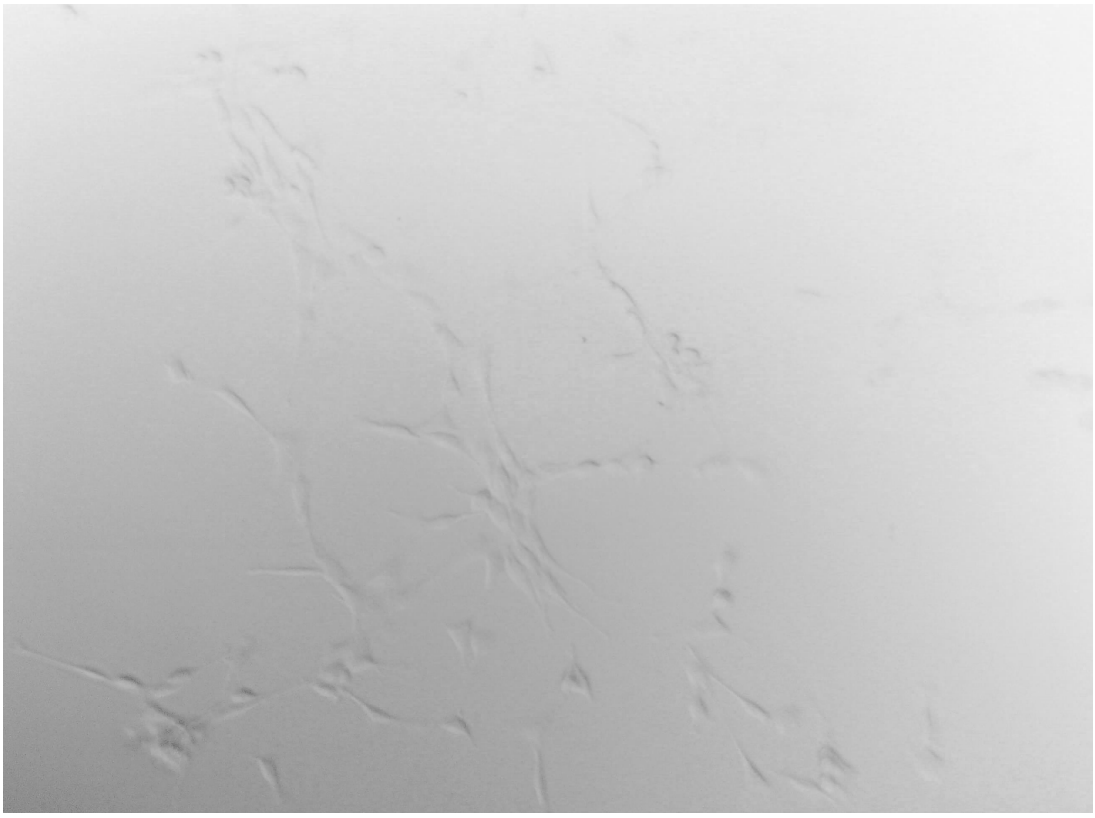


Figure 5.13: **Image of murine NIH/3T3 fibroblasts taken with our device, without tangential lighting ring.** Image was desaturated to provide a stricter comparison between imaging modes.

Additionally, mouse iPSC cells were obtained from the Mostajo lab. In order to maximize contrast at the boundaries of the cells, imaging was performed under the ring backlight illumination previously described.

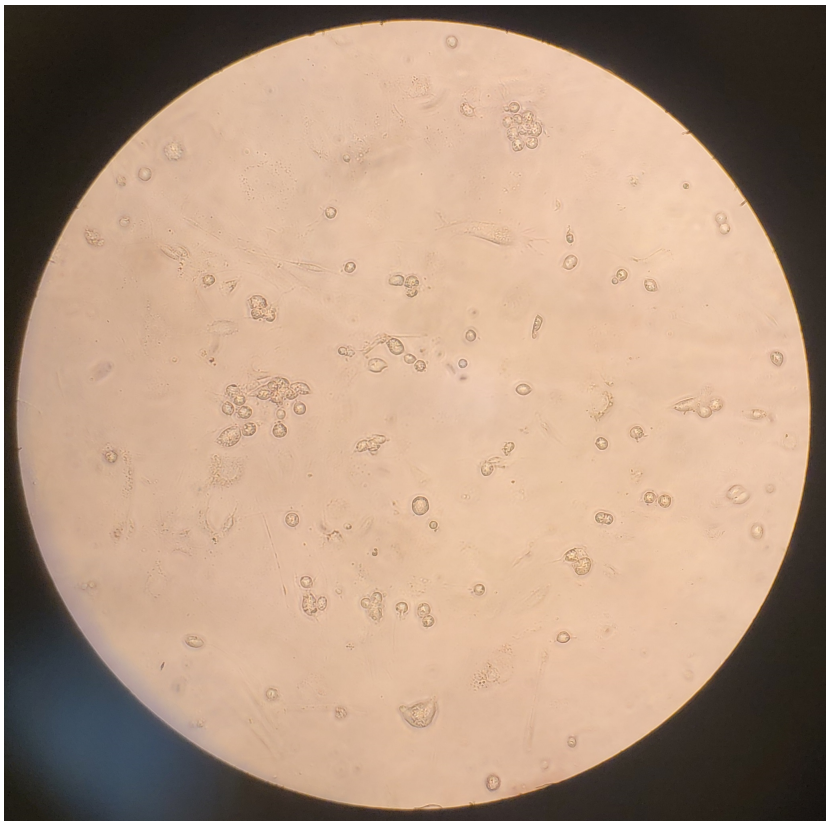


Figure 5.14: **Commercial microscope image of murine embryonic stem cells** A gold standard image was taken using a Zeiss Axiovert 25 inverted phase contrast microscope using a Zeiss Nº 440148 CP Achrostigmat 20x magnification, 0.30 numerical aperture phase contrast objective. The image was acquired with a mobile phone camera through the eyepiece of the microscope.

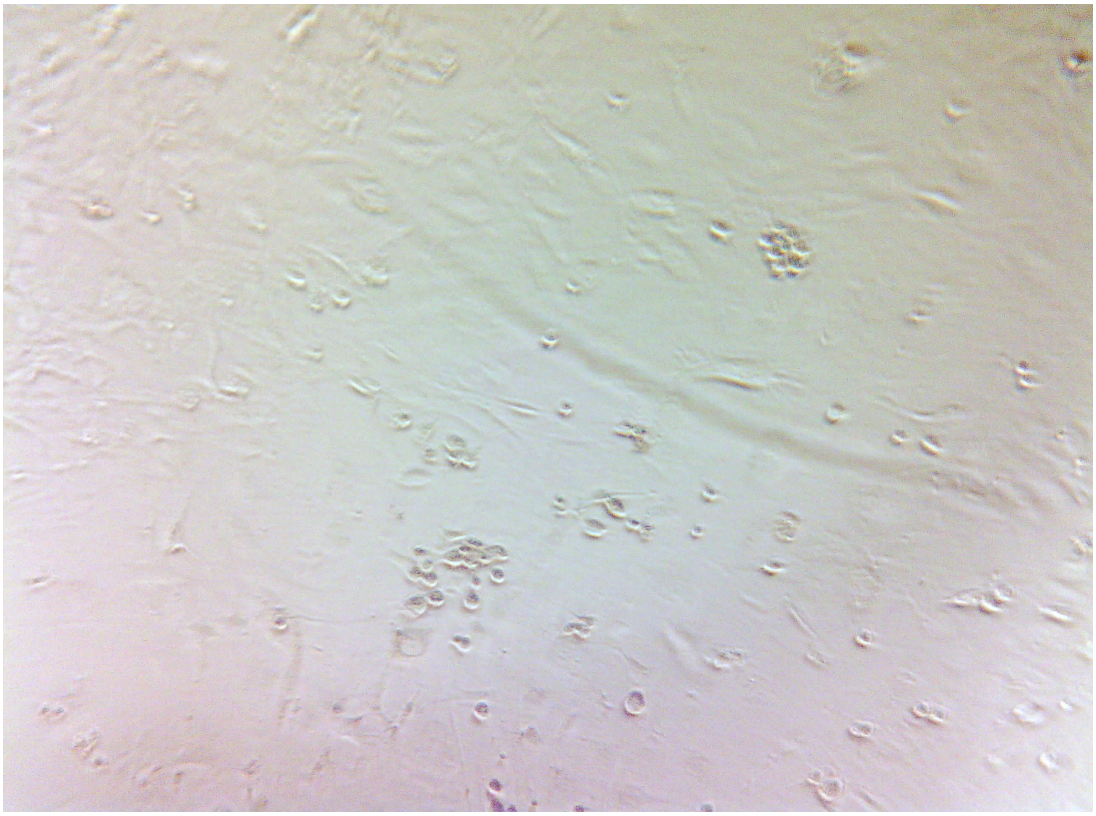


Figure 5.15: **Image of murine embryonic stem cells taken with our device** A field of view containing the same cells seen in the Zeiss Axiovert 25 image.

We qualitatively note spherical aberration impeding the focus of the image as compared to the Zeiss Axiovert 25 image, especially in the periphery. For precision applications, these lenses which are supplied with standard OV5640 USB camera modules are therefore not appropriate, however they do demonstrate the principle of cell microscopy using extremely low-cost single-lens optical assemblies. For higher quality imaging, we can build on these results using low cost aspheric lenses such as the cast plastic lenses available from Edmund Optics.

The image acquired on our device also shows a less homogeneous background than the Zeiss Axiovert 25 image. While similar background features are faintly and indistinctly

visible using the Zeiss image, in the image acquired with our device, they are only somewhat less in focus than the cells. This can be explained by the wider depth of field expected from the relatively low numerical aperture of the lenses used in the current iteration of our device: we are most likely capturing the vitronectin surface treatment of the cell culture plate as well as the cells in the same focal plane. As with the issues with spherical aberration, this problem of thick depth of field could be corrected with minimal modification of our device by replacing the default lens from the OV5640 USB camera module with a low-cost, higher numerical aperture lens.

We note that the ring lighting allows a similar degree of edge discrimination to the commercial phase contrast image, however, for some cells, especially round cells, (dead fibroblasts or embryonic stem cells) the cytoplasm of the cell also shows a shadow pattern as an artifact. While cytoplasmic detail can be appreciated in images of some 3T3 fibroblast cells, this is not consistent and is often impeded by this artifact. This demonstrates that the ring lighting scheme is useful for low-cost microscopy in settings where edge discrimination is important, such as certain computer vision cell tracking applications or for low-cost teaching labs.

Chapter 6

Data-rich biology needs new cell culture techniques

6.1 The current cell culture standard: air-gap respiration

Despite the potential for highly instrumented cell culture or some similar data collection approach to usher in a second data revolution in biology, the actual act of gathering these data is at odds with the standard labware and procedures used to house tissue cultures in the biological laboratory. Specifically, current technology achieves gas regulation of the contents of cell culture vessels using air gaps in these vessels, thereby relying on sealed incubator chambers to regulate the pH, temperature and osmolarity of cell culture media. This regulation scheme is hampered by the simple action of trying to observe the tissues being cultured, and conversely, this regulation scheme hampers attempts at observation, regardless of whether the observation is by manual or automated means.

These environment-regulating incubators are necessary because the classic ecosystem of mammalian cell culture workflows and equipment is built around labware which incorporates an air-gap for respiration: either dish-and-lid or vented-flask cell culture vessels. In the dish and lid system, the air-gap is a literal gap between subcomponents: the lid is elevated slightly over the dish to allow for airflow [68] between the immediate environment surrounding the cells in the dish and the external environment. This is the method used in both standard single well cell culture plates as well as in standard microwell plates. In the vented flask system, a microporous material built into the lid performs this same gas exchange function [21]. In either case, since air gaps are nonselective for gases, water vapor is also exchanged meaning that in order to reduce evaporation, water vapor must also be added to the carbon dioxide incubator through use of some sort of humidification apparatus.

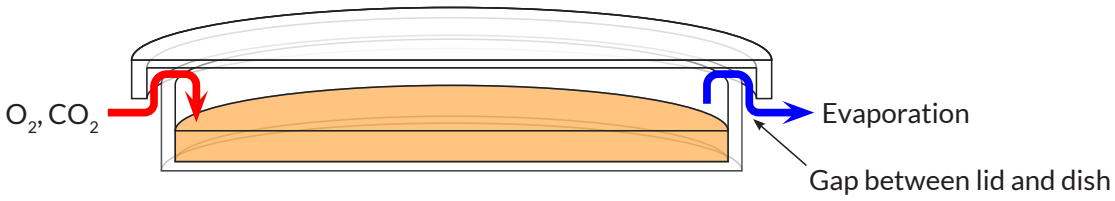


Figure 6.1: **Diagram of standard cell culture dish** An elevated lid allows oxygen and carbon dioxide exchange as well as water vapor egress. Microbes may also enter via the air gap, especially if a drop of media is splashed and bridges it.

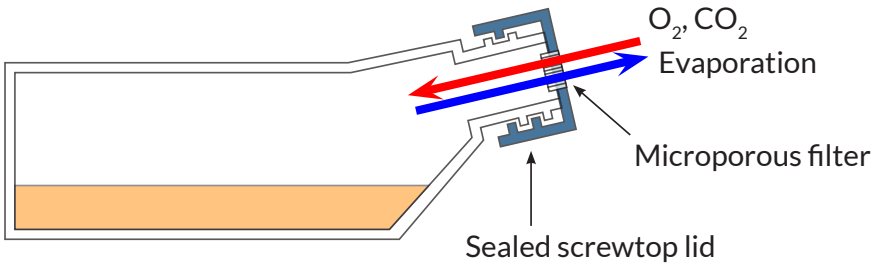


Figure 6.2: **Diagram of a standard cell culture vented flask** A microporous filter in a screwtop lid provides gas exchange by free transport of gases including oxygen, carbon dioxide and water vapor. Microbial ingress is limited by using a filter with submicron pores (ideally 0.1 μm, though 0.22 μm and 0.45 μm are common.)

The environment-regulating incubators function as follows: A mixture of room air, nitrogen and carbon dioxide provide regulated oxygenation and pH buffering respectively to cells in a bicarbonate based buffer system, with gas exchange occurring through an air gap in the cell culture vessel. Regulated levels of carbon dioxide are added to room air to participate in the carbonic acid-bicarbonate-carbonate buffer system (which is the same buffer system which exists in actual living beings.) These carbon dioxide levels generally range from 5–10% depending on the concentration of the buffer system used. In any case, this is a significant amount of added carbon dioxide versus the nearly nil concentration of the gas in room air. In "normoxic" cell culture, room air without addition of nitrogen is used to make a gas blend with the carbon dioxide, producing a regulated oxygen level slightly below that of room air but still well above physiological. Cell culture at lower oxygen levels (either physiological "physoxic" or

sub-physiological “hypoxic” oxygen levels) adds nitrogen to the gas blend in order to displace and lower the concentration of oxygen well below that of room air.

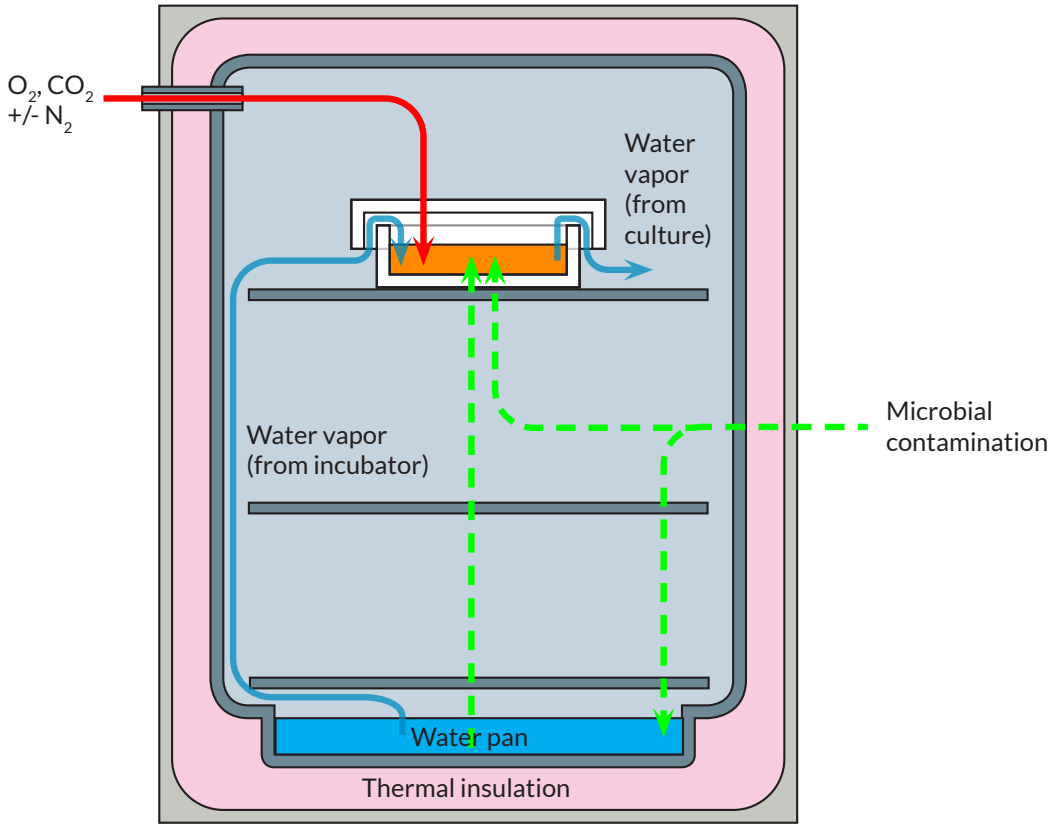


Figure 6.3: **Diagram of a cell culture incubator containing an air-gap cell culture dish** The incubator is a bulky, thermally insulated cabinet which surrounds the cell culture vessel. A regulated flow of oxygen, carbon dioxide and possibly nitrogen, under feedback control, maintain gas concentrations. (Red path) Water evaporates from the culture via the air gap in the vessel and is partially replenished by evaporation from a water pan at the bottom of the incubator. (Blue path) Airborne microbial contamination enters the incubator and deposits on the cell culture vessel as well as in the water pan or other condensed water, which then serves as a nidus for microbial growth.

6.2 Conventional air-gap cell culture hampers the use of instrumentation for data collection

Microscopy is among the oldest and most common methods of observation in biology, yet the standard paradigm of air gap cell culture is not compatible with long-term or frequent microscopy. Air gap cell culture is possibly even less compatible with other observation methods such as those which require inbuilt sensors in the cell culture device, such as electrophysiology or bioelectronic sensing or those which require frequent fluid sampling such as timeseries biochemical characterization.

6.2.1 Problems with microscopy

Microscopy is among the most versatile and data-rich methods of observing cell cultures, but in general, a laboratory will have one or several shared benchtop light microscopes

inside their lab but outside of their incubators as well as more sophisticated microscopes in a central microscopy core facility which might be as far away as in a different building from their lab.

This means that every time that a microscope observation needs to be taken, a technician must open the cell culture incubator (thereby releasing carbon dioxide and water vapor, neither of which will return to baseline for a number of hours), carry the cell culture vessel to the microscope, gather images, then return the vessel to the incubator, once again subjecting the entire shared incubator environment to the same disruption of carbon dioxide and water vapor concentrations. Each opening, closing and especially returning of a cell culture vessel to the incubator is a potential microbial ingress event, seeding a colony which will multiply in the warm, moist environment and potentially eventually contribute to experimental failure. These factors limit the number of times that an incubator can be open and close per day and still maintain the semblance of desired operation.

Furthermore, this microscopy setup also limits the frequency, number of replicates and the amount of data per replicate which can be collected. In general, microscope stages are at room temperature and do not provide cultures with mixed gases or humidification so as time outside incubator at the microscope increases, so does the magnitude of temperature, pH and oxygen perturbation of the culture environment. When these sorts of microscopes are available, their pricing complexity means that laboratories generally do not have enough of them to perform experiments with high replication. Since microscopes are also a shared resource between multiple experiments, they are also not available for imaging experiments that last the entirety of the day, day after day.

Lastly, the standard incubator set up does not lend itself to automation and so microscopy relies on human labor. This means that the frequency and timing of microscopy data acquisition is also limited by the schedule of the technician performing the experiments: for example a simple time series experiment with microscope images collected every four hours would not be practical from a labor point of view.

6.2.2 Problems with in-incubator approaches to observation

One option is to place instrumentation inside of the incubator, an approach which is used by some commercial XY stage based cameras, some commercial electrophysiology platforms and also by our own lab with the current iteration of our continuous imaging microscope. However these solutions come with their challenges.

It is difficult to engineer reliable devices which operate in the heat and humidity of the incubator. Specifically, many plastics suffer significant creep or worse, plain thermal deflection under load at incubator temperatures, which necessitates more sophisticated manufacturing. Most metals also risk corrosion and will also release metals into solution in a moist environment which compromises structural integrity, electrical integrity and may even present a serious electrical shock risk since not all incubators and laboratories are properly wired to ground fault circuit interrupter power supplies. The devices must also not generate significant heat themselves. Even powering, controlling, and sending data to and from these devices is limited, since the efficiency of wireless technologies is greatly decreased by the faraday cage-like structure of the incubator, and there is limited real estate to route wires from the inside to outside of an incubator. Any equipment placed into a cell culture incubator is also both a contamination

risk, especially if it is difficult to sterilize and may also perturb the thermal performance and convective gas transport of the incubator, leading to an inhomogenous and/or off target humidity and gas concentration throughout the incubator. Lastly, incubators take up valuable space, electricity, gas control and supply and maintenance resources which are generally in short supply under the budgets of normal academic research labs. Any microscopy or other apparatus placed inside the incubator takes up space which could be used for many more cell culture replicates. Supposing that one set of cell culture and observation devices, potentially including microscopy, microfluidics physiology, microfluidics and other devices takes up one incubator per 24 replicates, this requires that the lab purchase an additional incubator at the cost of $5000 to $10000 for every 24 replicates. Even if the additional costs of a non-incubator cell culture platform requires several thousands of dollars of additional equipment per 24 replicates, this cost is offset by the savings of not purchasing an incubator.

6.2.3 Other problems with modalities requiring inbuilt instrumentation

Other non-imaging modalities are emerging in terms of high bandwidth longitudinal observation modalities. In electrophysiology, commercial electrode arrays, especially CMOS multielectrode arrays are emerging as a powerful method to directly observe, via electrical impulses, the cell to cell communication of tissues such as brain or myocardium. However, these electrical systems require amplification and digitization of signals on a per-channel basis, requiring that large amounts of sensitive electronics must be colocated in the incubator with the cell cultures. Similarly, optochemical devices generally require laser light sources, sensors and actuators to be colocated with the cell cultures, again forcing sensitive electronics to be placed

inside the incubator. On top of the challenges and expense of making these electronics compatible with the warm, humid and, ideally, tightly regulated incubator environment, the electronic demands of the system also require power supplies and high bandwidth networking connections into the incubator. This means that a conduit into the incubator for wiring must be maintained, which also becomes a conduit for contamination and an egress for carbon dioxide and water vapor, compromising both the sterility and either the efficiency or the quality of the enviromental control of the incubator. Aside from the presence of a persistent conduit for contamination, it is also often difficult to sterilize instruments. Electronic equipment is often sensitive to autoclave and oxidizing sterilization due to corrosion of electronically integral metal components. Electrochemical probes and some optochemical probes are often sensitive to oxidizing chemical sterilants such as hydrogen peroxide or ozone. Other sterilization methods such as dry heat sterilization, milder oxidizing agents such as dilute peracetic acid or nonoxidizing chemical sterilants such as glutaraldehyde are either not accessible to standard laboratories or present safety concerns for routine use and may still not be compatible with these sterilization modalities.

Electrophysiological or biochemical instrumentation of cell cultures also increase demands on the fidelity of cell culture environments. For example, the electrophysiological phenotypes of tissues may be exceptionally sensitive to changes in the concentration of ions such as sodium, potassium, calcium or chloride, since the biological machinery of electrical signalling in the cell relies on sensing and manipulating the concentrations and transmembrane gradients of these species. Therefore, an approach to gas exchange for cell culture which allows drastic changes in electrolyte concentration as a consequence of its normal operation may not be appropriate for realistic electrophysiology experiments. Similarly, if experiments aim to

characterize metabolic processes, growth or cell stress to provide a richer experiments investigating the biochemical state of cells in experiments, it may no longer be acceptable to achieve nonphysiological and/or inconsistent oxygen or pH levels.

6.3 Water loss in air-gap systems

In addition to the difficulties which airgap systems with respect to acquiring observations for data rich biological experiments, airgap systems have significant physiological disadvantages which exist independent of the challenges of instrumentation or need for frequent access. As mentioned above, routine use of biochemical or electrophysiological observation modalities in cell culture will likely expose issues with the poor environmental fidelity of current cell culture approaches.

An immediate consequence of an air-gap scheme for and carbon dioxide exchange is that water vapor is also freely exchanged. Any conduit to an external environment with less than saturated humidity causes a humidity gradient to be established. This gradient in turn produces evaporation [101]. Since the size of the external environment also generally exceeds the internal environment by a large factor, this gradient can be approximately viewed as an infinite sink creating continuous loss of humidity.

Therefore, maximizing humidity in the external environment is essential to minimize a gradient which would draw water vapor out of the system. This is complicated by the need to access the shared incubator environment in order to service cell culture vessels for feeding, manipulation or observation. The need for access increases proportionally to the number of

vessels stored per incubator, creating a conflict between viability of cells being cultured and efficient use of space and equipment. This is a significant barrier to the scalability of cell culture based experiments in conventional laboratories. Potter and DeMarse [104] demonstrated that even brief routine openings of the incubator cause reductions in relative humidity lasting over an hour per opening. In an incubator in which even only 4 or 5 small experiments' worth of cell culture vessels needed to be accessed during a single day, one could extrapolate this finding to show that it is possible that the incubator is never at 95% relative humidity during the working day: operations which remove the vessels from the incubator generally need at least two openings of the incubator (one to remove vessels, one to replace) temporally spaced by the length of the operation, and since operations generally cannot be timed concurrently due to limited equipment such as microscopes and biosafety or laminar flow cabinets or workers' schedule conflicts.

However, even the best attempts to reduce incubator openings in order to prevent evaporation are in vain. Experimental results have shown that even with the complete elimination of any incubator openings, such that maximal humidity can theoretically be maintained without derangement from incubator openings, significant evaporation still occurs [31], since incubators do not maintain truly saturated humidities. This is most likely because fully saturated air in the incubator is both difficult to achieve and comes with its own drawbacks. Any reservoir of standing water, especially one maintained at a temperature on the order of 37 $^{\circ}$C is susceptible to microbial contamination. Therefore, the pooling on all surfaces which would occur in a fully saturated atmosphere would impart the system with significant contamination risks.

Removal of water vapor via evaporation does not just deplete volume from the cell culture device. It also causes significant derangements in concentrations of electrolytes and the other largely nonvolatile species whose concentrations directly affect metabolism and function of the cell.

Not only is there always evaporation in air gap based cell culture devices, it is also not even across the dish or device [31]. Differences in humidity gradient between the centre (farther from the air gap) and edge (closer to the air gap) are established in lid-dish type designs since the external humidity will always be below 100% whereas evaporation of media leads to local humidities approaching 100% near the center of the lid. Multiwell plates are very commonly used for replicate experiments in tissue culture but suffer a severe consequence: inhomogeneity of evaporation between wells depending on position on a supposedly even plate. In-house research by Thermo Fisher Scientific has shown that this causes meaningful differences in cell viability. In their experiments, after a 3-day incubation, differences in cell viability between central and perimeter wells, attributable to a sole variable of evaporation, exceed 4% [115].

In particularly sensitive experiments, various solutions have been tried, including not using edge wells of multiwell plates or using plates with built-in ”moats.” [115] However these strategies can still show measurable edge effect.

These effects are not even limited to within plates, but also differs based on position in an incubator: humidity in an incubator is usually established by evaporation from a water basin placed in the bottom of the incubator [31]. Thermo Fisher’s in-house experiments revealed as much as a difference of 5% versus 7% evaporation of equally setup dishes over 7 days between upper and lower shelves in a standard incubator maintained at the industry-standard setting of

95% relative humidity.

We will show preliminary data establishing that under standard cell culture conditions, using standard cell culture dishes, media becomes hypernatremic within a day. These metabolic derangements are even more serious in the context that common cell culture media formulations already contain supraphysiological levels [81] of certain electrolytes or nutrients and evaporation only heightens this derangement. For example, human plasma contains 136–145 mM sodium and 98–106 mM chloride versus baselines of 156 mM and 121 mM respectively in Dulbecco's Modified Eagle's Medium.

6.4 Review of existing approaches to instrumented cell culture

Existing devices designed to circumvent the problems caused by the difficulty of maintaining physiological gas concentration and temperature while also gathering time-series data are generally aimed at enabling live-cell microscopy, electrophysiology or both. Aside from methods which place the imaging or electrophysiology equipment inside a conventional cell culture incubator, discussed above, the remainder of the devices fall into one of three rough categories with respect to how they handle gas equilibration. The first is to exchange gases within the cell culture compartment itself by enclosing an open culture vessel in a miniaturized gas chamber. We will describe these as miniaturized incubators. The second is to flow gas through or across the cell culture chamber. We will describe these as air perfusion systems. The third is to equilibrate the cell culture media with gas outside of the cell culture vessel then inject it into the cell culture vessel. We will call these gas preconditioning systems.

While there also exist approaches which use carbon dioxide-independent media, for example the incubator-less, sealed, perfused system described by Minuth et al in 1996 [84], or the more recent Saalfrank et al [111] system aimed at time series simultaneous electrophysiology and microscopy, these are generally not acceptable to many labs for mammalian cell culture experiments. Carbon dioxide independent media largely eliminates the evaporation seen in carbon dioxide-based systems, since there is no gas-liquid interface, however it relies on nonphysiological pH buffers. This essentially renders it a non-solution for imaging: the standard carbon dioxide-independent HEPES buffer becomes cytotoxic on exposure to light, yet by necessity, microscopy requires optical windows to be present.

6.4.1 Miniaturized incubators

The basic concept of the miniaturized incubator is that instead of placing cell observation equipment inside the incubator, you put the cell culture incubator on the specimen stage of the instrument. Essentially all non-conventional incubator systems for temperature and gas control of adherent cells or 3D cell cultures during observation follow this format.

The general concept dates back to 1908 [99], when it was used for microscopy of live eggs. One of the earliest and simplest miniaturized incubators for mammalian in vitro cell culture was described by Bavister in 1988 [7]; it was designed for live cell imaging. It consisted of a simple plexiglass chamber into which a continuous flow of premixed carbon dioxide-oxygen-nitrogen blend was flowed, which in turn was placed inside a heated microscope chamber. More modern devices follow a similar design, using preconditioned gas blends under flow, except that in general the heating module is integrated into the device itself. In all cases, these devices are

wasteful of gas, requiring active flow of gas blends in order to maintain concentration [7]. This is not a problem for their usual application, which is occasional use for rare live-cell imaging experiments. However, this makes the solution impractical for routine use with many parallel replicate devices, as would be required by a cultural shift towards highly instrumented cell culture as a routine approach to biological experiments. Continually venting mixed gas to the room environment is not cost-effective or practical given the shipping, equipment and safety requirements of using compressed gas.

Ho et al [49] circumvent this problem by recirculating their gas perfusate. An actual cell culture incubator is co-opted for feedback control of gas concentration, also avoiding the use of pre-blended mixed gas, an inline pseudo-humidifier consisting of a bubbling apparatus is added to limit evaporation and the entire line is recirculated back into the incubator to avoid spillage of carbon dioxide enriched air.

Despite addressing the problem of gas wastage, miniaturized incubator systems are generally even more susceptible to evaporation than conventional incubators (which already produce evaporation.) This is due to the flowing gas which they use to eliminate the need for active carbon dioxide feedback control within the chamber: indeed, flowing gas over water's surface is a common method of intentionally promoting evaporation [101, 6]. We could imagine a system with humidification and carbon dioxide feedback control inside of the miniaturized incubator, however this is generally avoided due to the expense of the carbon dioxide sensors and computerized gas injection valves that are necessary for feedback control.

6.4.2 Gas preoxygenation

Integrated, sealed cell culture systems for continuous microscopy [64] have been developed using nonporous gas permeable tubing-based gas preconditioning devices which pass through incubator mixed gas chambers to control in-well gas concentrations rather than exposing the cell cultures on the microscope directly to mixed gas blends. However, these implementations have generally used polydimethylsiloxane cell culture chambers, which themselves leak carbon dioxide and water vapor to the environment. Therefore, these devices also require active flow to maintain gas concentrations or osmolarities in target ranges. This limits them to perfusion culture conditions unless the cell culture chambers are also placed inside a gas controlled and miniaturized incubator, defeating the purpose of a remote gas exchange preconditioning module.

As one of our proposed methods of gas control and osmolarity maintenance for highly instrumented cell culture, we have prototyped an extension of the preconditioned media perfusion bioreactor concept which, to our knowledge, has not yet been reported in the literature. First, we use a gas impermeable cell culture chamber and a less water vapor permeable gas exchanger material in order to prevent loss of carbon dioxide or water vapor to the outside environment. This approach has been demonstrated before, but only in perfusion culture [20]. Second, since our cell culture chamber is gas impermeable, we do not need active media flow to maintain gas concentrations within acceptable ranges, therefore we are able to use preconditioning of media without needing active perfusion of the cell culture. We have prototyped a cell culture device which uses gas preconditioning and static media cell culture with regular,

automated media changes.

Chapter 7

Air gap free cell culture

We have demonstrated prototypes of a system which diverges from air-gap designs, primarily to allow for a more flexible and scalable sort of cell culture which will enable denser instrumentation of experiments to allow for continuous multimodal data collection, allowing for a paradigm of live-cell genomics where rich phenotypic information is included in bioinformatic analyses of experiments. We also demonstrate how these devices produce more consistent and replicable metabolic parameters by eliminating evaporation and reducing the effect of inconsistencies in gas blend. We also describe how the design of our system has other consequences, such as allowing for otherwise inaccessible experiments which use different gas blends, hydrostatic pressures or toxins than normally available in the laboratory.

7.1 Nonporous membrane respiration

7.1.1 Nonporous gas exchange membranes enable air gap free and air free devices for standard mammalian cell culture

While the most conceptually straightforward way of exchanging gases between two compartments is via some sort of gap or porous material through which the bulk gas flows freely by diffusion and convection, this is not the only way to achieve gas exchange between two compartments. Indeed, the air gap scheme wherein the gas blend in the interstitial fluid is maintained by direct contact with a gaseous environment is seen almost nowhere in the body. This would be akin to the cells of the brain or the heart being in interstitial fluid gas homeostasis via a hole in the chest or in the head.

While some gas permeable "membranes" in common biological or medical use [102, 121] are simply microporous materials, there also exist non-porous but gas permeable solid materials [124, 70, 103]. We propose to develop new cell culture technologies which eliminate the air gap in cell culture altogether by using these materials.

This approach has precedent in related applications in biology and medicine. Fluorinated ethylene propylene, a mildly gas permeable transparent fluoropolymer has been used to replace the air gap lid of a cell culture vessel to reduce evaporation and contamination for electrophysiological experiments [104]. Modern artificial lung devices used in medicine employ non-porous polymethylpentene hollow fiber membranes for exchange of oxygen and carbon dioxide [129].

We propose two cell culture paradigms making use of this principle. The first is

what we will call “air gap-less cell culture” which achieves gas exchange between media and air through a poorly water permeable, nonporous membrane. In this simpler model, the cell culture does still depend on a surrounding controlled gas blend environment in order to maintain appropriate dissolved gas levels in the cell culture device.

The second is what we will call “incubator-less” or “air-free” cell culture, which does away completely with the use of an air compartment for gas exchange. This relies on the fact that carbonate buffer systems, if sealed from room air, can generate and maintain supraatmospheric dissolved carbon dioxide concentrations given appropriate initial stoichiometries. This concept can further be expanded to control of other dissolved gases such as oxygen by using chemicals which either scavenge or generate the gas in question.

7.1.2 Potential membrane materials

7.1.2.1 Polymethylpentene

poly(4-methyl-1-pentene), informally known as polymethylpentene (PMP) is a thermoplastic polyolefin. It is available commercially as Mitsui TPX. It has the highest oxygen and carbon dioxide permeabilities of any thermoplastics under common use [70]. It is commonly used in artificial respiration devices [129], food packaging and industrial gas separation applications due to its exceptional gas permeability.

It has a high melting point of 228 °C for commercial Mitsui MX004 grade and heat deflection temperature ($\frac{1}{4}$”, 0.45 MPa stress) of 100 °C. It furthermore is resistant to steam. These attributes make it ideal for standard laboratory autoclave sterilization. It is furthermore extremely chemical resistant, in particular, it is resistant to sodium hypochlorite, ozone, per-

acetic acid, isopropanol and ethanol. It is soluble in chloroform, weakly soluble in hexane and it swells in dichloromethane, making it suitable for coating techniques such as spin coating, knife coating or evaporation casting. Its solubility in these common solvents make it possible to bond it to dissimilar plastics despite its low surface energy.

7.1.2.2 Fluoropolymers

The common commercially available fluoropolymers are the class comprised of polytetrafluoroethylene (Teflon), the tetrafluoroethylene hexafluoropropylene copolymer commercially known as fluorinated ethylene propylene (Teflon FEP) and tetrafluoroethylene perfluoroether copolymers commercially known as perfluoroalkoxy alkane (Teflon PFA.) Among these, the latter two are known to have gas permeability sufficient to be used as gas exchange membranes in cell culture systems [104]. PFA is 36 times less permeable to oxygen than polymethylpentene and 50 times less permeable to carbon dioxide. FEP is 42 times less permeable to oxygen and 52 times less permeable to carbon dioxide. However, their advantage to materials such as polymethylpentene is their broad availability as transparent films.

The second, more recent class is the perfluorodioxolanes, commercially available as Teflon AF and Hyflon AD [103]. These have gas permeabilities which rival or even exceed that of polydimethylsiloxane [32]. The principal disadvantage of these materials is that, while they are commercially available, they are neither cost-effective to purchase, nor are they even available commercially in the film formats. They can theoretically be solvent cast (by spin coating, knife coating or otherwise) from perfluorinated solvents into films. At present, due to the lack of availability of films of these materials, they are not practical for use in mass produced

devices; if they become commercially available in the future they are strong candidates for use.

The fluoropolymers also suffer issues with bondability: their high chemical resistance, lack of solvents which can dissolve them or unusual solvents (generally perfluorocarbons, which in turn are nonsolvents for most commonly used engineering plastics) and low surface energy makes bonding by thermal, gluing or solvent means them a potential challenge.

7.1.2.3 Polydimethylsiloxane

Polydimethylsiloxane (PDMS) is an exceptional material for exchange of carbon dioxide and oxygen. Among commercially available materials, its oxygen and carbon dioxide permeability is only remotely rivaled only by other silicones and by the perfluorodioxolane copolymers (eg Teflon AF, Hyflon AD.) However it also has serious drawbacks for cell culture devices. First, it is also permeable to water vapor such that devices using it as an outer shell have been found to produce osmolarity shifts similar to those we see in air-gap systems [48]. In designs where the gas membrane is exposed to air, this partially defeats the purpose of eliminating the air gap. Second, PDMS is difficult to bond to common plastics, making it difficult to build cost effective devices incorporating it solely as gas exchange membrane: conventional PDMS devices are generally made of three dimensional PDMS bonded to flat glass plates. Third, its manufacturing techniques do not lend easily to mass production. Fourth, it acts in a solvent-like fashion towards many small molecules, which has been found to interfere with many types of cell culture. Specifically, it can sequester amino acids, hormones, neurotransmitters, small molecule chemical agents and other similar species, drawing them out of cell culture, interfering with experiments and producing different experimental results from experiments performed in

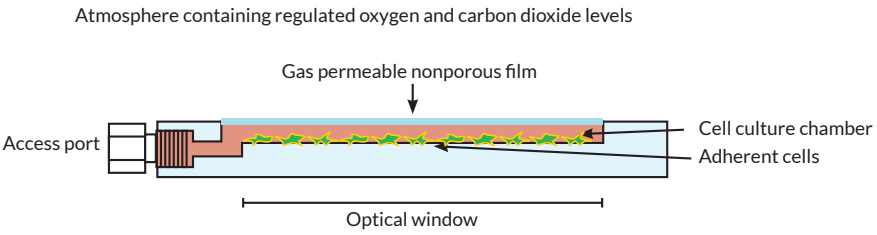
inert thermoplastic devices [105, 85, 119]. Up to 90% partitioning into PDMS has been shown in cell culture experiments [105]. This sequestration also shows capacitance, making time series chemical exposure experiments difficult to impossible [105].

7.2 Basic design of an air-gap free cell culture device

The simplest non air-gap devices extend the general design of air-gap devices by replacing their conduits for physical exchange of air with gas permeable nonporous membranes. We will describe these as air-gap free. There exists considerable precedent in the literature and commercial market for similar devices, however they are generally aimed at bulk cell culture for pharmaceutical manufacturing and the membrane material is used to increase available surface area for gas exchange. In general, these also use water vapor permeable materials for gas exchange and so they are still designed for use in an incubator environment [77, 43, 78, 139, 37, 138].

Even this air gap-less cell culture already provides an advantage over conventional cell culture. While it requires an incubator or equivalent for both carbon dioxide provision and heating, it eliminates the need for antisepsis of the incubating environment, allowing a much simpler incubation setup. Provided that a water vapor impermeable gas exchange material is used, it also eliminates any need for a humidified incubator environment, further simplifying environmental regulation.

2D cell culture air gap free design



3D cell culture air gap free design

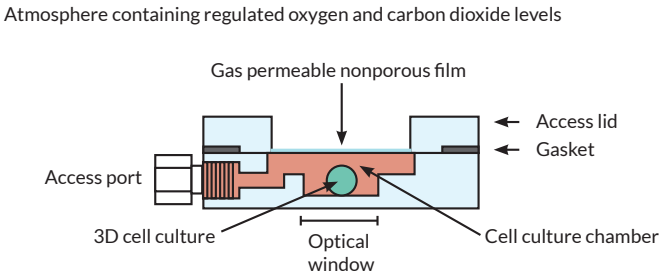


Figure 7.1: **Designs for simple air gap free cell culture devices** Top: A possible design for an air gap free cell culture vessel for 2D adherent cultures. The vessel contains a high aspect ratio chamber to provide a high surface area for cell growth while mimicking the relatively small media volume used in conventional cell culture dishes. The bottom surface of the chamber is made of a material which is suitable for cell culture in terms of adhesion. The upper surface of the chamber is a gas permeable nonporous film which is either inherently low-attachment for cells due to low surface energy or which has been treated to inhibit cell attachment. Note that there is no dead volume of air in the cell culture compartment. Since the device is sealed on all sides, no longer having a lid with an air gap underneath it, access ports are added to the cell culture chamber. Bottom: A possible design for an air gap free cell culture vessel for 3D cultures. This consists of a well which conforms to the 3D culture under study, one wall of which is made of a gas permeable nonporous film. In general it will have a separate, easy to access (such as luer lock) port for media changes as well as a large, removable, gasketed lid to enable access to the large 3D tissue specimen.

7.3 Basic design of an incubator-less cell culture device

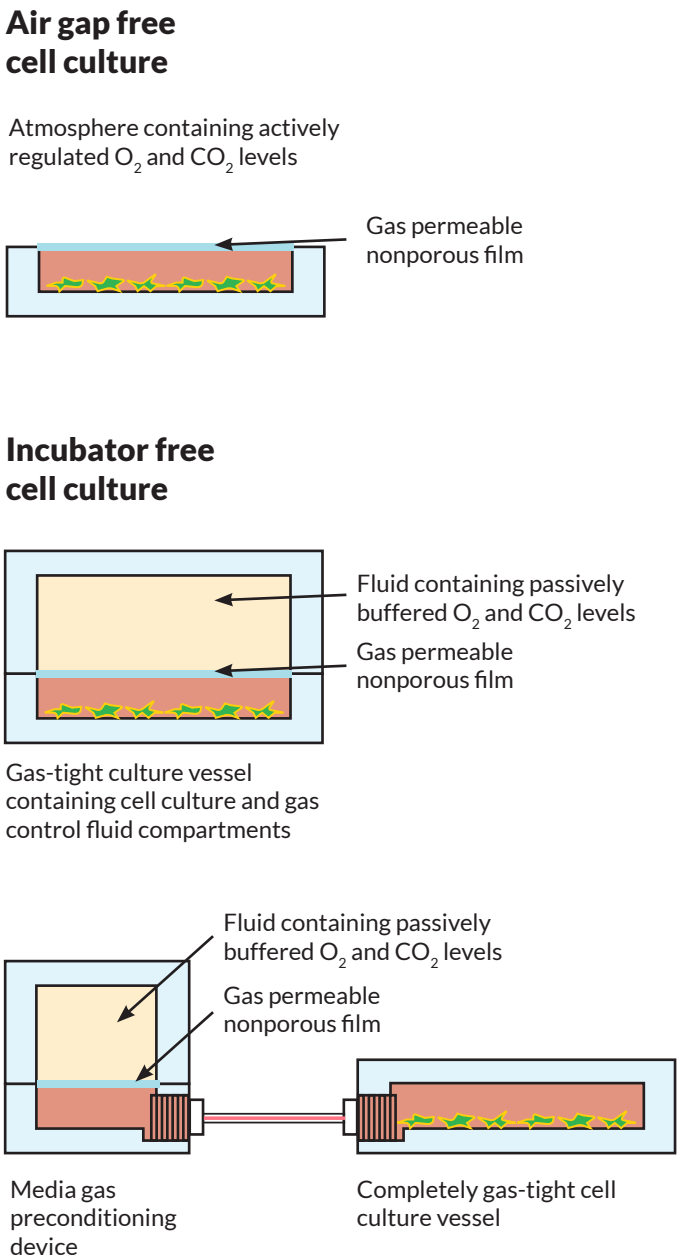


Figure 7.2: **Air gap-less versus incubator-less cell culture** Top: In an air gap free cell culture vessel, a gas permeable nonporous film separates an otherwise sealed cell culture chamber from the outside environment. No water vapor escapes through the film, however the vessel still depends on its surrounding environment having actively regulated oxygen and carbon dioxide levels. The gas atmosphere must also be temperature regulated since gas dissolution equilibria are generally temperature dependent. Middle and bottom: Two examples of possible incubator free cell culture vessels. Middle: An incubator free cell culture vessel comprising two chambers within a gas-tight vessel, separated from each other by a gas permeable nonporous film. A passive buffer system across the film from the cell culture compartment provides dissolved gas regulation, removing the need for a gas and temperature controlled atmosphere surrounding the cell culture vessel. Bottom: A second design for an incubator free cell culture vessel wherein gas exchange happens separately from the cell culture chamber, and media for each media change is instead preconditioned using a gas permeable nonporous film and passive gas control fluid, then automated fluidic actuators supply this preconditioned media to the cell culture chamber.

Beyond air gap free cell culture, we propose that the ideal system from the point of view of cell culture accessibility to instrumentation as well as from the point of view of biological fidelity is a system which eliminates the incubator entirely. Instead, we will take a more biomimetic approach by regulating the gas concentration of the immediate cell culture environment via gas exchange with another aqueous compartment, akin to the gas exchange between interstitial fluid and the bloodstream.

We build on the work of Takano, Tanaka and Futai [125, 126] who demonstrated that using fluid-filled, air-free two-compartments, one for cell culture and one for another fluid, separated by a liquid impermeable but gas permeable material, it is possible to perform cell culture where the dissolved gas regulation to the cell culture compartment is provided entirely by a system of pH and oxygen concentration buffers in the second compartment.

In contrast to our approach, Takano, Tanaka and Futai's systems are built from PDMS and use standard PDMS fabrication and bonding techniques to manufacture. Due to the biochemical disadvantages of a system having PDMS in contact with the cell culture chamber, as described above, as well as the inefficiency of using a system where the exterior walls are also made of a gas permeable material, we seek to develop a device whose shell is made from conventional thermoplastics.

Building upon these works, we propose a device which contains:

- A cell culture compartment, devoid of air
- A gas permeable nonporous polymethylpentene membrane
- A gas control buffer fluid, devoid of air

and whose components, other than the gas permeable membrane, are made from gas impermeable thermoplastics. The lack of air in the system allows for minimal bulk, allowing scalability, simpler control of gas concentrations and mathematical modeling thereof, simpler and more efficient heating and also prevents evaporation-condensation patterns which may occur in cell culture systems which are not uniformly heated. The use of gas impermeable thermoplastics maximizes efficiency by eliminating gas equilibration with the outside environment. It also

allows the devices to be constructed using simple in-house manufacturing techniques such as CNC milling or 3D printing as well as industrial manufacturing techniques such as injection molding.

The figure below illustrates the chemical principles underlying the fluid based gas control scheme which we are adapting from Takano, Tanaka and Futai. It further illustrates the chemical principle by which carbon dioxide (or other gases, such as oxygen) can be transported across the polymethylpentene membrane despite its lack of porosity.

We propose two general approaches to the relative positioning of the cell culture chamber and the gas exchange mechanism. It is possible to colocate the two: the gas exchange happens via one wall of the cell culture chamber. It is also possible to perform the gas exchange at a location distant from the cell culture. For example, one could use recirculated media which flows through an external exchanger. This is media-efficient since it reduces waste but increases and therefore dilutes the effective cell culture media volume. Alternatively, it is possible to pre-condition media prior to injection into the cell culture compartment, then regularly cycle the media to waste. This is media-inefficient but reduces the active media volume "in use" at any given time. The advantage of simple, local gas exchange devices is their simplicity, compactness and lack of need for microfluidic pumps. The advantage of the external gas conditioning designs is that they maximize the available real estate for cell culture observation and manipulation hardware, which is one of the primary goals of developing a new cell culture system for highly instrumented cell culture.

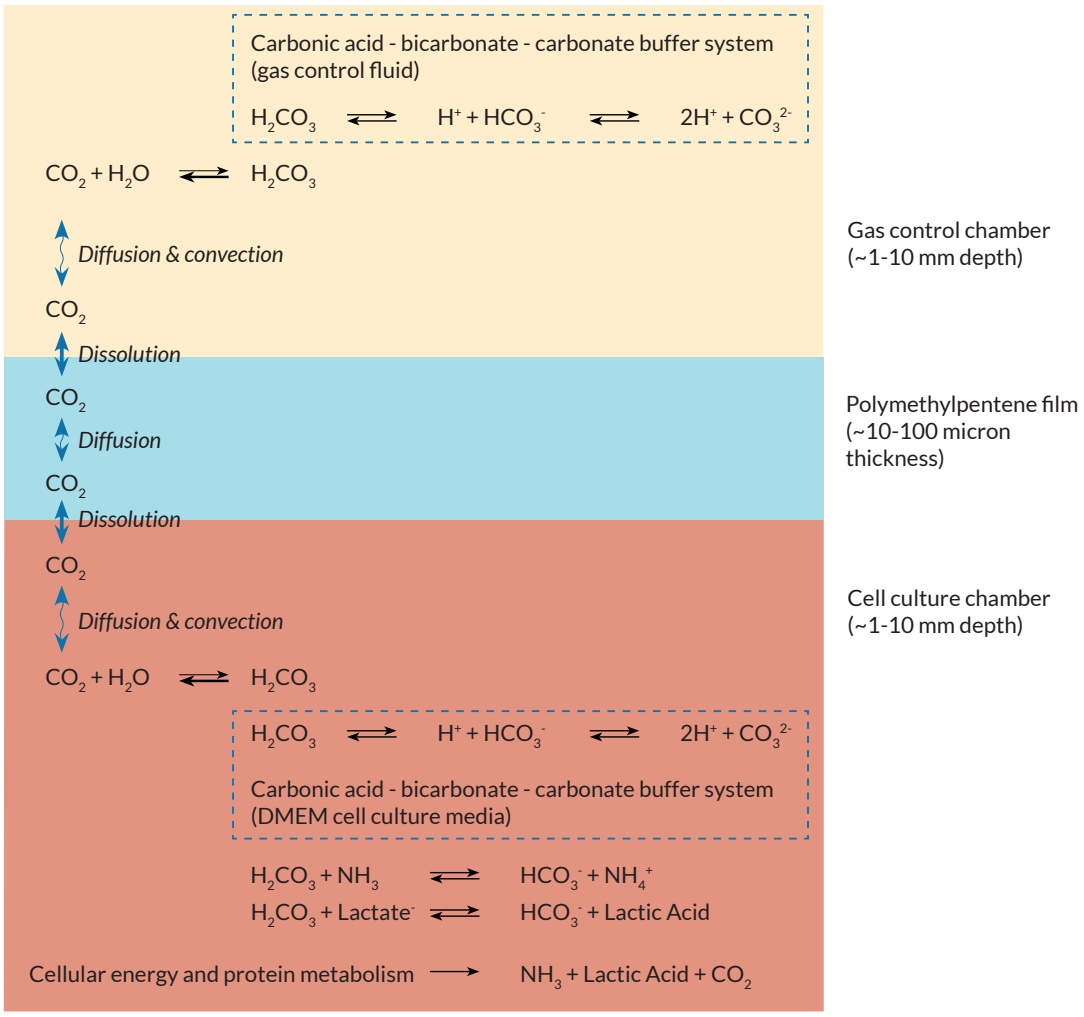

Carbonic acid - bicarbonate - carbonate buffer system
(gas control fluid)
$H_2CO_3 \rightleftharpoons H^+ + HCO_3^- \rightleftharpoons 2H^+ + CO_3^{2-}$
$CO_2 + H_2O \rightleftharpoons H_2CO_3$
Diffusion & convection
CO_2
Dissolution
CO_2
Diffusion
CO_2
Dissolution
CO_2
Diffusion & convection
$CO_2 + H_2O \rightleftharpoons H_2CO_3$
$H_2CO_3 \rightleftharpoons H^+ + HCO_3^- \rightleftharpoons 2H^+ + CO_3^{2-}$
Carbonic acid - bicarbonate - carbonate buffer system
(DMEM cell culture media)
$H_2CO_3 + NH_3 \rightleftharpoons HCO_3^- + NH_4^+$
$H_2CO_3 + Lactate^- \rightleftharpoons HCO_3^- + Lactic\ Acid$
Cellular energy and protein metabolism $\rightarrow NH_3 + Lactic\ Acid + CO_2$
Gas control chamber
(~1-10 mm depth)
Polymethylpentene film
(~10-100 micron
thickness)
Cell culture chamber
(~1-10 mm depth)

Figure 7.3: **Schematic of chemical processes, equilibria and transport involved in incubator-free cell culture** In a gas control fluid chamber, a carbonic acid - bicarbonate - carbonate buffer system is established to provide a buffered concentration of carbonic acid. Carbonic acid is itself in equilibrium with carbon dioxide; at and around physiological pH this equilibrium greatly favors carbon dioxide. Diffusion and convection are transport mechanisms for the dissolved carbon dioxide. At the interface between the gas control fluid and polymethylpentene film layer, the carbon dioxide is in equilibrium between dissolution in the fluid and in the polymethylpentene plastic. Carbon dioxide diffuses freely through the polymethylpentene plastic and on the other side of the film, it is in equilibrium between dissolution in the polymethylpentene plastic and in the cell culture media. It is transported throughout the cell culture media by diffusion and convection. Here, the dissolved carbon dioxide completes the carbonic acid - bicarbonate - carbonate buffer system via its interconversion with carbonic acid. This buffer system maintains physiological pH despite the generation of ammonium, lactic acid and other acid species by cellular metabolism.

7.4 Advantages of air gap-free cell culture

7.4.1 Evaporation

The primary advantage of air gap-free cell culture is in preventing the evaporation which plagues conventional air gap cell culture. The first consequence of preventing evaporation is that cell culture devices no longer need to be kept in a humidified environment. While mixed gas blends with elevated carbon dioxide levels are still necessary to achieve carbonate

bicarbonate buffer pH regulation, this removes one of the major constraints in cell culture which forces the use of incubators or incubator equivalents. The second consequence is that we will be able to maintain stricter osmolarity than even a perfect humidified incubator.

We also make distinction between our designs and polydimethylsiloxane microfluidic devices. While in general, conventional PDMS devices have the same benefits of being able to exchange carbon dioxide and oxygen through a microbially impermeable layer, the material is exceptionally permeable to water vapor, causing widespread issues with osmolarity shifts, potentially causing cell death or clogging of lines if the devices are operated in a heated but nonhumidified environment [48].

7.4.2 Antisepsis

Removing the air gap has significant benefits for antisepsis. Contamination of the cultures by inadvertent introduction of microorganisms or other material is much less possible than with conventional vessels. Specifically, the only ingress is through normally-closed access ports. Furthermore, it is easier to handle access ports aseptically than other devices: they have relatively small surface areas for accidental touch, they can be completely sealed with sterile hardware. Having a sealed, nonporous design also allows for routine decontamination of the outside of cell culture vessels. In an air-gap design, the conventional laboratory decontamination methods of spraying the exterior of objects with 70% alcohol or with dilute bleach cannot be done without exposing the contents to mists of these chemicals, which are also cytotoxic to the cells under study.

This also means that theoretically, secondary containment is not needed for handling,

transport or even storage of live cell cultures. Provided biosafety requirements are met, it would theoretically be possible to perform cell culture outside of a laboratory environment without any need for containment. This has immense practical applications for transport and for laboratory design.

It would theoretically be feasible to build laboratories where only the areas used for active manipulation of cultures would need to be kept clean since only actively opened cell culture devices would be susceptible to contamination. This would potentially make laboratory design both cheaper and more flexible. This could also potentially allow for more cell culture in resource-limited settings.

Furthermore, to take the concept of microbial barrier a step further, it would allow a particularly efficient method of heat maintenance: direct immersion of cell culture vessels in a water bath instead of a closed sterile incubator. Water has a specific heat capacity which rivals metals but at a much lower price. In addition, it has the advantage of convection as a heat transfer mechanism, which allows for very uniform heat distributions, even in complex geometries. In comparison, solids which rely only on conductive transport of heat have their evenness of heating highly dependent on geometry.

While the necessary validation is beyond the scope of this work, we also propose that an air gap-free cell culture ecosystem would provide an additional engineering control against direct exposure of laboratory workers to microorganisms from inside the cell culture device, potential release of these microorganisms into the immediate laboratory environment or beyond or against cross-contamination between microorganisms contained in separate vessels.

7.5 Advantages of incubator-free cell culture

The principal advantage of incubator-free cell culture is, as discussed above, to remove impediments to using external devices such as microscopes, electrophysiology equipment, optical or electrical sensors or other microfluidic liquid sampling and assay devices in concert with cell culture. As described, some of these impediments arise from the physical bulk and inflexibility of the cell culture incubator, which is not suited to observation of the cell cultures contained within using arbitrary instrumentation. Some of these impediments are also related to the challenges of using devices to observe the cells while still maintaining desired gas concentrations, osmolarity, temperature and sterility. Incubator-free cell culture removes the physical bulk of the incubator and also allows factors such as temperature or gas concentration to be controlled with devices (such as heat blocks or gas exchange modules) whose positioning can easily be adjusted to accommodate instrumentation.

The second set of advantages, which are in a sense an accidental consequence of the idea of getting rid of the troubles caused by the incubator, is that incubator-free cell culture can be more precise or versatile than conventional air gap cell culture when it comes to controlling gas concentrations or other environmental factors of the cells. Better precision and environmental fidelity can be achieved by eliminating the gas, humidity and temperature leakage inherent to the air gap design. But using a sealed, gas control fluid based system also allows simplicity and even flexibility in terms of factors such as chemical composition, temperature and pressure.

Using a liquid solution rather than an air chamber to provide dissolved gas blends to the cell culture compartment can simplify the maintenance of target gas concentrations. It is im-

portant to note that the dissolved gas concentration in this liquid solution can be in equilibrium with a chemical system that can include non-gas components. For example, carbon dioxide dissolved gas concentrations can be maintained by using an aqueous carbonic acid - bicarbonate - carbonate buffer system as a carbon dioxide generator. In this system, the non-gas species carbonic acid is in turn in equilibrium with its reversible decomposition to carbon dioxide and the equilibrium vastly favors carbon dioxide at and around standard temperature and pressure. This allows the concentration of carbon dioxide gas to be maintained by participation in a chemical buffer system. This is much more simple and cost effective than the two standard methods of maintaining carbon dioxide concentrations different from room air. The first, more common method is to use a carbon dioxide sensor and active metering of compressed carbon dioxide gas to maintain a gas concentration setpoint. The sensors and actuators involved are expensive, complex and susceptible to failure. The second method is to use actively replaced premixed compressed gas blends. While this does not depend on expensive and failure-prone sensors and controllers, it is not cost effective to routinely purchase and continually vent custom blends of compressed gas.

The same principles apply to using a liquid gas control solution to provide lower than room air (physoxic, hypoxic, anoxic) oxygen concentrations. Rather than active feedback control of oxygen, oxygen can be controlled by one of two methods. The first, applicable to physoxic and hypoxic conditions, is the ascorbate scavenger system described by Takano et al which displays its own buffering behavoir, allowing long-term maintenance of lower than room air oxygen concentrations without feedback control. For anoxia, a scavenger such as sodium sulfite can be used.

We also note that our sealed system can function with gas blends other than those commonly used in standard cell culture. It could conceivably achieve gas blends which are impossible outside of costly and impractical fully sealed glovebox culture systems. For example, it would be possible to study cells in carbon dioxide buffered yet completely anoxic conditions. In glovebox systems, low concentrations of oxygen are possible using exuberant nitrogen flushes and access through airlock but this sort of cell culture is cost-ineffective and difficult for technicians to manipulate. Furthermore, any brief exposure to an oxygen containing environment (such as during feeding or other manipulations) may cause oxygen to dissolve. This cannot be easily removed: negative pressure will pull other gases out of solution, sparging is impractical and likely to disturb the culture irreparably and oxygen scavenging chemical agents are toxic to cell cultures. We propose a system where an oxygen scavenger is used on the other side of the membrane in order to draw all oxygen out of the system. Takano et al [126] demonstrated a version of this principle with their hypoxic system where ascorbate in a buffered pH system acts as an oxygen scavenger with slow kinetics and is thereby is used to create and maintain a stable low oxygen environment. We propose to extend this concept by also using species such as cobalt-catalyzed sodium sulfite to completely eliminate oxygenation.

Hypoxic culture is important for research because relative hypoxia (to 5% or lower) is required to replicate the interstitial oxygen tensions of peripheral organs such as skeletal muscle [137] or brain [17]. Anoxic cell culture is used to study cellular behavior under pathological anoxic states such as anoxic injury to the myocardium (in myocardial infarctions) or brain (in stroke) as well as physiological anoxic states such as the gut luminal microbiome [54].

The cell culture principles underlying incubator-free cell culture are also theoretically

possible with water-soluble gases other than oxygen, nitrogen and carbon dioxide. Toxicological studies of dissolved gases such as carbon monoxide, ammonia, methane or nitrous oxide would be possible without the risk of releasing these species into the laboratory environment. These sorts of cell culture are currently possible, but only in a glove box environment. Similarly, cell culture in the presence of volatile toxins such as alcohols, organophosphates or organic solvents would be possible and safe in an incubator-free system.

The closed system also allows us to study tissue culture models under hydrostatic pressures other than atmospheric pressure. This could be important in the study of important disease states involving elevated pressures in tissue compartments: among them neoplasms, pericardial effusion, certain modes of heart failure, brain inflammation in the context of stroke, infection or seizure, among others. It is also well known that many tissues have hydrostatic pressures other than atmospheric pressure under normal conditions. Therefore, incubator-free cell culture may also allow for more realistic models of organ systems from a pressure point of view, even for organs in their healthy state.

Chapter 8

Cell culture devices for air gap-free and incubator-free 2D cell culture

8.1 Abstract

As our first proof of concept, we demonstrate that simple two-dimensional cell culture is possible using devices which are air gap free or incubator free. We use in-laboratory manufacturing methods to fabricate air gap free and incubator free devices from cytocompatible materials. We build a heat controller to allow for benchtop incubator free culture with these devices.

We will demonstrate that

- "Incubator-free" membrane gas exchange devices can provide stable pH control via maintenance of a carbon dioxide buffer system over several days without the use of any mixed gas blends or other active control.

- All membrane gas exchange devices provide consistent sodium levels as compared to the evaporation-induced hypernatremic conditions seen in standard air-gap cell culture devices.
- "Incubator-free" devices truly allow for cell culture without any incubator by performing cell culture with a heat block
- These devices can be easily manufactured using in-laboratory desktop CNC milling and bonding operations.

8.2 Design

In making our proof of concept labware, our goal was to design a cell culture vessel that is as comparable as possible in terms of culture area and volume to conventional cell culture vessels as to allow the most direct methodological comparison without other design choices as confounding variables. We chose to specifically design our parameters to match the prototypical airgap cell culture device, the cell culture plate.

Our device contains a cell culture compartment with a base surface area of 21.5 cm^2 and a cell culture compartment depth of 2 mm in order to closely match the well surface area and recommended media volume of standard 60 mm plates. However, unlike the conventional air gap plate, where a dead space and a lid with air gap lie above the top of the media, our design uses a gas permeable film placed directly contacting the media without any intervening air filling that space. Apart from this gas permeable film, there is no other site of ingress or egress of gas from the device: the walls are thick, gas impermeable polycarbonate and access

for cell seating or media change is through sealable airtight Luer lock ports. Note that our film material, polymethylpentene, has an exceptionally low surface energy such that even though the immersion of the film surface in the media presents an additional potential surface on which for cells to adhere, adhesion is energetically unfavorable. (This is in fact true of most non-silicone gas permeable films.)

In the version of this device used to test the principle of simple air gap free cell culture, the film is held in place by laminating between the bottom polycarbonate shell of the cell culture compartment and an upper polycarbonate bumper which also serves to protect the film. In the fully incubator free version of this device, this film is laminated between the lower cell culture compartment previously described and an upper compartment which is functionally comparable other than having a volume several times in excess of the lower cell culture compartment.

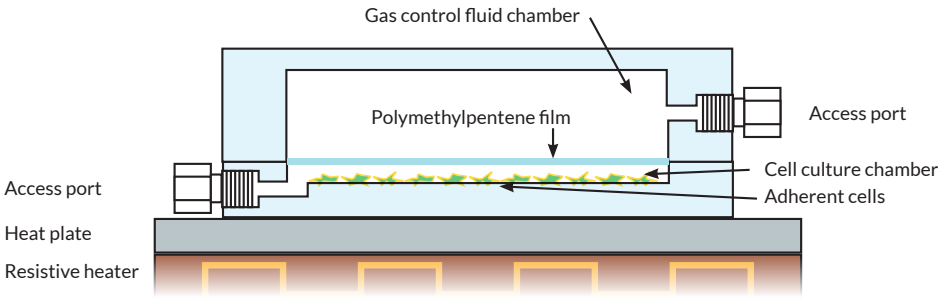


Figure 8.1: **General design of minimal 2D incubator-free cell culture device.**

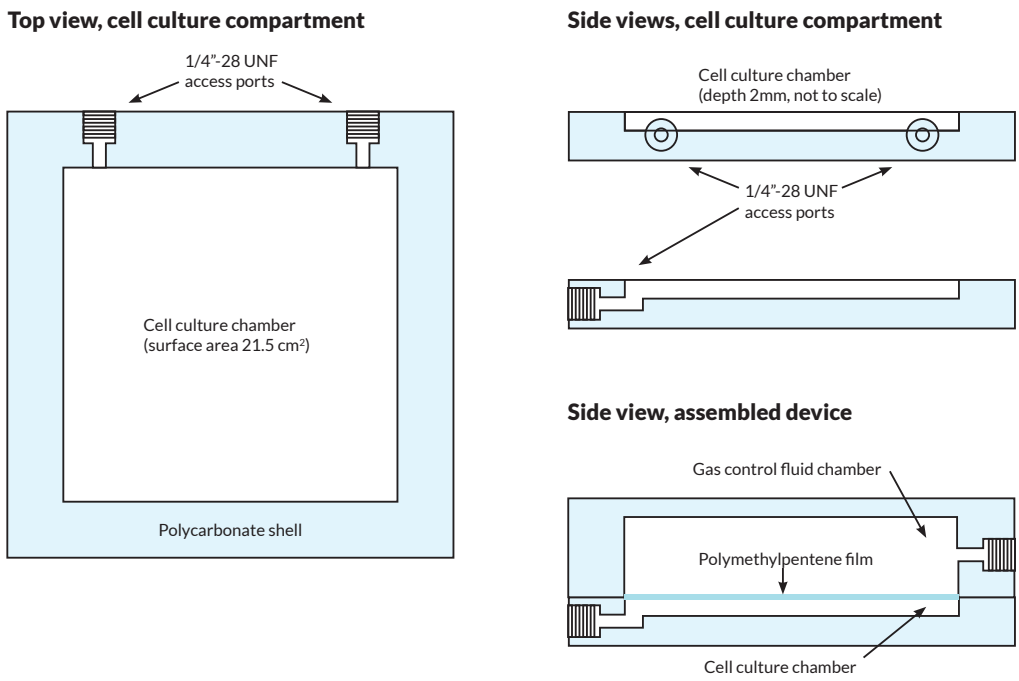


Figure 8.2: **Detailed design of components of our minimal 2D incubator-free cell culture device** Left panel (top view) shows the layout of the chamber along the transverse plane, showing shape of chamber and positioning of access ports. Top right panel (side views) shows a schematic of the shallow depth of the cell culture chamber and the positioning of the access ports viewed head-on (top) and side-on (bottom.) Note that the fluidic channels from the ports enter the cell culture chamber from its base due to the shallow depth of the chamber versus the clearance needed to insert the threaded ports from the side of the device. Bottom right panel shows a schematic view of the assembled device, consisting of the cell culture chamber, a similar but larger gas control fluid chamber, and a polymethylpentene film laminated between the polycarbonate shells of the two chambers.

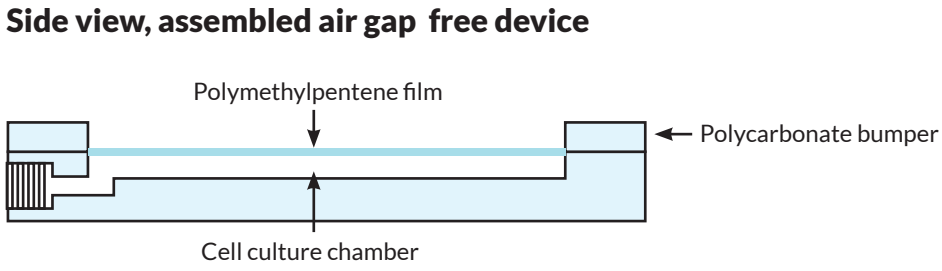


Figure 8.3: **Design of minimal 2D air gap free cell culture device** This is a straightforward modification of the incubator-free design, the difference being that the polymethylpentene film is laminated between the cell culture chamber subcomponent and a polycarbonate bumper rather than the cell culture chamber subcomponent and the gas control fluid chamber subcomponent.

8.3 Experimental design

The fundamental questions which we aim to answer are whether non-air gap devices are viable options for cell culture, whether they achieve physiological noninferiority to standard airgap devices and whether they have osmolarity benefits over airgap cell culture devices.

We will specifically compare our air gap and incubator free cell culture devices against the prototypical air gap cell culture devices: the cell culture dish and the vented cell culture flask. Specifically, in order to match the cell growth surface area of our devices, we will use the 60 mm cell culture dish and the T25 format cell culture flask with a 0.2 μm microporous filter vent. We will also include non-gas exchange controls of a T25 flask with sealed "plug" cap (which, while lacking gas exchange with the incubator environment does include a trapped gas volume exceeding the media volume with which to exchange gases) as well as a version

of our cell culture devices which is completely sealed to gas exchange on all sides rather than incorporating any membrane.

The first question, of whether we can use our non-air gap devices to achieve cell culture comparable in performance to airgap devices, will be addressed by making comparisons of viability and growth rate over the course of a multi-day cell culture experiment with comparisons between all experimental conditions. The second, of whether the devices achieve parity in physiological variables, and the third, of whether the non-airgap devices outperform airgap devices and osmolarity will be addressed by taking time point series measurements of physiological bioanalytes over a multi-day cell culture experiment.

Our experiments will be performed using a standard and fast-growing cell line, the NIH/3T3 murine embryonic fibroblast. We will use standard cell culture conditions, using Gibco brand Dulbecco's Modified Eagle Medium, a media which is representative of common media blends used in mammalian cell culture.

8.4 Device manufacturing

The cell culture chambers were prepared by a combination of CNC milling, manual drilling/reaming, manual thread tapping and solvent welding from polycarbonate sheet stock and polymethylpentene film stock. For production scale, injection molding, automated CNC milling and thermal welding could be used to fabricate equivalent vessels.

8.4.1 Construction of shell

8.4.1.1 Materials

Clear polycarbonate sheet stock (Plaskolite Tuffak or Covestro Makrolon brand) was obtained from McMaster Carr. Dichloromethane (ACS grade, 99.5%) was obtained from Sigma-Aldrich. 3 mil polymethylpentene film stock (Mitsui TPX) was obtained from CS Hyde and 2 mil transparent polymethylpentene film stock (Mitsui TPX Opulent) was obtained from Mitsui.

8.4.1.2 CNC milling procedure

All milling steps were done on a Carbide3D Nomad 3 desktop CNC milling machine. Toolpaths were generated in Carbide Create version 6. XY-axis zeroing was performed using a LCA mini LC/EF spindle mounted laser centering-edge finding tool and Z-axis zeroing was performed manually. Fixturing for basic cutting operations was achieved using Shurtape DS 154 double sided tape on an medium-density fiberboard (MDF) wasteboard which was faced using a fly cutter prior to use. Fixturing for side port cutting operations was performed using Carbide 3D Tiger Claw clamps on a worktable. For precise positioning and squaring to perform reverse milling operations, a right angle corner alignment jig was machined from MDF.

Milling cutters used were the Harvey Tool Nº 52808 $\frac{1}{8}$" diameter, $\frac{5}{8}$" length of cut square upcut single flute endmill for upcut operations and the Harvey Tool Nº 935408 $\frac{1}{8}$" diameter, $\frac{5}{8}$" length of cut square downcut single flute endmill for downcut operations.

8.4.1.3 Components

The first subcomponent prepared was the cell culture compartment. For this, the stock used was nominally $\frac{3}{8}$" thick clear polycarbonate sheet. To account for thickness deviations in the stock, the stock was faced to a thickness of 9 mm prior to use. The obverse side was milled first, using the upcut endmill.

Subsequently, the part was fixtured with its reverse side downwards using the reverse-cutting jig and a pocket was cut on the reverse side using the downcut endmill. A downcut geometry was used in order to counteract lifting forces: we found that using an upcut endmill would lift our our workpieces from their holding tape due to the reduced surface area left after the initial milling operation. The pocket was designed as to leave a 2 mm tall ledge, which when joined with the second component, would form a 2 mm deep cavity for cell culture. The second subcomponent prepared was the window for microscopy. This is a clear slab of polycarbonate which fits into the pocket milled in the first part. This was routed from a nominally $\frac{1}{4}$" polycarbonate sheet. No facing operation was performed since thickness of this part was not critical to the part but facing would ruin the transparency of the window. The reverse side of the part was marked with a blade for orientation when the part was assembled. This is in order to not expose the cells being cultured to the same surface which had been in contact with the double sided tape in order to reduce the possibility of contamination of the culture by the adhesive from the tape.

The two parts were then assembled by solvent welding. Solvent welding operations were performed by wetting opposed surfaces with dichloromethane via capillary action, sup-

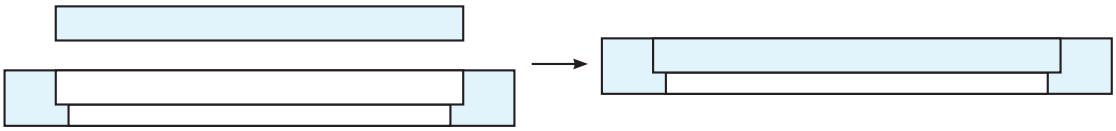
plied via a disposable 27 gauge needle and a Hamilton Nº 1001 TLL glass/polytetrafluoroethylene gastight syringe. Disposable syringes were not used for fear of contamination of cell culture contact surfaces, since swelling of disposable syringe (BD PlastiPak) plunger elastomer material was observed on exposure to dichloromethane.

The design of the first and second subcomponents incorporates a 25 μm gap on all mating sides to allow for capillary action to form a solvent bond between opposed polycarbonate surfaces. The second part was inserted into the pocket of the first, and pressed so that it fully rested on the lip. Solvent was flowed into the gap and bonding was confirmed by viewing under tangential light.

Obverse milling, first component

Reverse milling, first component

Insertion of second component



Bonding of second component

Figure 8.4: **Schematic of two-piece assembly of cell culture device half-shells containing a cavity with a clear window** In the first obverse milling step, the workpiece is faced to create a bonding surface and pocketed to create the cavity for the cell culture compartment. In the subsequent reverse milling step, the part is fixtured on its other side to insert a pocket, in which a separately milled clear window is inserted, leaving only 2 mm of the initial pocket. Finally, the two components are solvent welded along their edges, yielding a single polycarbonate part incorporating a clear windowed cavity.

Two other versions of the part were made, using different techniques to produce a transparent window for cell culture. In both, the cell culture chamber was formed by a depression machined to a depth of 2 mm in a single sheet of polycarbonate. In the first, in order to produce a transparent surface, a Datron Nº 0078330E 3 mm diameter polished single-flute end-mill was used for a pre-finishing operation followed by a final finishing operation using an LMT Onsrud Nº 75-110 polycrystalline diamond polishing tool router bit. This produced a surface which was grossly transparent but with sufficient chatter marks from the poor rigidity of the Nomad 3 mill that microscopy was difficult. We recommend this method for laboratories which have access to a rigid and powerful enough mill to maintain the necessary feedrate demanded by the polishing tool without generating significant chatter.

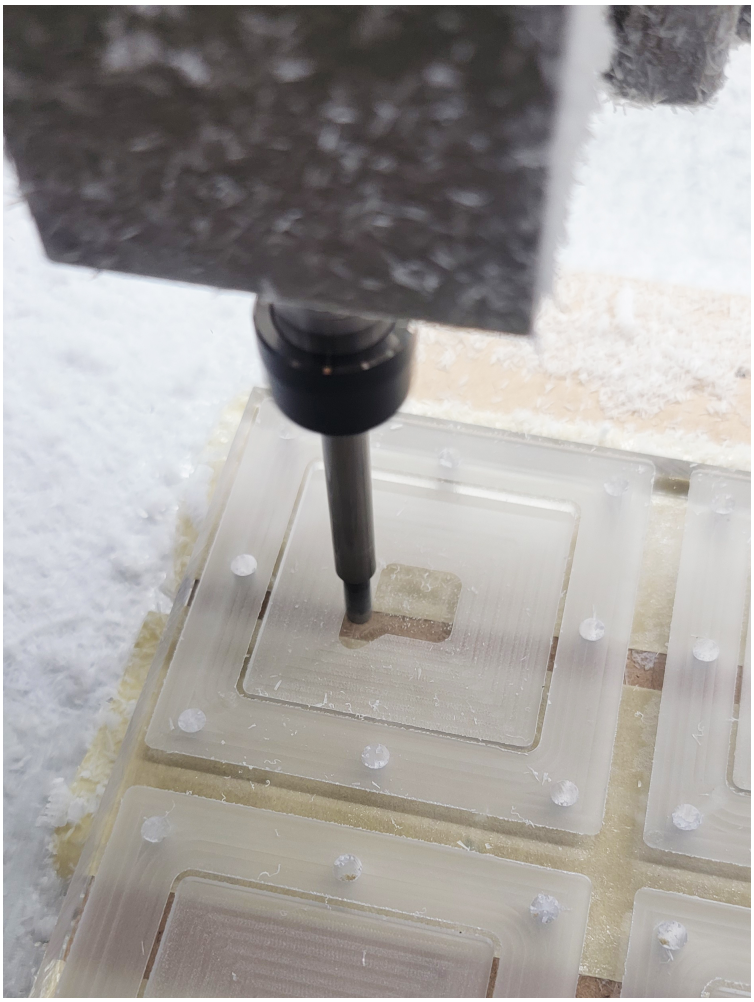


Figure 8.5: **Results of fabrication of optical window using CNC polishing tool** Left: the LMT Onsrud polycrystalline diamond polishing tool in action as a finishing pass on the Carbide3D Nomad 3 CNC mill. Right: Image under tangential illumination shows an unacceptably rough surface finish including grooving and ringing patterns, most likely due to vibration of the consumer-grade Nomad 3 CNC mill.

In the second method the pocket was roughed but not finished, then a mist of dichloromethane was applied with a spray bottle to solvent-smooth to a transparent finish. This produced a

serviceable transparent finish slightly inferior to the two-piece window version. The solvent welded method may also require a longer degassing time due to the additional solvent load.

Subsequently, a second upper chamber for gas control fluid was milled in a single piece from nominally $\frac{1}{2}$" polycarbonate sheet faced to 12 mm thickness.

8.4.2 Construction of fluidic ports

The parts were fixtured side-on as previously described. First a pair of $\frac{1}{4}$" tubular channels were drilled by peck drilling using the upcut endmill, and then a pair of 5.41 mm diameter tubular channels concentric with the first were milled.

$\frac{1}{4}$"-28 UNF thread fluidic ports were chosen for fluidic access due to their versatility and ease of construction by subtractive manufacturing as opposed to luer lock connectors. They are compatible with both Luer connectors via $\frac{1}{4}$"-28 UNF to luer lock adapters as well as to standard chromatography fittings such as compression ferrule systems.

The parts were then fixtured side-on to the work table of a drill press using a bench vise with soft jaws. The upper channels were then manually tapped to $\frac{1}{4}$"-28 UNF threads as described above. Lastly, a drill press and a clean $\frac{1}{4}$" drill bit was used on the cell culture face of the assembly to link the threaded channel to the cell culture cavity.

Thread tapping was performed manually using a Tapmatic Nº 13033 RX 30 self-reversing tapping head and an uncoated high-speed steel closed-end bottoming chamfer $\frac{1}{4}$"-28 UNF thread chip-clearing tap (McMaster Carr) on a Craftsman benchtop drill press running at its minimum speed of 610 rpm. Workholding was with a machine vise and soft jaws. Where required, tapped threads were cleaned using a dental pick and compressed air.

Polycarbonate $\frac{1}{4}$"-28 UNF male thread to female luer lock fittings (Masterflex Nº 45501-66) and polyvinylidene fluoride male luer lock plugs (Masterflex Nº 45513-56) were obtained from Cole-Parmer.

The $\frac{1}{4}$"-28 UNF male thread to female luer lock fittings were installed in the $\frac{1}{4}$"-28 UNF threaded holes and tightened with a ratcheting wrench. A drop of dichloromethane was added to the threads to seal and prevent leaks. This was found to be necessary for a leakproof seal for polycarbonate-on-polycarbonate fitting-to-body threaded interfaces, since the rigid on rigid threads did not provide a leakproof seal on their own. This would be unnecessary if a softer material such as polypropylene is used; polypropylene was not used in this case due to its lower stability to ozone sterilization.

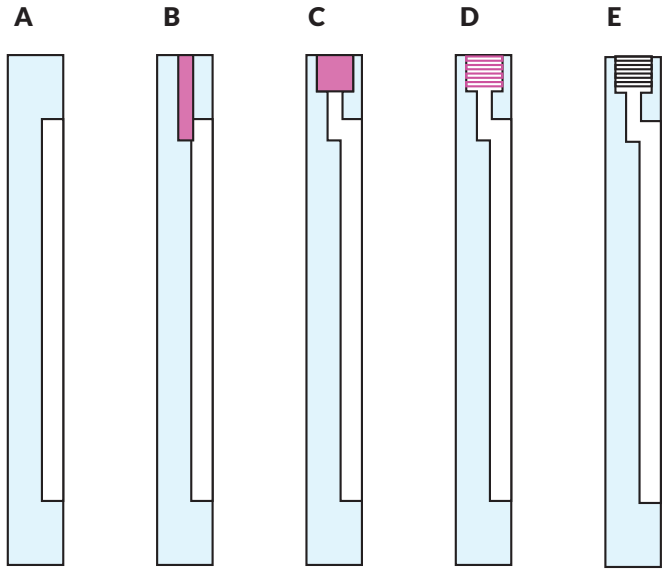


Figure 8.6: **Side access port manufacturing steps** (A) Part is fixtured side-on. (B) A $\frac{1}{8}$" access channel is milled. (C) A 5.41 mm pilot hole for threading is milled. (D) The pilot hole is threaded to $\frac{1}{4}$"-28 UNF specifications using a machine tap and tapping apparatus. (E) The final part now contains a side channel with $\frac{1}{4}$"-28 UNF fluidic fitting in direct communication with its interior cavity.

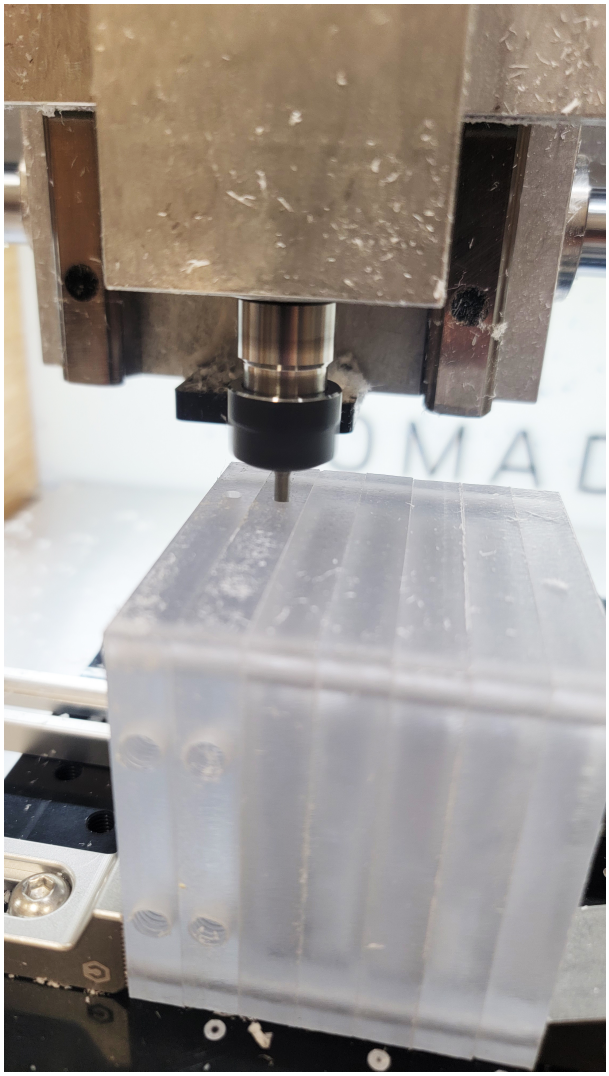


Figure 8.7: **Side milling procedure** A set of parts is fixtured side-on using Carbide3D Tiger Claw clamps. A Nomad3 CNC mill is used to mill consistently sized and positioned side channels in this set of parts parts using a square endmill.

8.4.3 Membrane installation and final assembly

Where required, parts were deburred using an S-blade deburring tool and a razor scraper. Debris and dust was coarsely removed from parts using compressed air. Prior to closure of devices, all cell culture contact surfaces were cleaned with Texwipe Nº TX3273 sterile 70% USP-grade isopropanol and Kimtech low-linting non-abrasive wipes and allowed to briefly air-dry. Prior to solvent welding, surfaces were cleaned of remaining milling debris and fingerprints by this same procedure to prevent particulate from interfering with the bond. A cleanroom environment was deemed unnecessary for bonding since the bonding procedure is tolerant to small (sub-25 μm) particulate. Such an environment was also deemed unnecessary for cell culture surfaces since equal to better cleanliness can be achieved by flushing these clear prior to use in cell culture.

TPX polymethylpentene films were cut to size by hand with a razor blade, to a size equal to the area cell culture cavity, plus a 3–5 mm margin on all sides. The film was placed over the cell culture cavity of a freshly cleaned cell culture side assembly. The glossy side was used for the cell contact surface since, per manufacturer communication, the matte side is in contact with a rubber roller during the sheet manufacturing process, which risks transfer of potentially contaminants to the matte surface. Dichloromethane was flowed under the edge of the membrane by hand to spread via capillary action and glue the membrane in place.

Dichloromethane was applied to and around the remaining surface of the TPX film to wet the whole surface. The two halves of the assembly were opposed under pressure for 15 seconds. Dichloromethane was flowed into the remaining gap by capillary action and the parts were allowed to de-gas.

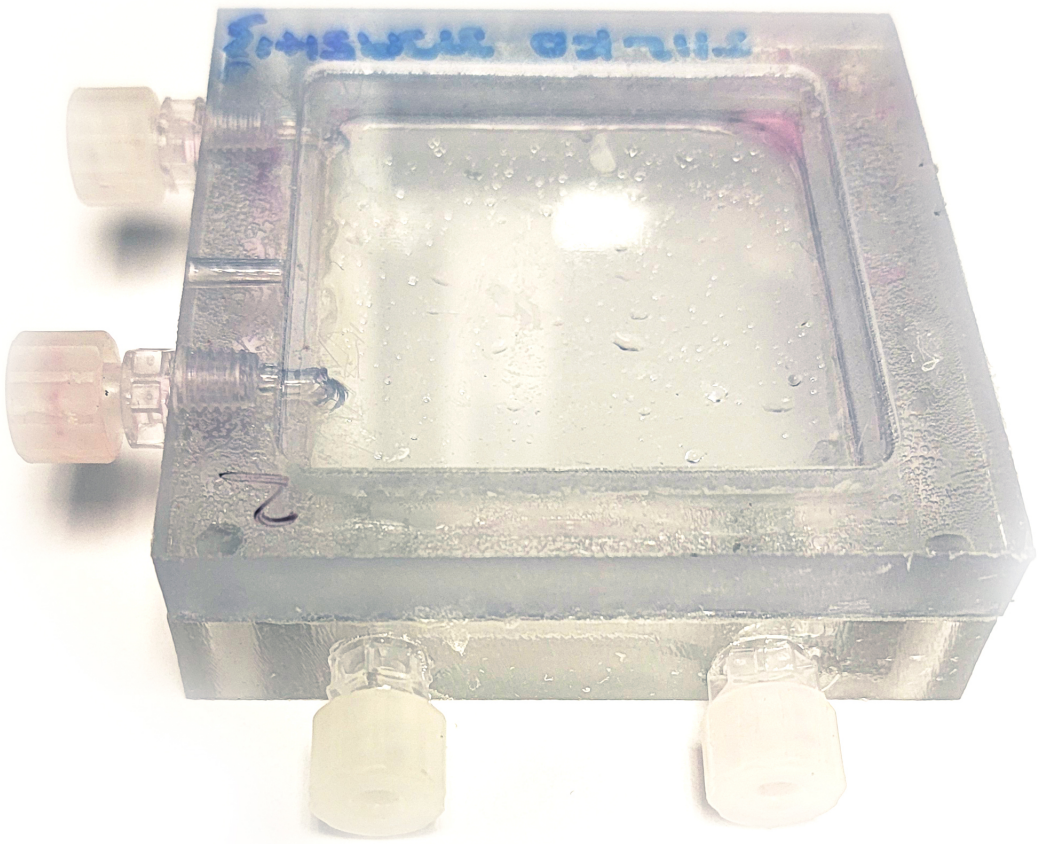


Figure 8.8: **An assembled incubator-free 2D cell culture device** This is a 2D cell culture device, shown after it has been used for cell culture. The cell culture chamber is on top. The inset optically clear window can be seen, showing a clear view of the residual liquid in its cavity.

8.4.4 Post-processing and sterilization

After solvent welding, parts were degassed in a vacuum chamber for at least 4 hours at room temperature and at vacuum of 25 inHg below atmospheric pressure. Heating was not used for this initial degassing due to bubbling of small amounts of retained solvent distorting the part at the join. Similarly, heat-based sterilization methods were avoided for solvent-welded parts.

Parts, which had been prepared without cleanroom or aseptic precautions, were grossly decontaminated on all of their cell culture surfaces and other wetted surfaces prior to sterilization using 70% isopropyl alcohol and KimTech low-linting wipes. Due to temperature limitations caused by solvent welding, ozone sterilization [28] was used as our low temperature sterilization modality of choice.

PlastCareUSA self-seal sterilization paper-backed pouches were used for sterile part containment during and after sterilization. These pouches are designed for autoclave or ethylene oxide sterilization but are also ozone permeable. An Ozilla Ozone Sterilizer was used as ozone source and a plastic storage bin was used for containment. Ozone sterilization was performed outdoors for ventilation in case of a potentially hazardous ozone leak from the containment vessel. A 2 hour ozonation time and a 30 minute scrub cycle was used.

8.5 Heating system design and assembly

8.5.1 Design and assembly of heat controller

We built a cell culture device heating system based on a commercial industrial temperature controller, aluminum heat block and resistive heater. While laboratory heaters are commercially available, ones with the necessary safety features for unattended operation are generally expensive: the entry level IKA RCT Basic hotplate is $ 1132 USD and the Ohaus Guardian 5000 hotplate is $ 500 USD at the time of writing. Furthermore, these are integrated hotplate units rather than systems which are able to control the temperature of heated devices of arbitrary shapes and sizes.

Safety was key to the design of the heating unit, since a thermal runaway condition caused by a failure of the control unit could lead to sample loss in the best case and fire in the worst case. Safety is even more of a concern since part of the design philosophy of our systems is to be scalable and automated as to supplant human labor in cell culture to then enable us to operate more replicates of equipment than can be monitored by a human. Therefore, the system must be designed from the ground up to not cause laboratory catastrophes.

The worst outcomes are those which either threaten human life or threaten entire laboratories. The former is less of a concern in daily operation given that the system is designed to be run without humans physically interacting with it, however, a system cannot be considered safe if it runs any risk of causing a fire in a building occupied by humans.

Some options for heat control are obviously inappropriate: First, lower cost feedback control hot plates generally completely lack the control firmware and fuses to detect and or

abort a thermal runaway condition. Second, while it would be in the ”maker” spirit of our other work to use low cost commodity microcontrollers or temperature control units, these are not fail safe. It is important to use devices which have been tested and rated for long term critical application.

The basic principles of design for safety are:

- To be able to detect and abort a loop break condition where the thermal sensor is failing to detect changes in temperature. This sort of condition can be caused by a wiring failure on the temperature sensor, a detachment of the temperature sensor from the heated device or a malfunction of the temperature sensor.

- To be able to detect and abort an unintended rise in temperature of the device such as from a short circuit or a failure of a relay in the closed position. This can be implemented in firmware with output to a signal operated circuit breaker, or this can be implemented using a thermal fuse. The latter is the cheaper and more fail safe option but the least customizable and resettable.

- For the control logic to be robust to network failure, network congestion and competing tasks in the overall device. This demands that local feedback control with a completely hardwired system be used and that the local feedback controller is independent from other functions of the overall cell culture system.

According to these principles, we have designed our control and power unit with five key safety decisions. The main power supply to the entire system has a fuse to cut power in the case of a short circuit. Temperature control is operated from a dedicated industrial temperature

controller with built-in safety functions including loop break detection. A high quality solid state relay is used for switching power to the resistive heater to minimize the chances of failure. The temperature sensor is rigidly mounted to the heat block to minimize the chance of detaching causing a loop break condition. A thermal fuse is used to irreversibly cut power in case of firmware or relay failure.

8.5.2 Components of heat controller

Omega Nº CN32PT-220-DC process controllers, Omega Nº CN401-11445 PID process controllers or Novus Nº N1040-PR USB 24V process controllers were used for feedback control of temperature. These were chosen due to their reliability and safety features: in particular, it was important to choose controllers from legitimate industrial suppliers which include firmware loop break detection. The Omega Nº CN32PT-220-DC and Novus Nº N1040-PR USB controllers are suitable for integration with automated systems; the former was used in early experiments due to its full featured software monitoring capabilities.

The input to the process controllers was via type K thermocouples (Omega Nº TC-PVC-K-24-180 or Ecubmaker M4 screw mount Type K thermocouple) which were mounted to the heat block by either thermal adhesive (MG Chemicals Nº 8329TCM) or tapped M4 screw hole.

The process controllers controlled a direct current, normally open solid state relay (either a Sensata-Crydom Nº DC60S3 60V 3A DC-DC relay or Senstata-Crydom Nº GN 84134750 60V 10A DC-DC relay.) A solid state relay was chosen over an electromechanical relay for switching due to its long lifetime.

The circuit was protected against device overheating primarily by the electronic safety routines of the PID process controller and secondarily by a thermal fuse (Cantherm, 72 °C AC-DC 10A axial thermal fuse) which was mounted to the heat plate assembly. The latter provides only protection against catastrophic overheating events caused by failures of the PID controller, relay and/or thermocouple in the case that primary protection also failed. The former provides protection against both these catastrophic events (by detecting failures of the thermocouple to detect changes in temperature or failures of the solid state relay to close) as well as protection against overheating to lower temperatures in the range of > 41 °C which would not damage the system components but would place experiment-ruining thermal stresses on the cells themselves.

Except where geometrically impossible, commercial resistive heaters were used as heating elements. Other than supplying appropriate voltage and ensuring that the relay chosen is suitable for the current of the heater, the system is agnostic to the resistive heater chosen. If the calculated resistance of the heaters was found to exceed the rating of the solid state relay used, a current limiting resistor was added to the circuit.

Heat blocks were manufactured in manners described below. For insulation, 1" thick 3.85 m^2K/W R-value melamine foam insulation was obtained from McMaster-Carr. This was chosen for its balance of R-value, ease and safety of handling and flame retardant properties.

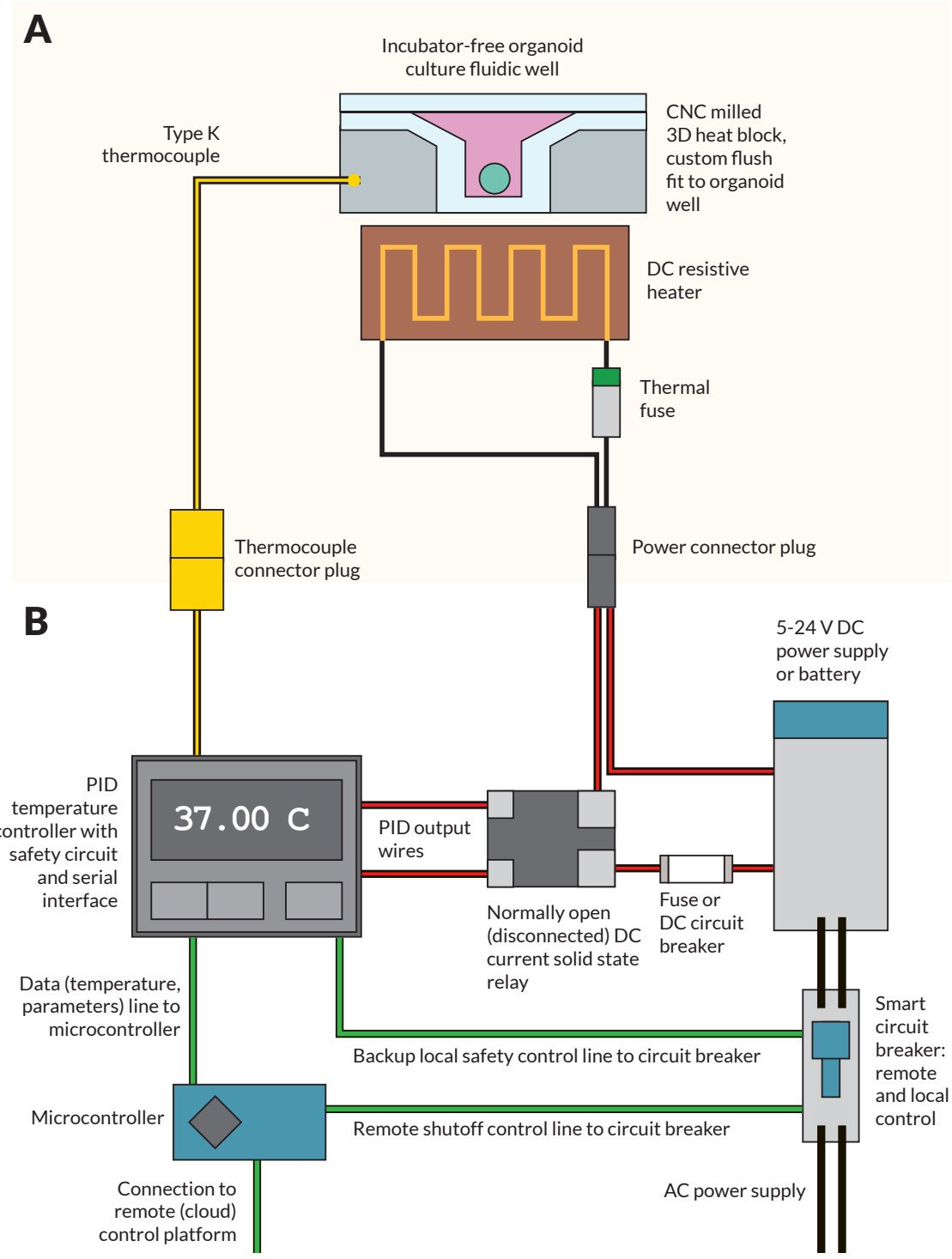

A
Incubator-free organoid culture fluidic well
Type K thermocouple
CNC milled 3D heat block, custom flush fit to organoid well
DC resistive heater
Thermal fuse
Thermocouple connector plug
Power connector plug
B
5-24 V DC power supply or battery
PID temperature controller with safety circuit and serial interface
37.00 C
PID output wires
Normally open (disconnected) DC current solid state relay
Fuse or DC circuit breaker
Data (temperature, parameters) line to microcontroller
Backup local safety control line to circuit breaker
Smart circuit breaker: remote and local control
Microcontroller
Remote shutoff control line to circuit breaker
Connection to remote (cloud) control platform
AC power supply

Figure 8.9: **Schematic of temperature regulation system** Above (shaded panel A): A heat block is milled from aluminium with precise dimensions to fit the cell culture device for which it is intended. A resistive heater is used to heat this heat block, a Type K thermocouple is inserted and fixtured for temperature feedback and the assembly is insulated (not pictured.) A thermal fuse is installed inline with the power to the resistive heater, to automatically and irreversibly cut power in case of overheating. Standardized electrical connectors are used to enable the experiment-specific equipment (shaded panel A) to be removed and exchanged. Below (panel B): This part of the schematic shows the reusable, experiment-agnostic control system electronics. Temperature feedback control is local using an industrial PID temperature controller with fault detection. The output from the controller is used to control a normally open DC solid state relay, supplied by a DC power supply protected by a current limiting fuse. A serial (eg USB) interface reports thermocouple measurements and allows parameter manipulation by an external computer. In a future implementation (not currently implemented) a smart (ie externally digitally controllable) circuit breaker is used as a method to turn on and off the system as well as to cut power in the rare event that solid state relay might fail in the closed position.

8.5.3 Manufacturing of heat block

The heat block consisted of a large plate sized to fit four cell culture devices to which four smaller heat blocks were attached using screws. The smaller feed blocks were designed to fit into the slight depression in the cell culture devices at the base of the cell culture compartment.

Aluminum blocks were laser-cut from 6061 aluminum by SendCutSend according to

.dwg files we provided. If burrs were present, parts were deburred by hand using an S-blade deburring tool. Screw-holes were countersunk manually using a high-speed steel vibration-resistant O-flute through-hole countersink on a Craftsman benchtop drill press, using a pecking motion. LPS Nº 40330 # 1 Tapmatic Gold cutting fluid was used for lubrication and coolant.

Thread tapping in the large plate for mounting of the sub blocks was performed manually in a single downstroke and single upstroke using a Tapmatic Nº 13033 RX 30 self-reversing tapping head and self-clearing black oxide cobalt steel through-hole threading machine taps. Anchorlube All-Purpose Metalworking Compound was used as tapping fluid. The sub blocks were then mounted using M4 steel Torx screws.

8.6 Cell culture and experimentation

8.6.1 Cell culture

8.6.1.1 Basic biological setup

NIH-3T3 murine embryonic fibroblasts which were obtained at passage number 151 from Sigma Aldrich. Media used was Dulbecco's Modified Eagle Medium (DMEM) with GlutaMAX from Gibco supplemented with 10% donor bovine serum (Gibco Nº 16030074, New Zealand Origin.) Gibco DMEM uses the standard 3.7 g/L sodium bicarbonate concentration.

Prior to beginning the experiments, the cells were expanded to generate homogeneous stocks to use for experiments. This process used at most 10 additional passages beyond the passage number 151 at which the cell line was obtained. In each experiment, all replicates of all conditions were seeded at equal seeding densities from cells at the same passage number grown

from the same plates.

All cell handling operations were performed using aseptic technique in a Class II biosafety cabinet. Except where otherwise noted, cell culture was performed in a shared humidified carbon dioxide incubator with temperature setpoint of 37 °C and carbon dioxide setpoint of 5% . This carbon dioxide setpoint was used due to the laboratory being a shared facility, even though it is not appropriate for standard DMEM.

8.6.1.2 Bioanalyte measurement

Point measurements of pH and of concentrations of oxygen, carbon dioxide, sodium bicarbonate, sodium ion and chloride ion were taken using a Beckman Coulter Vi-CELL MetaFLEX Bioanalyte Analyzer. Measurements were taken from 0.2 mL aliquots gathered in 1 mL syringes. Measurements were taken as soon as possible after aliquot collection; if more than 2–3 minutes elapsed between sample collection and measurement, syringes were capped tightly with polypropylene luer-lock caps to prevent gas exchange with room air. Any bubbles aspirated were expelled prior to sample loading, and care was taken to avoid bubbles of air contaminating any sample.

Bioanalyte analyzer measurements were not taken from closed-lid T25 flask samples except for initial and terminal measurements, since daily opening and closing of their lids for sampling would effectively cause gas replenishment of their internal chamber, confounding potential differences between conditions.

For all other conditions, 0.2 mL aliquots were collected with a 1 mL polypropylene luer syringe for bioanalyte measurements on a daily basis.

8.6.1.3 Initial cell seeding conditions

The in-house fabricated cell culture devices were loaded as follows. First, both luer caps of the cell culture compartment of the device were removed to allow for pressure equalization, then the vessel was laid on its side. Cells were deposited into the well via pipette filler. The well was then filled with media by the same port and the air bubble was allowed to escape via the other port. The cell compartment was then closed by re-insertion and locking of the Luer caps.

For air gap free devices, the seeding protocol was adapted from standard protocols for culture of NIH 3T3 cells using standard laboratory plasticware. The cell culture chamber was uncapped. Cell culture chambers were rinsed in Dulbecco's phosphate-buffered saline via the access ports. An aliquot of 350,000 cells, suspended in 4-4.5 mL of media was loaded into the cell culture component and the media was subsequently topped off to fill the cell culture component, if necessary. The access ports to the cell culture chamber were closed, and the plates were placed in a cell culture incubator. The device was placed with the cell culture surface downwards such that the cells could settle and adhere, and the plate was kept in this position for the remainder of the experiment. The plate was kept undisturbed in the incubator for 24 hours to the cell culture surface of the chamber after which time the media was exchanged and the experiment started.

On day four after plating, at which point the cell cultures in the cell culture devices had reached 80 to 90% confluency, experiments were terminated.

Cells were counted using a BIO-RAD TC 20 automated cell counter after dissociation

under standard protocols using 1 mL TrypLE Express dissociation reagent (Gibco). Briefly, after a 7–8 minute incubation with TrypLE Express, the cells were diluted to a total volume of 5 mL with media, collected, centrifuged for 3 minutes at 210 g and resuspended in 1 mL of media, after which point automated cell counting was performed.

8.6.1.4 Cell culture using standard cell culture labware

Control groups were cultured in either Thermo Scientific № 12-556-001 60 mm tissue culture dishes, VWR brand № 10861-570 T25 closed lid tissue culture flasks or VWR brand № 10861-572 T25 microporous membrane vented lid tissue culture flasks. Plating cell counts, media volumes, media change schedule and experimental endpoints were identical to air gap free vessels. Tissue culture dish and flask cultures were grown in the same cell culture incubator shared with other experimental conditions.

8.6.1.5 Incubator-free plate experimental procedure

Preparation of gas control fluid Aqueous gas control solution was fresh prepared from solid sodium bicarbonate and sodium carbonate (Sigma) in water. Stock mixtures were not used due to the possibility of decomposition and subsequent carbon dioxide loss from stock solutions which would cause the experiment parameters to deviate from those intended. Reagent mixtures were estimated using the nomogram published in Takano [126]. Our default mixture was 0.8 M sodium bicarbonate, 0.07 M sodium carbonate which we found resulted in a carbon dioxide concentration of approximately 7% and provided a cell culture media pH on the order of 7.4.

The first step in the gas control solution preparation was to ensure an oxygen concen-

tration in equilibrium with room air at 37 °C. First, Type I deionized water was heated to 37 °C in an open vessel under magnetic stirring for 15 minutes, at a rate sufficient to generate a vortex. This allowed the water to attain a partial pressure of oxygen approximately in equilibrium with room air at physiological temperature.

After 15 mintues, dry reagent was added, and the vessel was sealed to prevent carbon dioxide loss and stirred for an additional 5 minutes to ensure dissolution. A 300 μL aliquot was collected fresh to confirm pO_2 and pCO_2 measurements using the MetaFlex bioanalyte analyzer.

Loading of cell culture device A 10 mL syringe was filled with gas control fluid. The two luer caps of the gas control fluid compartment were uncrewed and the syringe was attached to the lower luer fitting. The fluid was injected into the compartment, expelling the air through the top port.

Incubator-less heating Prior to loading of the cell culture devices, the heating system was turned on and set to a set point of 40 °C which we found would give an average internal temperature on the order of 37 °C in the well. This was left to heat up and equilibrate during the cell culture device loading process. Once loaded, the cell culture devices were placed, cell culture chamber down, on the heat block previously described.

Media changes Media changes were performed analogously to the initial loading. Both Luer caps to the cell culture compartment were unscrewed to allow for free flow of media. The device was placed on its side and media was removed from the bottom Luer port. The media was then refilled via the same port with the air bubble being allowed to escape via the other port.

8.7 Bioanalyte and gas exchange results

8.7.1 Conventional air gap cell culture causes increases in media osmolarity via evaporation

Media sodium concentrations were measured using the Metaflex bioanalyte analyzer. We use hypernatremia as a quantitative proxy for extent of water loss via evaporation, since Weigmann et al. showed that sodium ion concentration in an evaporation-susceptible environment is an accurate proxy for measuring evaporation in cell cultures [134].

Over a four-day period, air gap cultures grown in 6 cm dishes showed significant day-over-day increases in sodium, indicating evaporation. On average, over the four days of cell culture, sodium concentration in these dishes increased by 13 mM over a baseline of 151 m, corresponding to an average volume loss of 9.4%. This is expected, both based on prior published results [31] and based on the humidity gradient and opportunity for free transport of water vapor that is inherent to air gap designs.

This effect was significantly blunted when using filter-cap flasks versus cell culture dishes, showing volume depletion limited to 1.8% over four days. This is likely due to a relatively lower effective cross-sectional area of the air-gap venting conduit in the filter cap versus the completely barrier-less air gaps of the conventional plate.

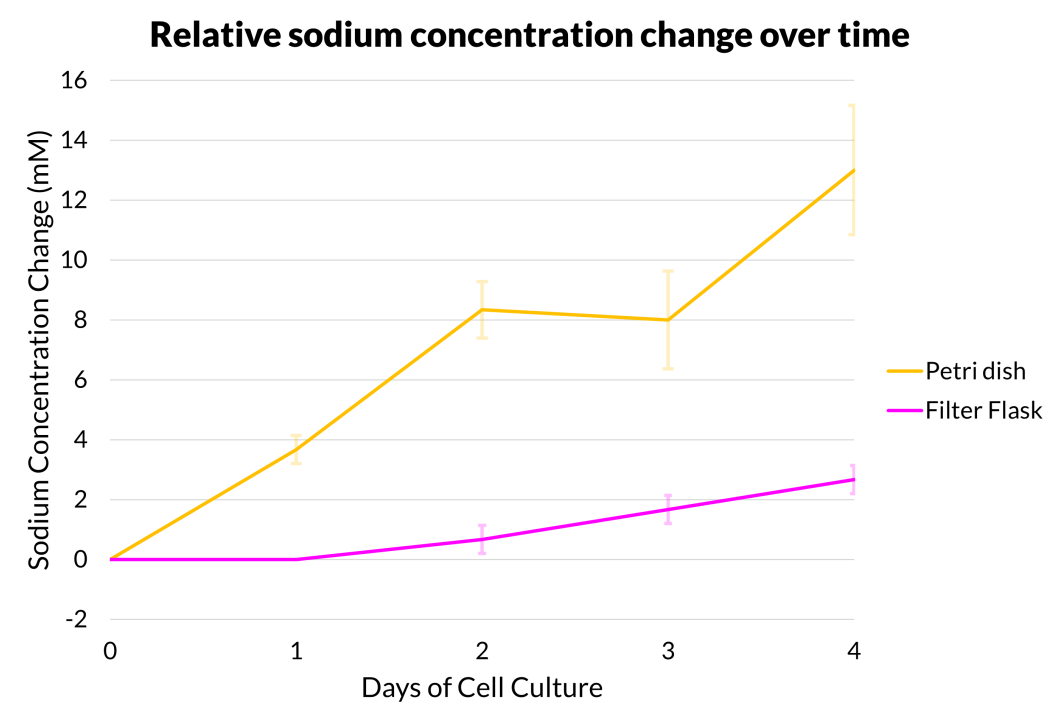


Figure 8.10: **Relative sodium concentration versus days of cell culture for Petri dish and filter flask culture** We observe increasing trends in sodium concentration for both Petri dish and filter flask culture versus a baseline level of 151 mM measured for fresh media on Day 0 of cell culture. Error bars represent standard deviations.

8.7.2 Non-air gap cell culture devices do not demonstrate definite trends in sodium concentration

In contrast to the cell culture dish results, we do not observe as clear a monotonic trend in sodium concentration (and by proxy osmolarity) in either the air gap free or the incubator free device. While the observed sodium concentrations generally exceed the baseline 151 mM of the fresh media, there is no definite monotonic trend and the standard deviation of the measurements

is comparatively high. While we cannot conclusively determine from these data that there is not any evaporative water loss, there is also no conclusive evidence that there is. To gather more conclusive data, we propose an experiment using continuous measurement of some proxy for osmolarity, either sodium concentration or conductivity of the solution.

The sodium concentration experiment also has a secondary role in assessing the incubator free device. The lack of definite monotonic increase in sodium concentration in these samples indicates that the polymethylpentene membranes and their sealing is indeed fluid-tight: even microporous holes or small failures in sealing would result in trans-film sodium leaks which would be evident since the transmembrane gradient between the cell culture media (151 mM sodium) and the gas exchange fluid (860 mM) is in excess of five-fold.

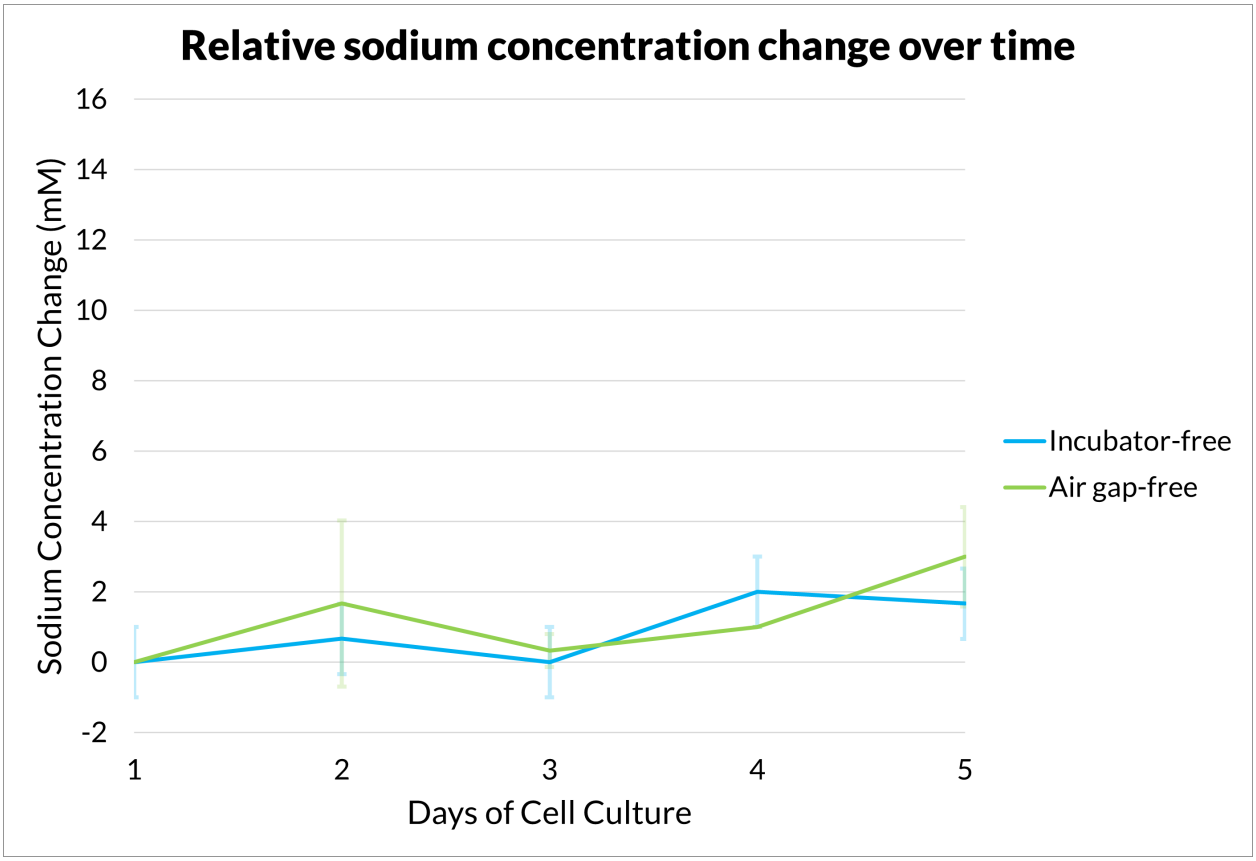


Figure 8.11: **Relative sodium concentration versus days of cell culture for incubator-free and air gap-free culture** We observe increasing trends in sodium concentration for both Petri dish and filter flask culture versus a baseline level of 151 mM measured for fresh media on Day 0 of cell culture. Error bars represent standard deviations.

8.7.3 Non air-gap, fluid-based respiration cell culture devices maintain stable carbon dioxide concentrations and pH values within target ranges

Physiological human arterial pH is tightly regulated in the range of 7.36-7.44 [44]. By tradition, cell culture laboratories often set their incubator setpoints to 5% carbon dioxide: however, if 5% carbon dioxide can truly be maintained, then absent acid products of metabolism, expected pH is 7.67 based on the Henderson–Hasselbalch equation, assuming a pKa of 6.1 for carbonic acid. In long-term cell culture, accumulation of acid products of metabolism, primarily lactic acid and ammonium, decreases this pH marginally, however the 5% carbon dioxide setpoint will still produce an alkalotic media pH.

In our results, we observe that, apart from on Day 0, where the pH measured is of the fresh media with a 7% dissolved carbon dioxide content, neither the air gap cell culture vessels (Petri dish or filter flask) nor the air gap free vessel (which exchanges carbon dioxide with the incubator mixed gas via its polymethylpentene gas permeable film) achieve any pH in the physiological range. We observe as well that at no point does the dissolved carbon dioxide level in any of these vessels reach the incubator setpoint of 5%: among these vessels, the average dissolved carbon dioxide partial pressure is 4.40% and the maximum observed is 4.83%. This suggests that the underprovision of carbon dioxide at the level of the incubator

mixed gas setpoint might be worsened by incubator openings and/or inhomogeneity, as has been observed with humidity levels. While the elevated pH seen in all three of these conditions is primarily indicative of the fact that a 5% CO2 setpoint is not appropriate for a standard DMEM buffer system, it suggests two other advantages of an incubator-free cell culture system.

First, the difference between the measured dissolved carbon dioxide concentrations and the intended incubator setpoint makes it clear that incubator openings are a major barrier to achieving precise carbon dioxide (and therefore pH) control. Second, using shared cell culture incubators requires that all users agree jointly on a single gas blend which they all must share: often a cultural default such as 5% carbon dioxide is chosen rather than individual researchers being able to choose the concentration which is actually ideal for their experiment.

The pH and carbon dioxide data is also evidence that the air gap free cell culture vessel, which consists of a fluidically sealed cell culture chamber equipped with a nonporous polymethylpentene gas permeable film, freely exchanges dissolved gas with the incubator mixed gas environment. This is shown by the concordance between the dissolved gas levels of these air gap free vessels and the petri dishes and filter flasks.

In contrast, incubator free vessels employing a 800 mM sodium bicarbonate, 70 mM sodium carbonate buffer for gas generation and regulation showed a pH within the physiological range of 7.36 to 7.44 for the duration of the experiment.

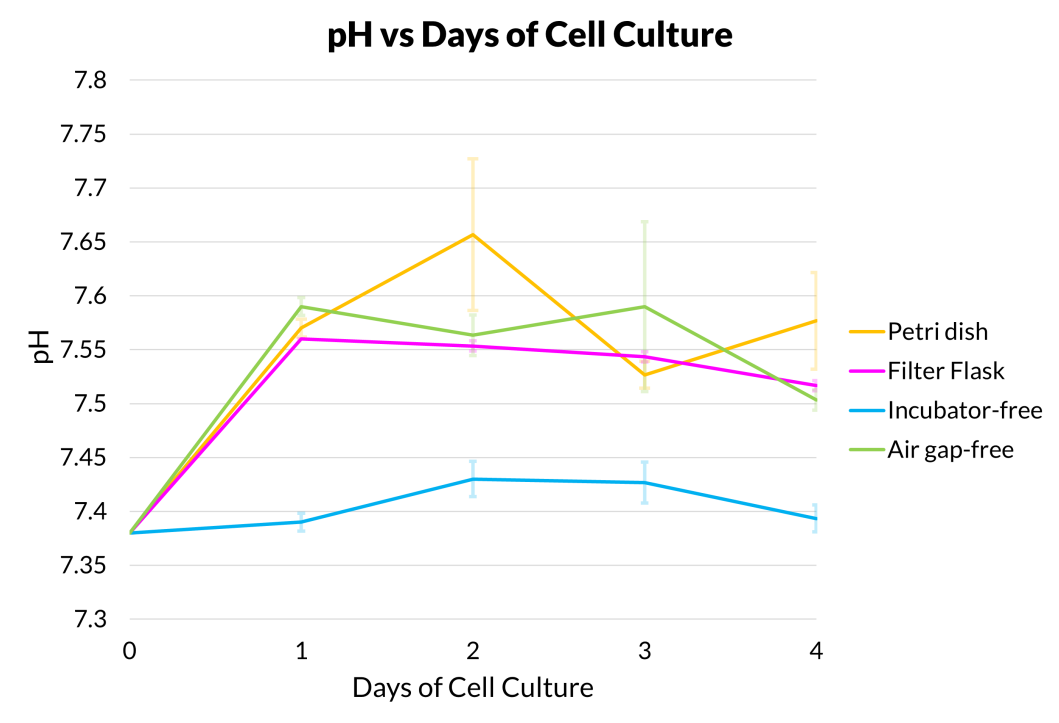


Figure 8.12: **pH in air gap and air gap-free cell culture vessels over four days of cell culture** The media buffer was Gibco DMEM with 3.7 g/L bicarbonate buffer. The incubator setpoint was 5% carbon dioxide. The incubator-free used an 800 mM sodium bicarbonate, 70 mM sodium carbonate buffer for gas generation and regulation. Error bars represent standard deviations.

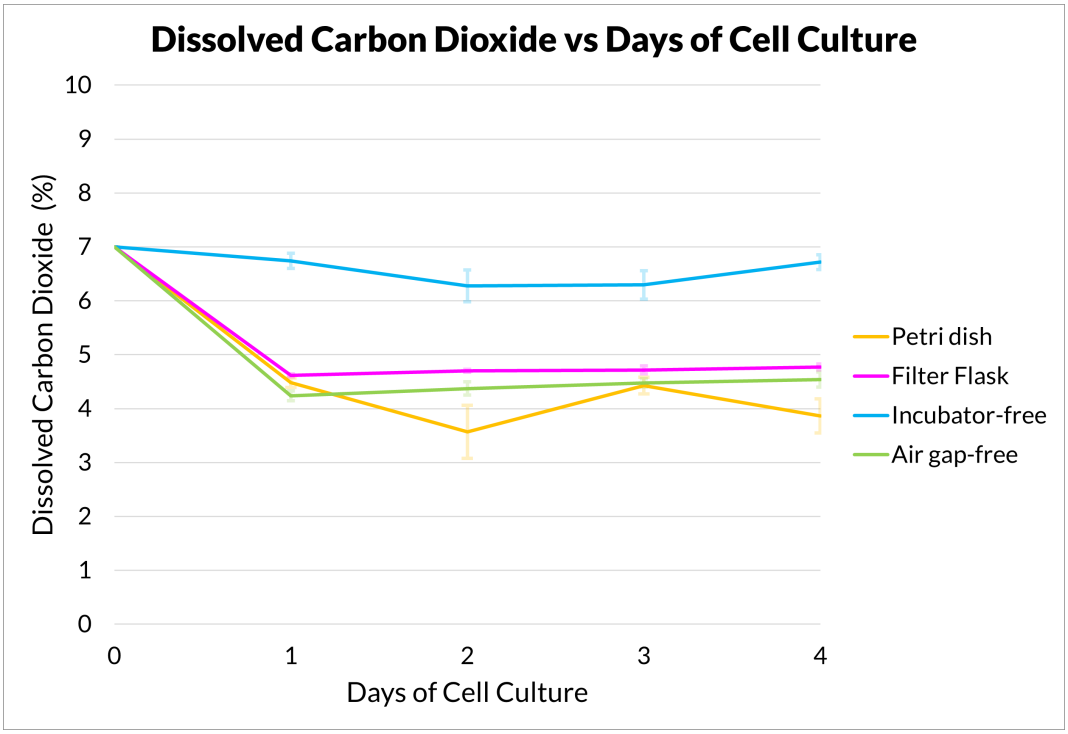


Figure 8.13: **Dissolved carbon dioxide in air gap and air gap-free cell culture vessels over four days of cell culture** The media buffer was Gibco DMEM with 3.7 g/L bicarbonate buffer. The incubator setpoint was 5% carbon dioxide. The incubator-free used an 800 mM sodium bicarbonate, 70 mM sodium carbonate buffer for gas generation and regulation. Gibco DMEM nominally requires 5-10% carbon dioxide setpoint to achieve physiological pH, note that no device exchanging gas with the incubator environment was able to maintain a partial pressure of CO2 at the setpoint. Error bars represent standard deviations.

8.7.4 Non air-gap, fluid-based respiration cell culture devices maintain stable oxygen concentrations within target ranges

Measurement of dissolved oxygen content of the cell culture devices did not show clear trends or differences between experimental conditions. Importantly, neither air gap free (membrane respiration of incubator mixed gas) nor incubator free (membrane respiration of gas control fluid) devices showed differences in oxygen trends as compared to conventional air gap devices. We hypothesize that near-atmospheric oxygen concentrations have too much of an excess of oxygen relative to metabolic needs for there to be any measurable consumption in any cell culture system tested. Experiments evaluating control of oxygen at physoxia and hypoxia are outside the scope of this work but would be necessary to measure the ability of the incubator free system to compensate for oxygen consumption in cell culture.

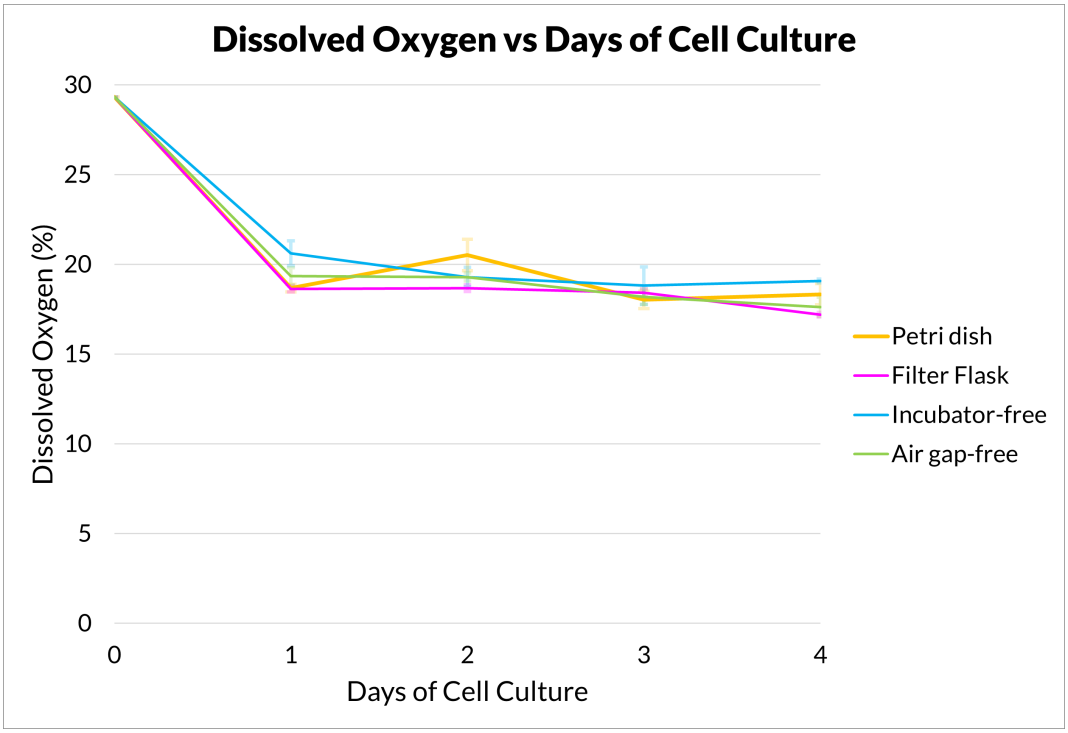
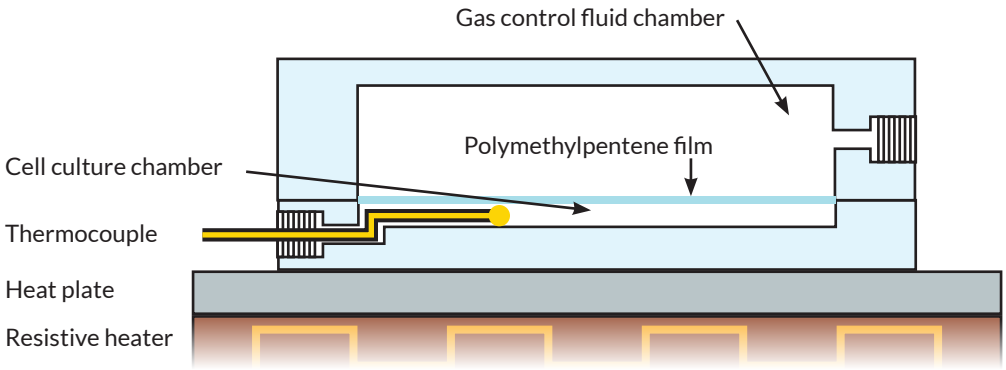


Figure 8.14: **Dissolved oxygen in air gap and air gap-free cell culture vessels over four days of cell culture** Dissolved oxygen trends are approximately equal between all four vessel types: standard air gap petri dish and filter flask vessels, as well as air-gap free and incubator-free vessels. Note that at atmospheric pressure and 37 °C, water in equilibrium with room air has a partial pressure of 20.9% oxygen.

8.8 Thermal characterization results

8.8.1 Outside of incubator heat block temperature characterization

We tested the ability of our heating system, combined with a heat block specifically designed for our 2D culture plates, to maintain a target temperature for cell culture within our cell culture devices. This was achieved by inserting a thermocouple probe into the cell culture compartment of the incubator-free device (see figures below.).



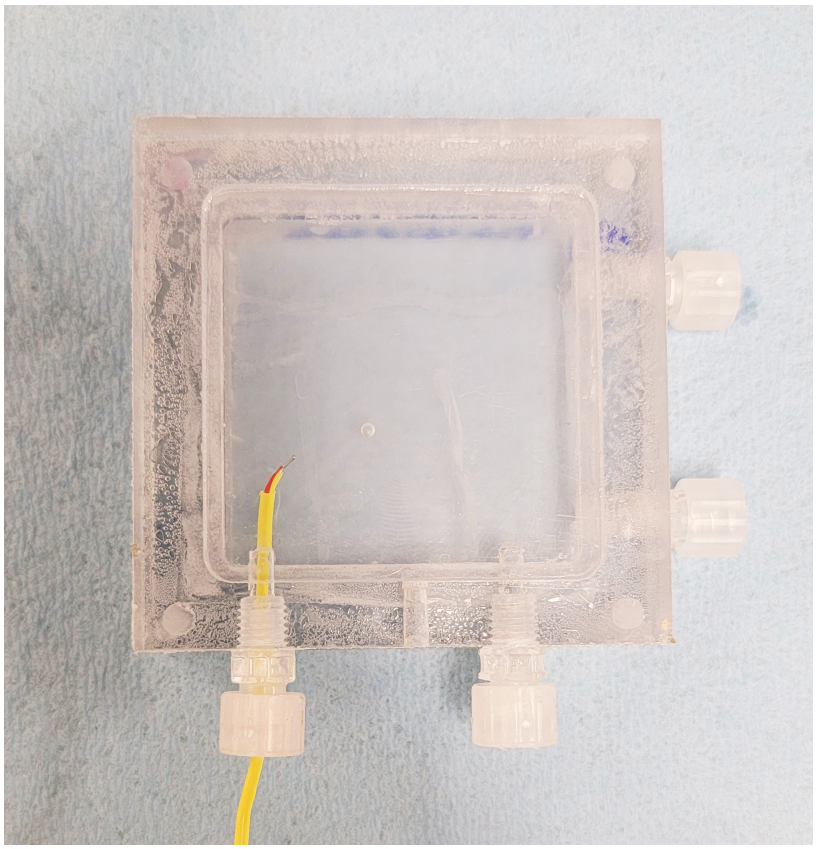


Figure 8.15: **Thermocouple positioning for minimal 2D incubatorless cell culture experiment** Top: Cartoon of thermocouple positioning. Bottom: Photograph of thermocouple positioning. The thermocouple is inserted via one of the two filling ports, then sealed in place using a gasket improvised from wrapped parafilm and a compression fitting improvised from a luer plug fitting with a hole cut in its plug.

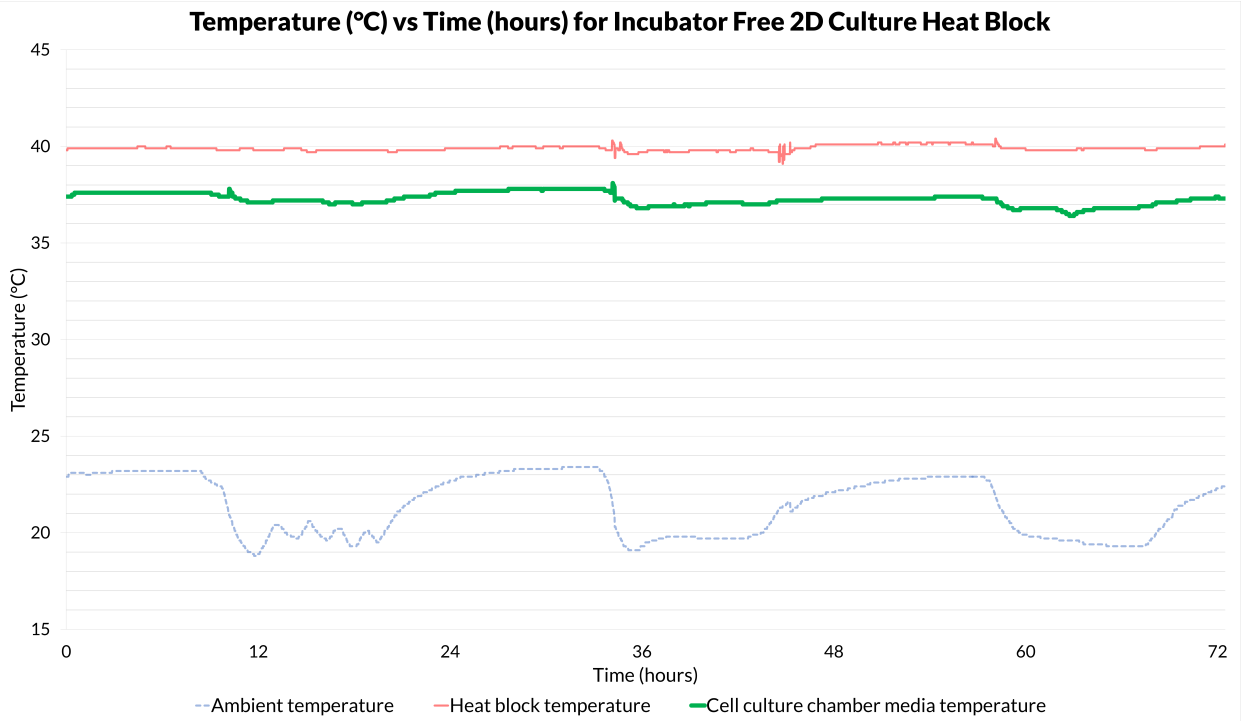


Figure 8.16: **Temperature control using heat block for 2D incubator free cell culture** Measurements were taken over 3 days. Temperature traces correspond to a thermocouple attached to the lower surface of the heat block, the thermocouple inside of the cell culture chamber detailed in the previous figure, and a thermocouple measuring ambient temperature which was suspended above the lab bench and did not contact any solid surface. Note the muted drop in temperature of interior of the cell culture chamber relative to drops in room temperature.

Temperatures of the media in the device varied between 36.4 °C and 38.1 °C, with a mean and median of 37.3 °C and 50% of temperatures lying between 37.1 °C and 37.6 °C and 99% of temperatures lying between 36.5 °C and 37.8 °C. Meanwhile, ambient temperature around the device varied from 18.8 °C and 23.4 °C, with a mean of 21.4 °C and median of 21.7 °C and 50% of temperatures lying between 19.8 °C and 22.9 °C

Chapter 9

Devices for 3D organoid culture

9.1 Design and manufacturing of heating system

9.1.1 Design of heating control

The heating control system was designed and assembled as previously described. Owing to the controller and power supply system being agnostic of the resistive heater layout and heat block design, the previously described system was adapted to these devices without any modification.

9.1.2 Design and manufacturing of heat blocks

9.1.2.1 Design

Since these 3D culture devices are intended for continuous long term culture in a single cell culture device, rather than short-term culture requiring media changes or passaging every few days, we designed heat blocks to envelop the bottom part of the cell culture devices

from five sides rather than only heating from below. This was achieved using a well structure which holds each cell culture device. The heat block can be designed in a modular manner to hold one or more cell culture devices in repeating subunits.

Each cell culture well device subunit of the heat block was designed to incorporate four full thickness M4 threaded holes. These can be used with M4 screws in order to mount and compress a lid against the cell culture device which is seated in the heat block. Using full thickness M4 threaded screw holes in a thick block of aluminum allows considerable compression force to be exerted by the screws without danger of stripping the threads.

Due to a desire for compatibility with an earlier well design wherein there were fluidic ports attached to two sides of the bottom of the cell culture device, we prototyped the device as a four-well heat block. However, the modules are designed such that a six well version of the heat block positions the cell culture wells to be identically spaced to a standard cell culture plate at a pitch of 39 mm. A single well, open bottom version of the heat plate was also designed to fit the Maxwell Technologies MaxOne cell culture well.

9.1.2.2 CNC milling of heat blocks

$\frac{5}{16}$" 6061 aluminum stock was obtained from McMaster Carr. Heat blocks were CNC milled from this stock on the Carbide3D Nomad3 desktop milling machine using a Harvey Tool Nº 73199-C4 amorphous diamond coated 4-flute $\frac{1}{4}$" square endmill for coarse operations and Harvey Tool Nº 73125-C4 amorphous diamond coated 4-flute $\frac{1}{8}$" square endmill for fine operations. The Carbide3D Nomad threaded table and clamps were used for fixturing. On account of the weak motor and low stiffness of the Carbide3D desktop mill, shallow passes

of 0.05 mm were used. Despite this, we experienced inconsistent success in accurately or completely cutting and recommend machining on a higher powered mill or via a service bureau.

Screw holes were pilot-drilled to a depth of 2 mm on the CNC mill, then drilled to full depth with a 3.5 mm carbide drill bit using copious coolant. The holes were then tapped to M4 thread using the TapMatic tapping head as previously described. Parts were post-processed manually using an S-blade deburring tool, a belt sander and a Dremel rotary multitool equipped with a sanding drum.

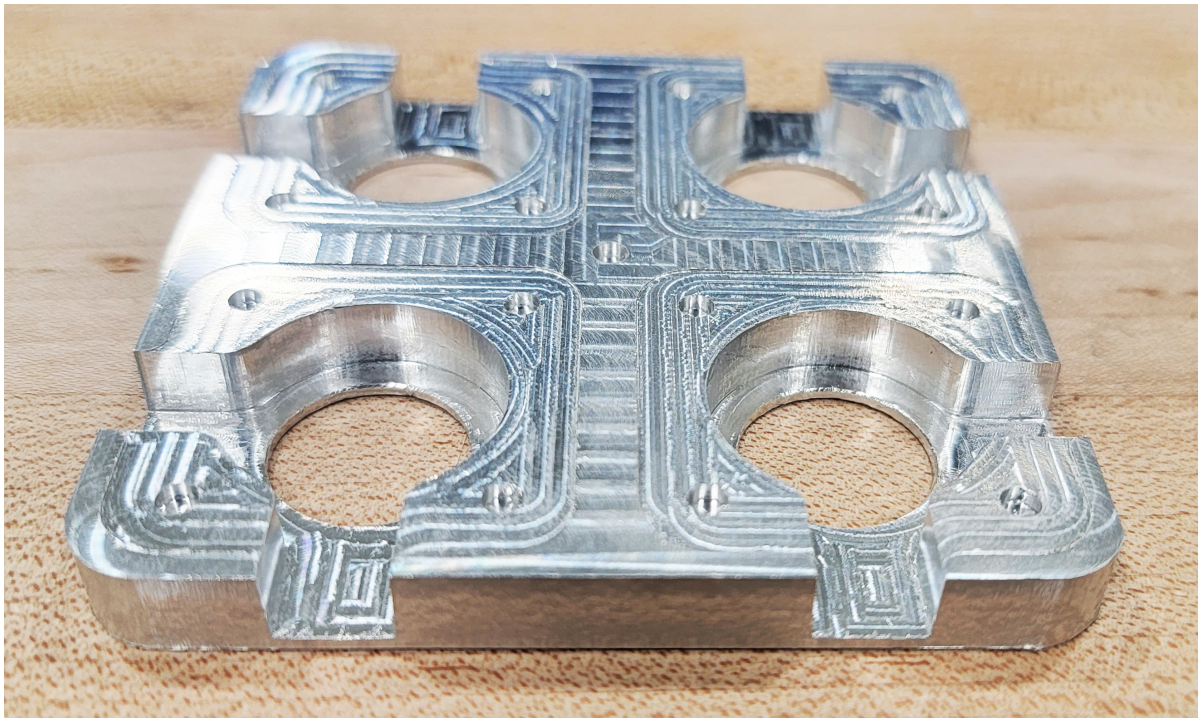


Figure 9.1: **Detail of a prototype CNC milled heat block designed to hold multiple fluidic devices for organoid culture**

9.2 Design and manufacturing of cell culture devices

9.2.1 Cell culture well design and manufacturing

9.2.1.1 Cell culture well design

Our goal was to design an incubator-free device which would support the tissue culture of small cortical organoids while confining them to a small footprint for microscopy. For initial prototypes we chose to develop a device which would accommodate four organoids, up to 3 mm each in diameter within a 6.5 mm square footprint.

The fundamental requirements for the design were: it would need to hold organoids, it should contain a small dead volume of media for dilution of waste products, it should accommodate imaging and it should possess inflow and outflow ports for manual or automated media exchange and a removable lid with a port large enough to insert or remove organoids. In this specific iteration of our design we also engineered the dimensions of the device to perfectly fit the horizontal and vertical footprint of the MaxWell Biosystems MaxOne CMOS electrode array well. By doing this, we will easily be able to generalize our cell culture device to one which allows simultaneous microscopy and electrophysiology experiments outside of an incubator environment. A schematic of the axial real estate considerations used in the well design is shown below.

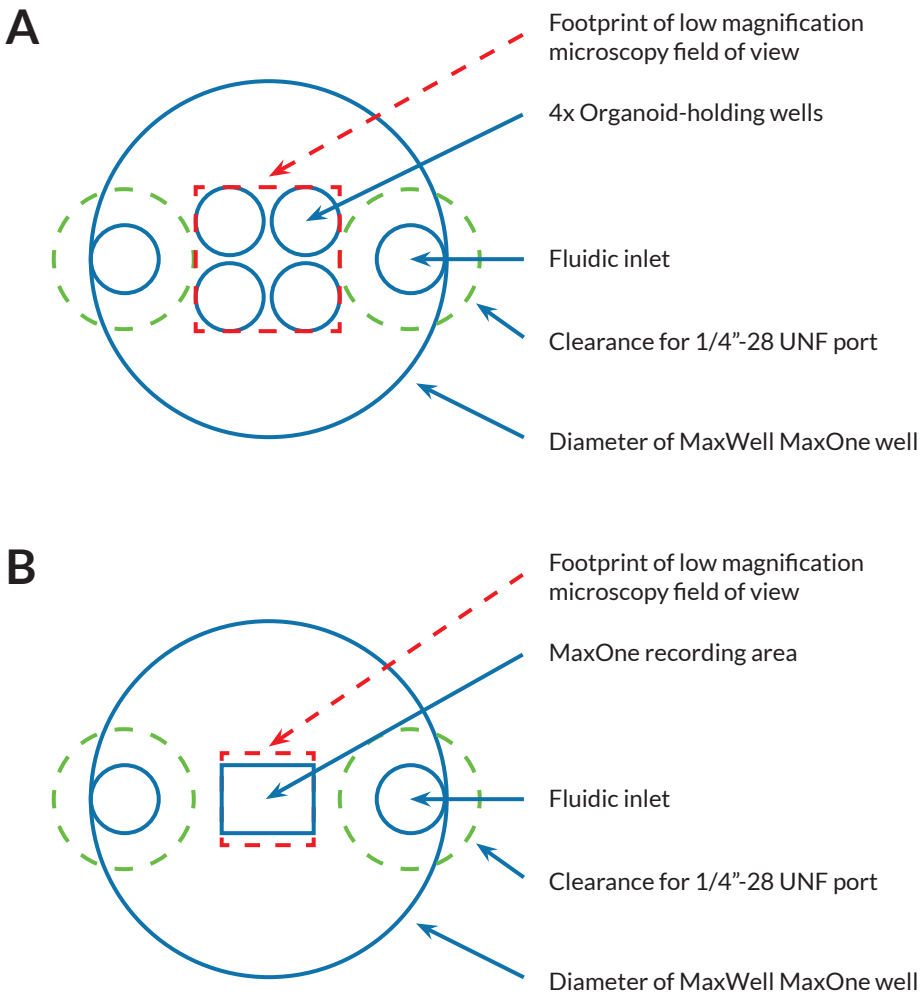


Figure 9.2: **Axial plane real estate considerations for design of organoid culture well** Panel A shows the four-organoid version of the device which places for 3 mm wells for organoids in the tightest packing possible given the limitations of our CNC milling capabilities. Clearance from the four fluidic ports is required both on top of one of the order to enable imaging with transformation. Panel B shows an adaptation of the same axially layout for use in simultaneous electrophysiology (from bottom of device) and microscopy (from top of device.)

We designed the device in two main subcomponents. A cell culture chamber subcomponent contains a large well which holds media and connects to fluidic ports. The bottom of the cell culture chamber contains smaller wells for holding individual organoids. The bottom of the smaller wells is optically clear. A lid subcomponent seals the well from above and also provides fluidic ports for media changes, provision of gas-equilibrated media or other fluid manipulation operations. The two subcomponents are sealed together using a watertight and gastight gasketing scheme. Screws apply pressure between the lid subcomponent and the heat block previously described. This sandwiches the sealing gaskets and the cell culture chamber subcomponent in the middle of the assembly. The pressure applied compresses the gaskets to form a seal and also presses the cell culture chamber to the heat block to maximize heat transmission.

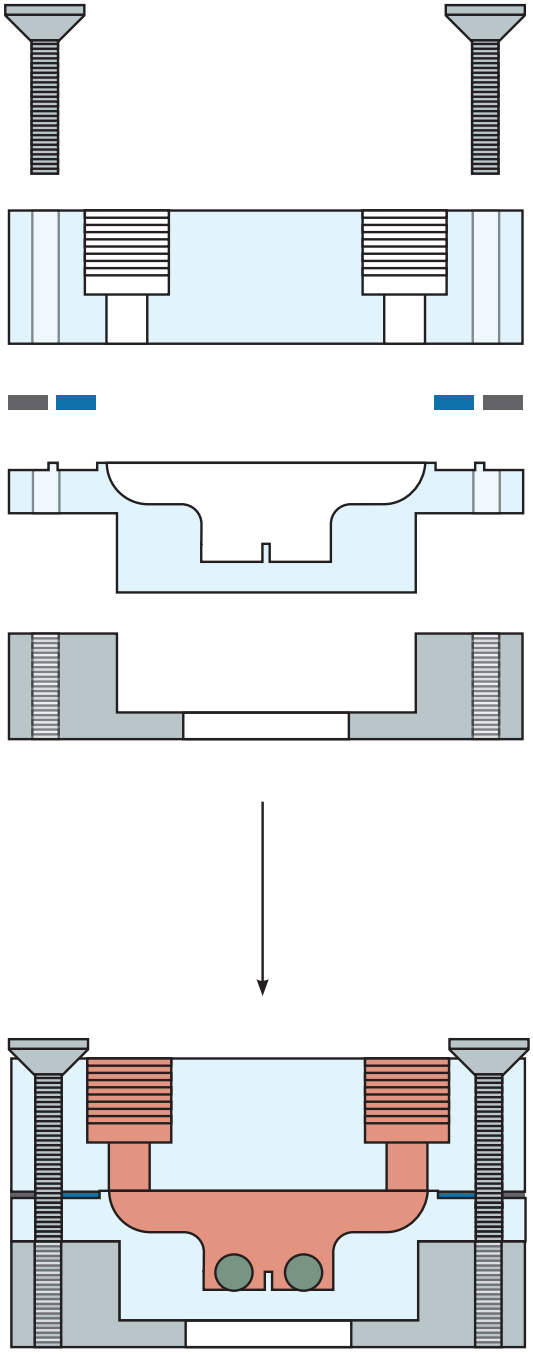

M4 screws (for sealing pressure)
Lid subcomponent
Sealing gaskets
Cell culture chamber subcomponent
Heat block with threaded channels for screw mounting
Assembled incubator-free organoid culture device

Figure 9.3: **Organoid cell culture device** Top panel shows the subcomponents which comprise the organoid cell culture device assembly. Bottom panel shows the assembled device with all components sandwiched together. Organoids are shown in the wells of the cell culture chamber, with the optical microscopy window and a window in the heat block below.

9.2.1.2 Cell culture sealing surfaces

In order to allow for access to insert and remove organoids, it was necessary to be able to access the entire well. This required that a lid and bottom to subcomponent design was needed for the cell culture device. In order to be watertight under pressure while using manufacturing techniques with $>$ 1/1000 inch precision and while using rigid on rigid thermoplastic connections, gaskets were necessary to form a seal between the two subcomponents of the device. For initial prototypes we have opted for flat rather than O-ring gaskets. Outside of the current rapid prototyping phase, it may be beneficial to explore the use of O-ring gaskets. For reasons that will be explained shortly, our gasketing scheme uses two sequential gaskets. To establish a seal with these gaskets, grooves of precise depth were milled in the bottom subcomponent to see to the gaskets and to allow them to protrude a set distance above the lip of the well. The opposite meeting part on the lid of the well was capped with a smooth surface for gasket sealing.

9.2.1.3 Inlet and outlet positioning

Early prototypes used media inlets and outlets positioned below the highest point in the cell culture chamber. (See figure below, left.) We discovered that having the outlet in this

position caused accumulation of any bubbles in the fluidic perfusate to accumulate in the cell culture well. In one case, this was severe enough to contact the organoids on the bottom of the well and cause their drying. This will also affect the effective media volume in the well, which could cause more rapid accumulation of waste products. In retrospect, it is clear that any design with a low-positioned outlet will suffer this bubble trapping. Therefore, we have designed our wells to have their outlet at the highest point in the well, as a component of the well lid rather than well bottom subcomponent. (See figure below, right.)

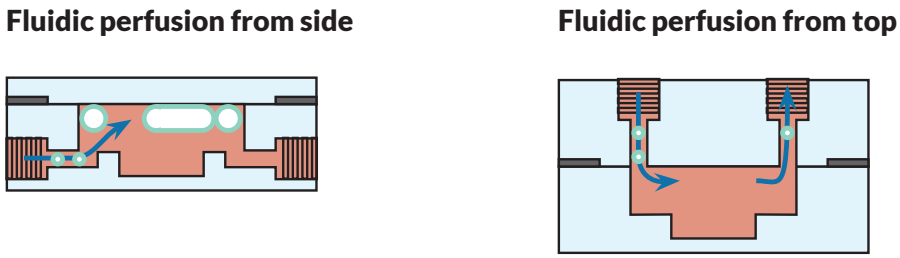


Figure 9.4: **Perfusion channel geometry and bubble trapping** This diagram shows, on the left, the bubble trapping which occurs when the outlet of a fluidic chamber is below the level of the roof of the chamber. The right shows a device where an outlet at the level of the roof allows bubbles to exit.

9.2.1.4 Cell culture device milling

$\frac{1}{2}$" thick polycarbonate sheet stock (Plaskolite Tuffak or Covestro Makrolon brand) was obtained from McMaster Carr. All milling steps were done on a Carbide3D Nomad 3 desktop CNC milling machine. Toolpaths were generated in Carbide Create version 6. XY-axis zeroing was performed using a LCA mini LC/EF spindle mounted laser centering-edge finding

tool and Z-axis zeroing was performed manually. Fixturing for basic cutting operations was achieved using Shurtape DS 154 double sided tape on an medium-density fiberboard (MDF) wasteboard which was faced using a fly cutter prior to use. For precise positioning and squaring to perform reverse milling operations, a right angle corner alignment jig was machined from MDF.

Milling cutters used were the Harvey Tool Nº 52808 $\frac{1}{8}$" diameter, $\frac{5}{8}$" length of cut square upcut single flute endmill and Harvey Tool Nº 929762, $\frac{1}{16}$" diameter, 0.186" length of cut square upcut single flute endmill for upcut operations and the Harvey Tool Nº 935408 $\frac{1}{8}$" diameter, $\frac{5}{8}$" length of cut square downcut single flute endmill for downcut operations.

The first subcomponent prepared was the bottom half of the cell culture well. In order to manufacture a cell culture well with an optically clear bottom, we followed a procedure analogous to the manufacturing of the incubator free plates for two dimensional cell culture previously described. Specifically a reverse milling operation using an upcut endmill created a pocket for insertion and bonding of a polycarbonate window as well as the outside contour of the device. Next, an obverse milling operation using an upcut endmill was used to shape the outside contour and the interior pocket of the cell culture chamber.

Specifically, in order to mill the reverse features of the bottom component, the $\frac{1}{16}$" tool was used to mill a shallow pocket to fit a $\frac{1}{16}$" polycarbonate window, designed to cover four $\frac{1}{8}$" diameter organoid-holding wells, each separated by a gap measuring 0.5 mm at its minimum. Next, the $\frac{1}{8}$" tool was used to drill said wells by peck drilling and subsequently to mill a contour and square flange for mating the cell culture wells to the pockets in the heat block previously described. Screw holes were drilled in the corners of the flange such that screws could pass

through for compression of the lid component against the heat block, sandwiching this cell culture well component and the gaskets in between. A border wall was left around the flange to provide additional surface area for fixturing as well as an XY registration point for alignment during the obverse milling operation.

The parts, which had been milled on their obverse sides, were aligned and fixtured on their opposite sides. The square $\frac{1}{8}$" downcut endmill was used for all subsequent operations to reduce lifting forces. A facing operation was used to obtain consistent measurements between parts and to remove the surface which had been in contact with the acrylic tape. Grooves were milled for the gaskets. A tapered cell cultured chamber was milled, which has a large cross section at its top to interface with off-center fluidic ports while maintaining a small diameter at its bottom to center the organoids within the viewing window. Finally, the parts were separated from the sacrificial border wall left over from the reverse operation.

Next, the optical window, designed to be 25 μms smaller than the window-fitting pocket was routed from $\frac{1}{16}$" polycarbonate. This was cleaned with Texwipe Nº TX3273 sterile 70% USP-grade isopropanol and Kimtech low-linting non-abrasive wipes. It was pressed into its mating pocket using a metal probe, then solvent welded into place with dichloromethane via capillary action, supplied via a disposable 27 gauge needle and a Hamilton Nº 1001 TLL glass/polytetrafluoroethylene gastight syringe.

The parts were optionally post-treated on the inner surface of the cell culture well by sparing application of liquid dichloromethane. This operation smoothed the matte surface of milled regions such that they would be optically clear for observation.

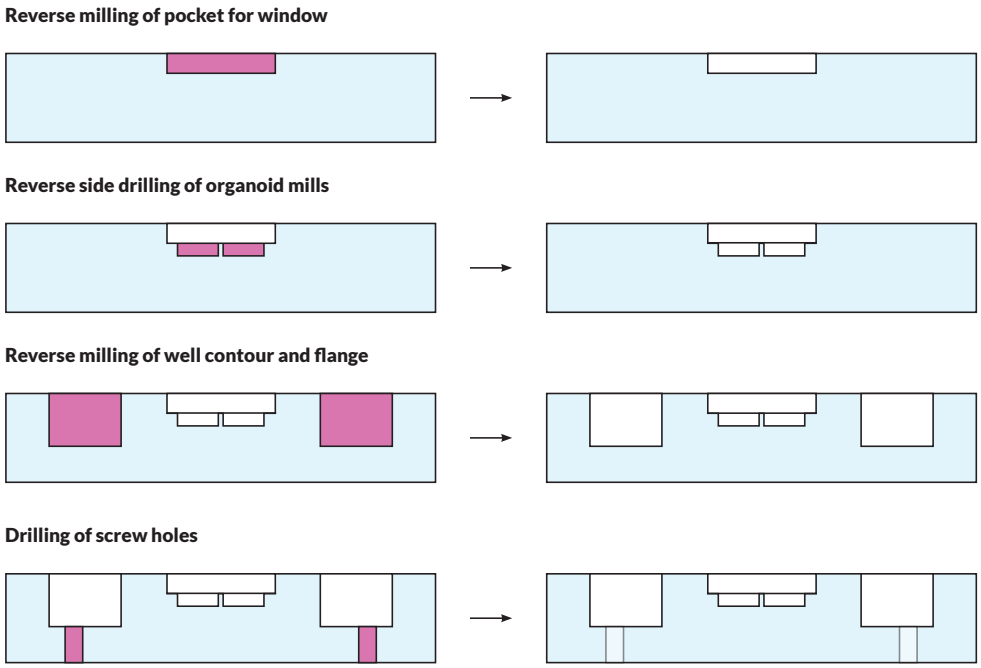


Figure 9.5: **Reverse milling operations for fabrication of organoid culture device well subcomponent** The reverse side of the stock is machined first. A pocket for later insertion of an optically clear window is milled and the organoid wells are also drilled from the reverse side. The external culture of the well and the gasket sealing flange are subsequently milled along with through holes through which the sealing screws pass.

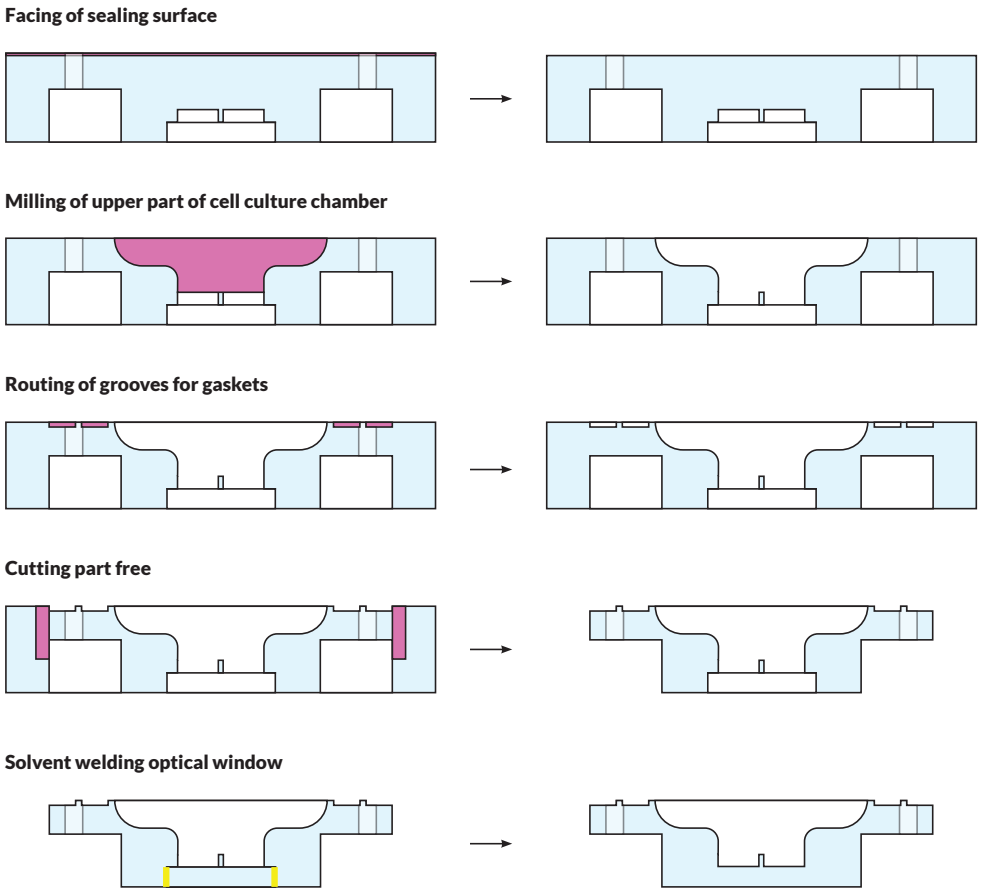


Figure 9.6: **Obverse milling operations for fabrication of organoid culture device well subcomponent** The obverse side is faced to leave a clean and dimensionally accurate top surface for sealing. The cell culture chamber is milled to join the organoid wells milled from the reverse side. Grooves are routed for gaskets. The part is then cut free of the additional material which had been left to support the part during the obverse milling operation. Finally, an optical window which had been routed separately from stock is solvent welded into the previously left pocket, sealing the bottom of the well with an optically clear window for microscopy.

After milling and assembly of the bottom subcomponent of the cell culture device,

the lid component was milled. This was again performed using $\frac{1}{2}$" thick polycarbonate sheet stock milled using the Carbide3D Nomad 3 desktop CNC mill with set up analogous to previous milling jobs. The lid incorporated two $\frac{1}{4}$"-28 fluidic ports positioned outside the optical path of the observation window in the bottom of the cell culture well subcomponent. It also incorporated four screw holes for compression mounting using the threaded holes of the heat block.

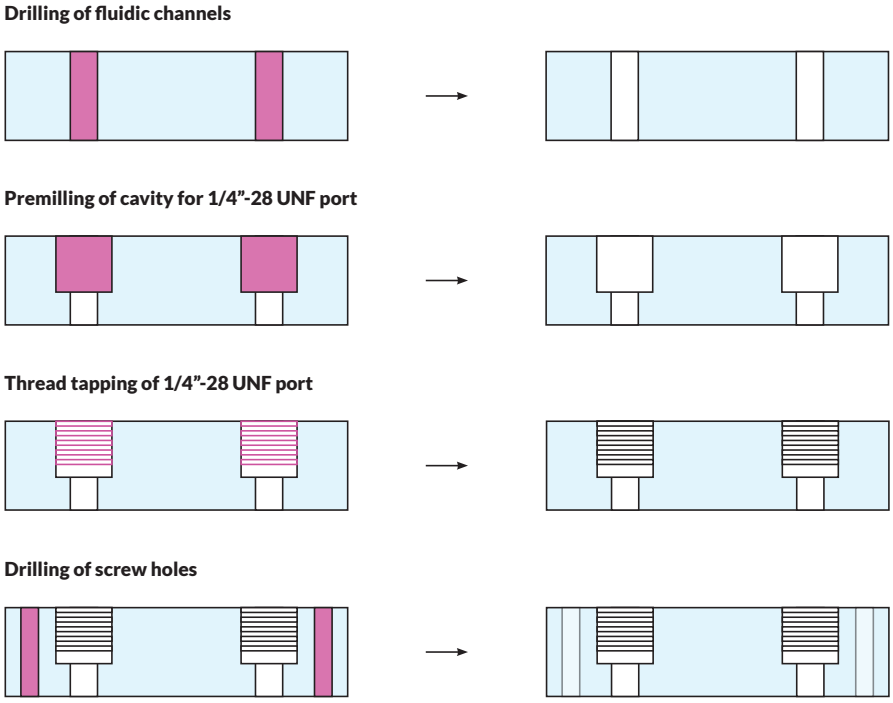


Figure 9.7: **Milling operations for fabrication of organoid culture device lid** After the outer shape of the lid is routed from polycarbonate stock (not pictured), fluidic channels with ¼"-28 UNF threaded ports are fabricated via combination of milling and thread tapping as previously described. Screw holes are CNC drilled in the part to match the holes in the cell culture well subcomponent and heat block.

9.2.2 Gasket design and manufacturing

9.2.2.1 Gasket design

Early prototype cell culture devices used solitary fluorosilicone gaskets. However, we found that despite using a wide gasket of thin material, $\frac{1}{64}$"" in thickness, in an attempt to minimize gas transport, there was still sufficient carbon dioxide leakage through the gasket to rapidly deplete the gas exchange fluid, as observed via continuous pH monitoring. Despite initial successful pH control over the first few hours of use of the device, we observed pH to drift into the >8 range, indicating depletion of dissolved carbon dioxide via a leak in the device. Later elimination of this drift in solvent welded experimental devices incorporating no gasket whatsoever confirms that this leak was indeed via the gasket material.

Unfortunately, fluorosilicone was also the only elastomer we found which was robust to one or more of our preferred sterilization methods, ozone, dry heat and autoclave and which we also found to not induce significant decreases in viability of cells.

Literature review revealed that even though silicones (including but not limited to fluorosilicones and polydimethylsiloxane) are extremely gas permeable, elastomers are, as a general class, all relatively gas permeable with the the exception of isobutylene isoprene rubber. However, it is well known that this rubber releases significant leachates by direct contact with water which are also known to be toxic to multiple cell lines.

With these benefits and restrictions of each material in mind, we came up on a solution: double gasketing. An inner gasket provides the non-cytotoxic liquid seal and an outer gasket provides the gas seal.

Figure 9.8: **Double gasket for organoid cell culture device** An inner fluorosilicone gasket is surrounded by an outer isobutylene isoprene rubber gasket.

9.2.2.2 Gasket manufacturing

We also experimented with multiple techniques for producing gaskets: CNC drag knife cutting, CNC rotary knife cutting, manual die punching and CNC routing.

For CNC drag knife cutting and CNC rotary knife cutting, we used a Cricut Explore Air 2 and a Cricut Maker respectively. We found that the cuts were inconsistent and different from what was specified because of sticking of fluorosilicone to the knife, stretching and peeling or moving it. Furthermore, the roller mechanism of the Y axis of these devices left additional defects in the machine surface. Manual die punching was also attempted, using commercial circular die punches, however we were not able to adequately coaxially center the cuts.

Finally, we developed a method of CNC routing fluorosilicone and isobutyl-isoprene

rubber. This involved a combination of optimizing fixturing to prevent detachment of the work-piece and optimizing cutter geometry, cut depth and feed rate to minimize lifting and tearing forces. For routing fluorosilicone, heterogeneous silicone-acrylic double sided adhesive tape (Marian Chicago Inc) was used for fixturing. For routing isobutyl-isoprene rubber, Shurtape DS 154 acrylic double sided tape was used. A Harvey tool Nº 52462 $\frac{1}{16}$" square upcut endmill was used with a shallow 50 μm depth of cut in order to reduce chances of lifting or tearing of the material.

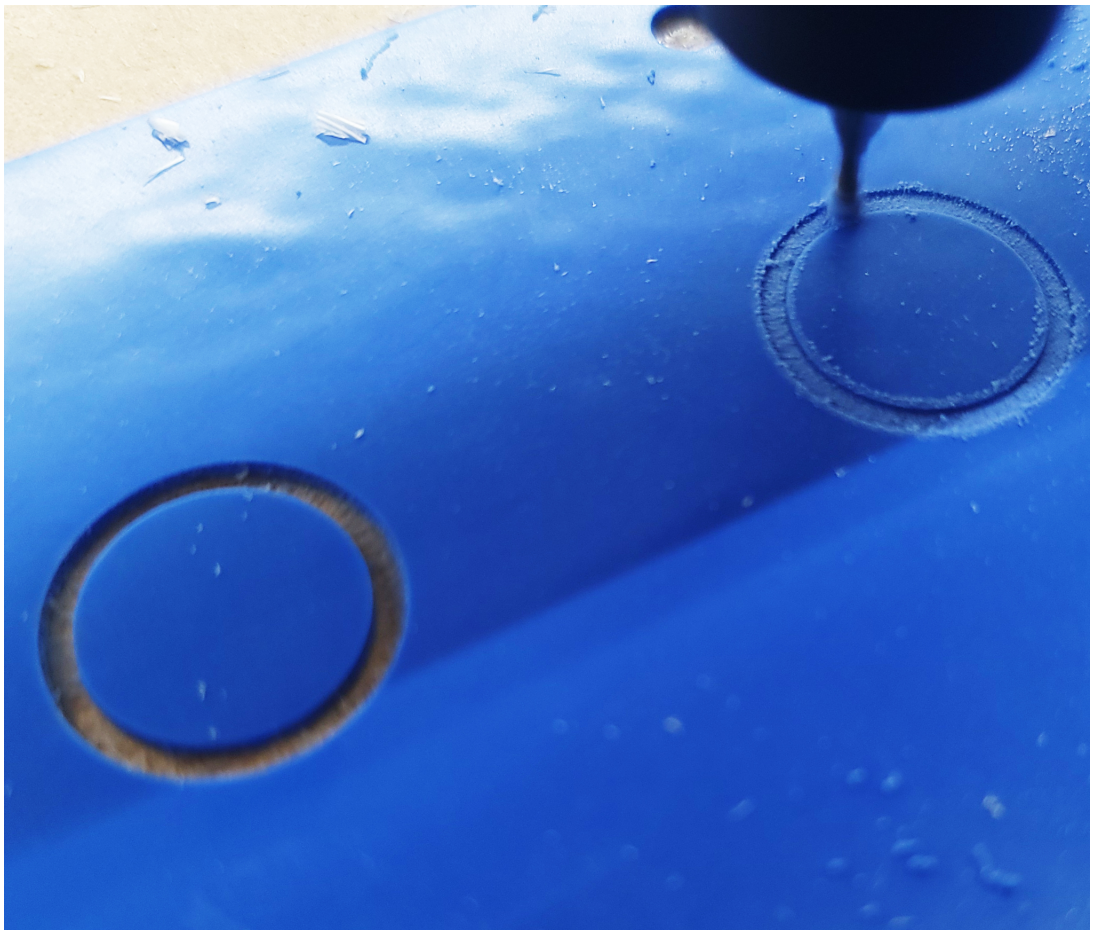


Figure 9.9: **Fluorosilicone gasket milling** Fluorosilicone gasket milling was achieved using a combination of shallow, <50 μm passes and the use of a $\frac{1}{16}$" downcut geometry square single-flute endmill to minimize lifting forces. On the right, the endmill makes repeated shallow passes to remove material. On the left, we see the final appearance of the cut with precise geometry and smooth walls.

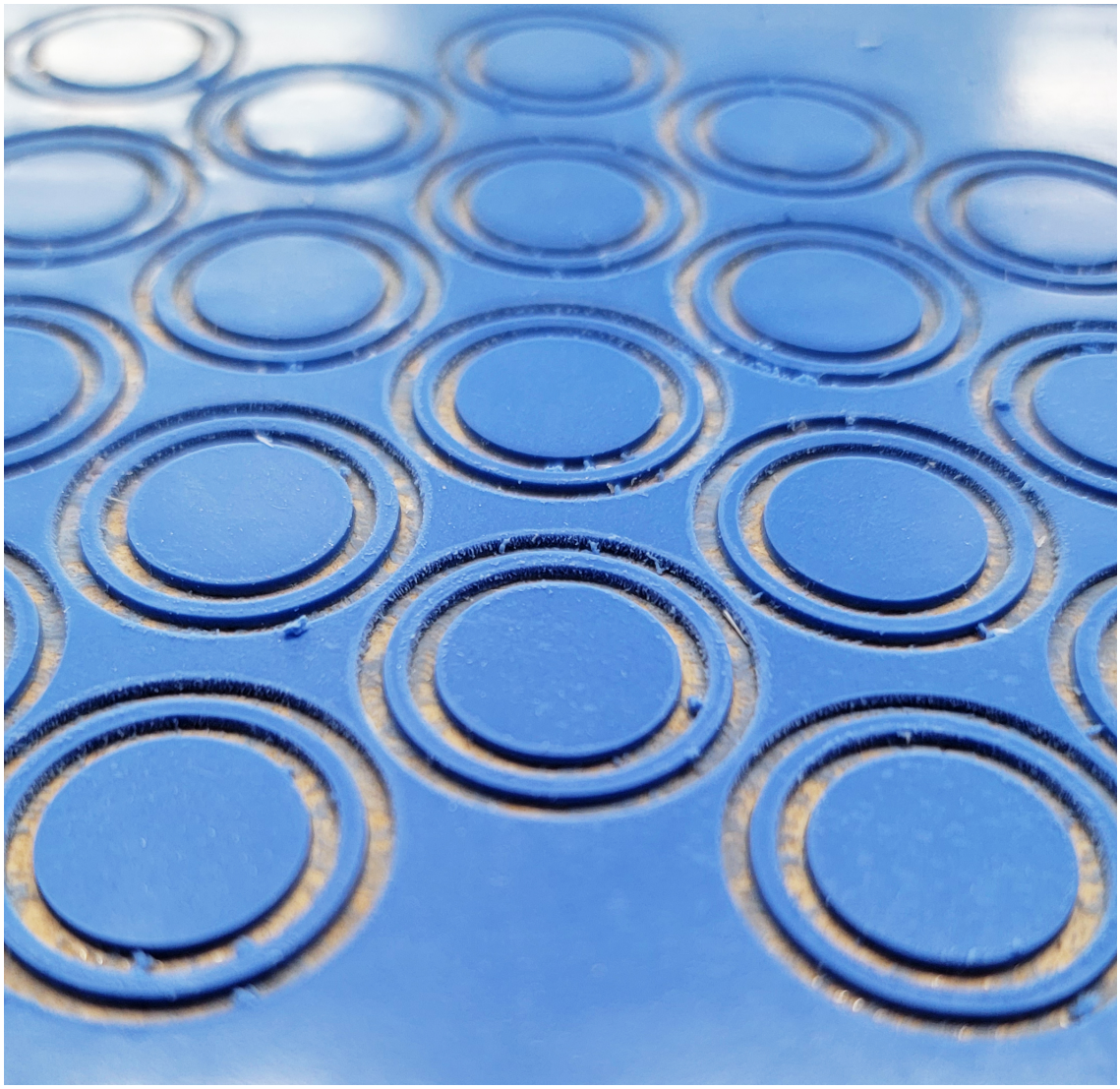


Figure 9.10: **The finished product from CNC milling of a set of fluorosilicone gaskets** Note the precise, smooth contours of the gaskets. Some debris is present, this can be easily removed using a rolling motion with a finger. This can also be removed using compressed air, though doing this manually is more laborious due to the small size of the gaskets. A tumbling or a liquid immersion method could also theoretically be used for part cleanup.

9.3 Design and manufacturing of fluidic actuators

For fluidic cell culture systems which involved gas exchange remote from the cell culture chamber, fluidic actuators were needed to supply gas equilibrated media to the cells. Specifically, a fluidic system with a flowthrough recirculating gas exchanger requires that the media be circulated between the cell culture compartment and the gas exchanger in a continuous flow. A fluidic system with a gas preconditioning cell and subsequent metered media delivery to a cell culture compartment also needs a fluidic actuator which can transfer media from the preconditioning cell and replace it with fresh unconditioned media.

While there exist potential methods to provide continuous fluidic flow through switched collections of syringe pumps, one of the most straightforward methods of perfusing a recirculating device is through use of a peristaltic pump. A peristaltic pump is also appropriate for a fluidic system which must both deliver fluid from a chamber and replace the fluid in said chamber simultaneously. This is because a peristaltic pump is naturally capable of acting as a transfer pump, simultaneously pushing fluid from and pulling fluid into a chamber. If syringe pumps were to be used, at least two syringe pumps and some valve manifold would be needed

to perform these tasks simultaneously. Other pump types such as those which use impellers can also act as both recirculating and transfer pumps, however these are generally more complicated than peristaltic pumps. Indeed, centripetal impeller pumps are only used over peristaltic pumps in human artificial lung applications due to the mechanotrauma to red blood cells experience during use of peristaltic pumps. Otherwise peristaltic pumps are simple to manufacture and have the added benefits of containing no wetted components of the tubing and furthermore, assuming adequate occlusion, of acting as fluidic pinch valves when they are not moving.

While there exist open hardware peristaltic pumps in the literature [8], these are generally complex and laborious to build, consisting of an unnecessary number of parts, and generally have 3D printed plastic rotors with inherently poor durability. Therefore, we have devised a device which is constructed simply from two simple laser cut metal parts, one CNC milled part and a small number of standard hardware store components. Since the part is laser cut from metal rather than 3D printed from plastic, the manufacturing dimensional tolerances and durability are superior to previously described open hardware pumps. Furthermore, the labor cost of manufacture is much lower due to the extremely simple assembly process and the rapid, highly parallelizable manufacturing techniques employed.

We have designed a novel fixed-diameter peristaltic pump designed around a NEMA 17 standard stepper motor and standard bearings. We specifically chose inexpensive and widely available standard motor and bearings for our design. A triangular rotor was designed and then ordered laser cut in 304 grade stainless steel from an external laser cutting bureau (SendCutSend, Reno NV.) The rotor incorporated a hole custom fit to the standard halfmoon geometry driveshaft of the NEMA 17 stepper motor around which three mounting screwholes were po-

sitioned at uniform distances and angular positions from the axis of the driveshaft. The screw-holes were manually tapped to an M3 thread. Thread tapping was performed manually in a single downstroke using a Tapmatic Nº 13033 RX 30 self-reversing tapping head and self-clearing black oxide cobalt steel through-hole threading machine taps obtained from McMaster Carr. Standard Nº 693ZZ 3mm x 8mm x 4mm bearings were mounted to the rotor by clamping their inner body against the rotor using conical flat head M3 screws.

This rotor was then mounted to the NEMA 17 motor driveshaft atop an M5 PTFE washer (McMaster Carr), and then fixed in place using an M5 drill depth stop (McMaster Carr). The drill depth stop was found to be the simplest and most cost-effective method of fixing the rotor in place. Alternatively the rotor could be brazed, welded or epoxied to the driveshaft to further simplify assembly.

A shell and backplate for compression of the peristaltic tubing was CNC milled from polyoxymethylene or polycarbonate stock. CNC milling was performed with a single flute $\frac{1}{8}$" square upcut endmill on a Carbide3D Nomad 3 desktop mill. The shell consisted of a semicircular backplate spanning 180°with a diameter chosen such that the rollers would produce full occlusion when pressing the tubing against the backplate. We determined this diameter first by calculating widest fully occlusive diameter of the backplate, then producing a number of versions of this device differing by increments of 25 μms from this diameter, then testing the pumps with an inline fluid flow meter to determine the minimum diameter at which no backflow was observed.

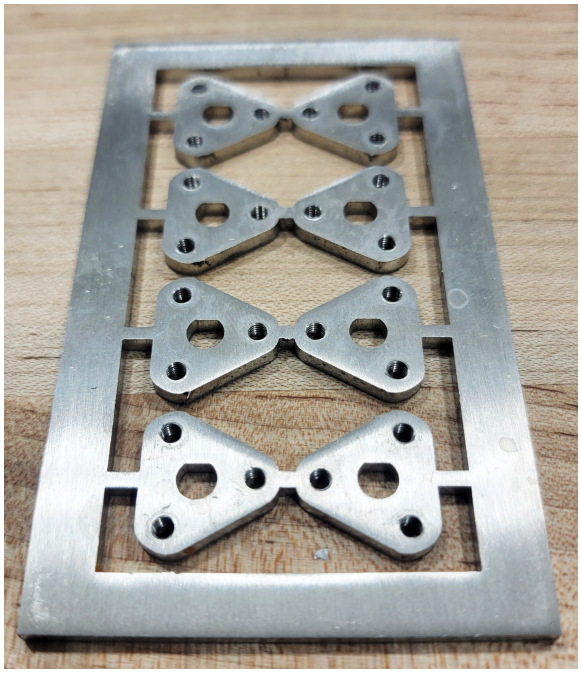


Figure 9.11: **Laser cut and tapped sprued stainless steel rotor bodies** Note the half-moon driveshaft hole. Laser cutting and tapping operations are performed with multiple sprued parts for ease of handling, then cut out using a rotary tool.

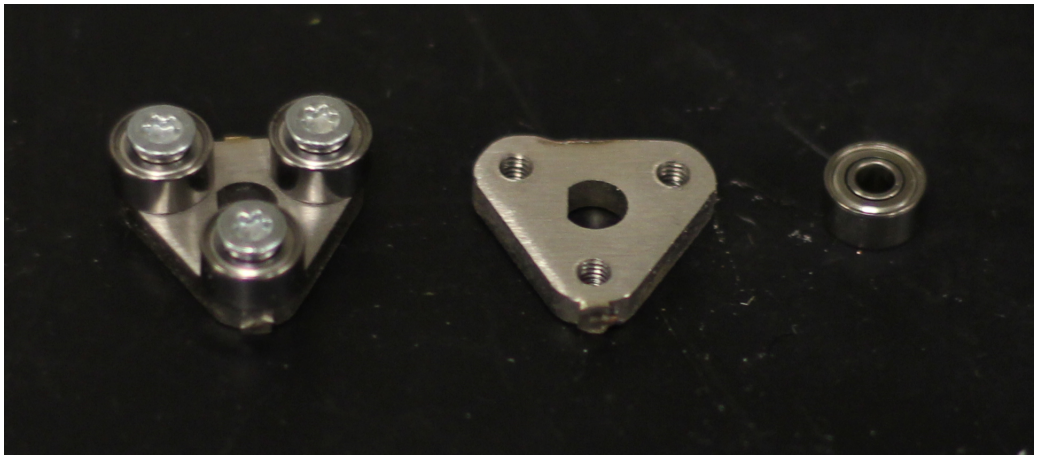


Figure 9.12: **Peristaltic pump rotor and assembly** 693ZZ bearings (right) are mounted to tapped rotor body (middle) and clamped to the rotor with conical flat head screws to form a 3-bearing, fixed diameter peristaltic pump rotor (left).

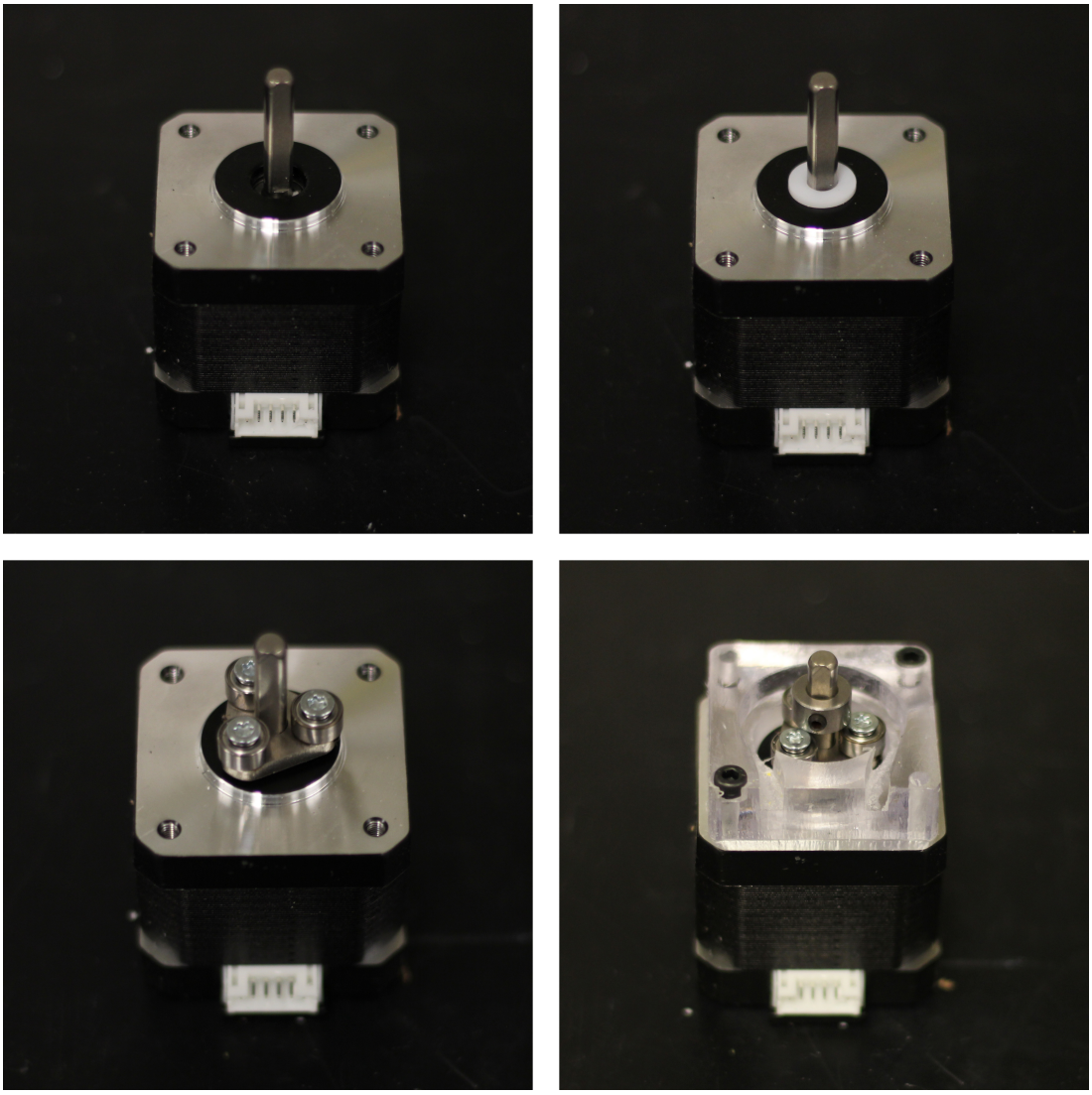


Figure 9.13: **Assembly of the peristaltic pump** A) A NEMA 17 motor is used as the basis for the peristaltic pump. B) First, a PTFE washer is placed on the driveshaft. C) Next, the rotor (described above) is mounted on the driveshaft. D) The CNC milled pump backplate assembly is mounted to two screwholes of the NEMA 17 motor. The depth stop is mounted on the driveshaft to hold the rotor in place. E) The peristaltic pump tubing is mounted and a steel plate is mounted over top of it to hold it in place, using the two remaining screwholes of the NEMA 17 motor.

9.4 Design and manufacturing of gas exchange devices

9.4.1 Gas exchange for recirculating fluidics

9.4.1.1 Gas exchange device design

Our approach to gas exchange in a continuous high-flow system is inspired by the device used for the equivalent purpose in intensive care medicine. In the case of respiratory or circulatory failure such that the lungs cannot be used for uptake of oxygen and/or removal of carbon dioxide, extracorporeal membrane oxygenation (ECMO) is used to perform these functions using a device outside of the body [129, 61, 118].

In the human medicine setting, the goal for carbon dioxide exchange is generally to remove carbon dioxide from the body: while the peripheral circulation is, under normal conditions, dramatically enriched in carbon dioxide versus atmospheric carbon dioxide, without egress, such as in the case of respiratory failure, the carbon dioxide concentration will rise

beyond physiological levels and begin to cause toxicity.

In general, the ratio between the metabolically active (and therefore carbon dioxide producing) tissue and intravascular space is much lower, such that the quantity of carbon dioxide released from the tissues is too low to make up the difference between physiological and atmospheric concentration. Therefore, in our case, our goal with carbon dioxide is precisely the opposite of standard ECMO: we wish to use our gas exchange device to add carbon dioxide into the luminal space in which our cell cultures reside. This is necessary to complete the buffering function of the media's carbonate-bicarbonate buffer system.

In human medicine, the extraorganoid membrane oxygenator uses one of three materials: silicone, microporous polypropylene and polymethylpentene. In current clinical application, all devices are manufactured using polymethylpentene oxygenation cells, indeed, we were inspired by their design.

We opted for a design which offered lower gas exchange surface area but was significantly simpler to manufacture than the hollow-fiber oxygenators of ECMO machines. Instead, our oxygenator simply consists of a high aspect serpentine channel whose roof offers a large surface area exposed to a thin layer of fluid. While this offers a smaller surface area to volume ratio than a hollow fiber gas exchange module, it is simple to build and does not require difficult to source specialty material.

9.4.1.2 Gas exchange device manufacturing

Briefly, the device was manufactured analogously to the previously described two-dimensional cell culture devices. Both the high aspect ratio serpentine channel and the gas

exchange fluid reservoir were CNC milled from polycarbonate. The part was post-processed, polymethylpentene film was incorporated and the part was sterilized using ozone as previously described.

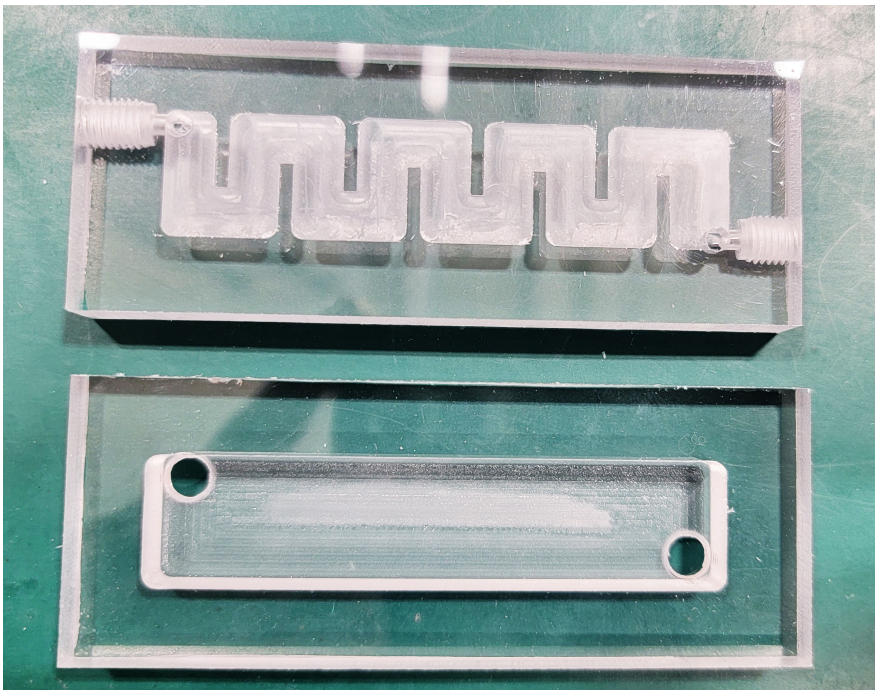


Figure 9.14: **Subcomponents of recirculating gas exchange device** The recirculating gas exchange device was milled in two components from polycarbonate. The top subcomponent consists of two $\frac{1}{4}$"-28 UNF fluidic ports feeding into a high aspect ratio serpentine channel. The channel has a 400 μm depth and a 6.4 mm width. The bottom subcomponent consists of a large reservoir for gas control fluid, also equipped with two $\frac{1}{4}$"-28 UNF fluidic ports. The open base of the gas control fluid reservoir and the serpentine channel occupy rectangles of the same dimensions with the same wide flange around them. The flange is used to laminate a polymethylpentene gas exchange film covering these rectangular regions between the two subcomponents.

9.4.1.3 Fluidic circuit design

A peristaltic pump was assembled from a stepper motor peristaltic pump and $\frac{1}{16}$" inner diameter Pharmed BPT peristaltic pump tubing as previously described. A Pololu Tic T825 stepper motor controller, controlled via computer via USB was used to control the peristaltic pump. The stepper driver was set to $\frac{1}{32}$ microstepping to reduce fluctuations in fluid velocity and noise and to 500 mA current limiting to limit heat.

A loop circuit was designed and constructed, consisting of the cell culture well, the peristaltic pump, the membrane oxygenator and inline sensors as needed. Either $\frac{1}{16}$" inner diameter Tygon E-3606 non-DEHP plasticized polyvinylchloride fluidic tubing, or $\frac{1}{16}$" inner diameter polyvinylidene fluoride nonplasticized tubing was used for the circuit components. Luer lock to $\frac{1}{16}$" barb fluidic connectors, $\frac{1}{4}$"-28 UNF compression fittings and adapters between the two were used for tubing assembly.

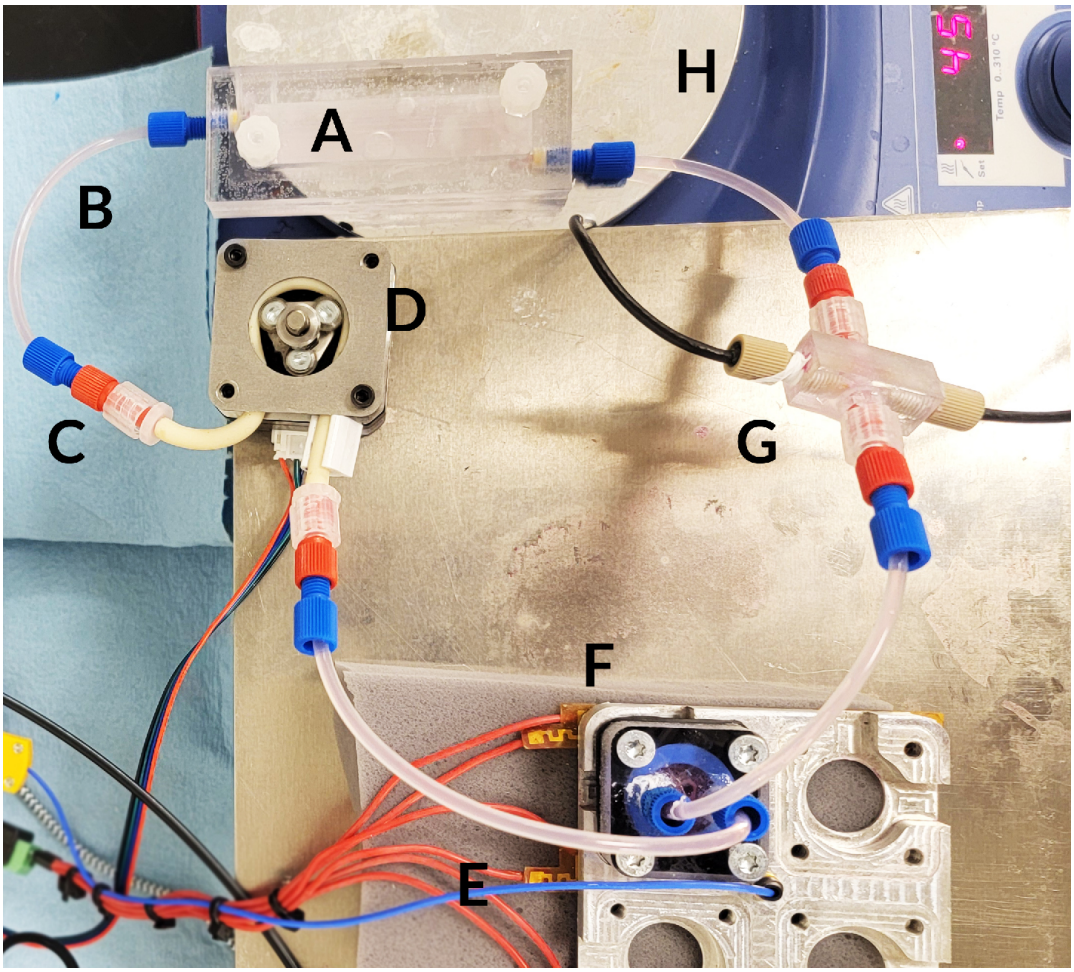

A
B
C
D
E
F
G
H

Figure 9.15: **Implementation of recirculating fluidic circuit** Counterclockwise from top: A) Membrane respiration device. B) Polyvinylidene fluoride (gas impermeable) tubing C) $\frac{1}{4}$"-28 UNF compression fitting for polyvinylidene fluoride tubing and adapter to hose barb fitting for PharMed BPT peristaltic pump tubing. D) Stepper motor based, computer controlled peristaltic pump, designed and assembled as described previously. PharMed BPT peristaltic pump tubing. Controlled via computer using USB stepper driver module. E) Resistive heating apparatus and heat block for organoid wells, as previously described. Under computer monitoring and control via USB interface. F) Organoid well, attached to circuit via $\frac{1}{4}$"-28 UNF chromatography compression fittings. G) Inline pH monitoring module. H) Ika RCT safety hotplate used as interim heating device for maintaining temperature of membrane respiration device; future prototypes will use a heat block which dually heats the cell culture well and membrane respiration block.

9.4.2 Gas exchange for non-recirculating fluidics

9.4.2.1 Design

The basic idea of the gas preconditioning device for provision of regulated gas levels for cell culture in non-recirculating fluidic systems is that, similar to with the 2D incubator free cell culture devices, a thin media compartment with a high surface area gas exchange membrane roof will rapidly equilibrate its gas concentrations with a gas control fluid on the other side of this membrane. We augment the design from the 2D incubator free cell culture device by incorporating a captive stir bar which will allow for more rapid equilibration of the media in the gas preconditioning chamber.

9.4.2.2 Manufacturing

Briefly, the device was manufactured analogously to the previously described two-dimensional cell culture devices. The main manufacturing difference was that prior to polymethylpentene film incorporation and final solvent welding assembly, a captive 1.5 mm x 15 mm magnetic stir bar was inserted into the media reservoir of the device. Devices were sterilized by ozone sterilization, it is important to note that in general, magnetic stir bars are incompatible with high heat sterilization methods such as autoclave or dry heat sterilization.

9.4.2.3 Fluidic circuit design

A peristaltic pump, constructed as previously described, was used for fluidic actuation. An Arduino microcontroller was used to drive the motor using a Pololu № TB67S249FTG stepper motor driver. Control software was written in order to provide intermittent feeding at specified volumes and frequencies.

A fluidic reservoir constructed from a glass GL45 vessel, a $\frac{1}{4}$"-28 UNF compression fitting and a $\frac{1}{4}$"-28 UNF luer adapter equipped with a 0.1 μm microporous filter. $\frac{1}{16}$" inner diameter Tygon E-3606 non-DEHP plasticized polyvinyl chloride fluidic tubing was used to route fluid from the reservoir into the gas preoxygenation device. The preoxygenation device was placed on a safety stirring hotplate (IKA RCT) for heating and stirring; future iterations will use lower cost homebrew heating and stirring analogous to the cell culture heating devices previously described. The outlet to the preoxygenation device was attached to the peristaltic pump using fluidic tubing. In turn, the outlet to the pump fed into the cell culture vessel. From the cell culture vessel, a final line led to a waste reservoir. In future implementations, the outlet

will also lead into the same peristaltic pump via a second pump channel; this will allow the pump to function as a valve to prevent either backflow and contamination into the system as well as prevent escape of fluid via hydrostatic fluid difference in case of seal failure in the system.

9.5 Cell culture and experimentation

9.5.1 Experimental design

Our initial set of experiments described herein primarily focus on the functional validation of the device in terms of temperature and chemical parameters, and only secondarily on cell culture. The extent of our analysis of cell culture at this stage in development will be to qualitatively determine whether or not the system is immediately toxic to the cells under culture.

Our proof of concept measurements of the performance of the system will include measurement of heat block and intra-well temperature stability, measurement of carbon dioxide gas exchange and maintenance in a recirculating oxygenator system, measurement of carbon dioxide gas exchange rate in a non-recirculating gas preconditioner system, and measurement of stability of carbon dioxide concentration in a sealed well, a marker of the performance of the double gasketed system for preventing gas exchange.

We will furthermore qualitatively assess the system for leaks, bubble trapping and ability to circulate media.

9.5.2 Cell culture

Mouse iPSC derived organoids were obtained from the Mostajo-Radji lab. These were transferred from their facility in Sasai 3 media, then kept in the incubator in a closed conical tube until loading.

9.5.3 Cell culture device loading and assembly

A titanium bitholding precision screwdriver, a Wera Torx 20 hex bit, a Wera Torx 20 L-key and a Knipex Cobra wrench were sterilized by dry heat at 450 °F in an instrument tray in a convection oven for 30 minutes. All other parts were sterilized by ozone as previously described.

The bottom halves of the cell culture wells were mounted to the heat block inserting them into the cavities and seating them by press fit. The butyl rubber outer gasket and inner fluorosilicone gasket were placed in their respective grooves, aligned and pressed lightly into place.

Organoids were transferred by aspiration into a P1000 pipette set to 200 microliters volume and deposited into wells. Due to the small size of mouse cortical organoids, this could be achieved with standard P1000 pipette tips; for human or other larger cortical organoids, tips cut to a larger orifice size might be necessary. After loading, the pipette was used to remove media to such a level that the organoids were just covered, such that the gaskets would not become wetted before the top of the well was installed.

The tops of the wells were placed loosely on the gaskets and aligned with the bottoms. Four Torx 20 drive M4 screws were inserted in the holes in the well top, then screwed into the

receiving threads of the heat block. These were tightened loosely, then fully screwed down in a Z-pattern order until they evenly pressed down the gaskets throughout the well. The well was then partially filled with an additional 200–400 μL of media.

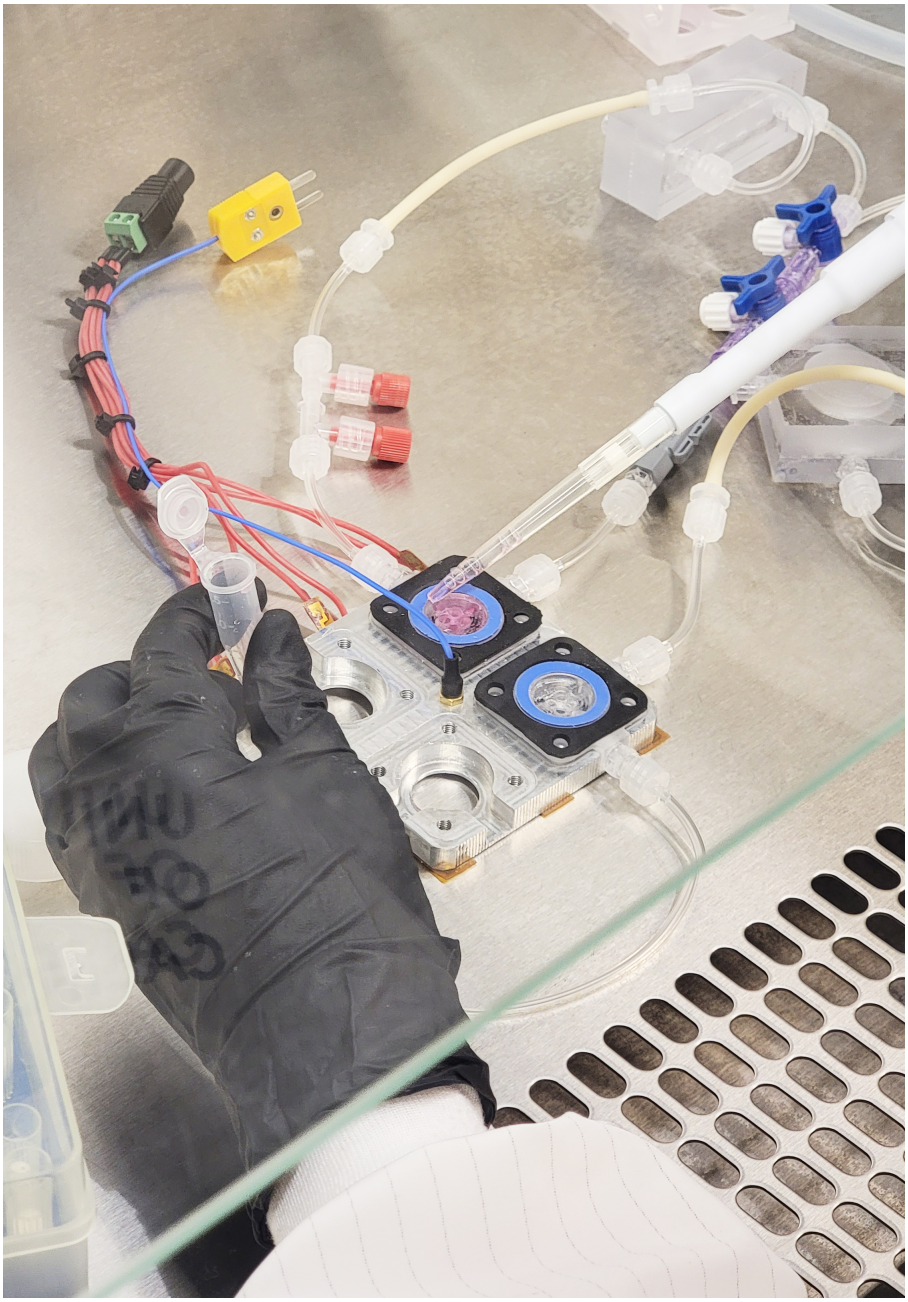


Figure 9.16: **Organoid culture device organoid loading**

9.5.4 Automated intermittent feeding circuit assembly and priming

The outflow and inflow lines to the well were attached. The outflow line was fitted to a waste reservoir. The inflow line was attached to the outflow port of the gas exchange preconditioning reservoir. The inflow port of the gas exchange preconditioning reservoir was attached by Tygon tubing and double sided $\frac{1}{16}$" hose barb to coupling to the peristaltic pump tubing, which was in turn connected by $\frac{1}{16}$" hose barb coupling to the outflow from the media reservoir.

9.5.5 Recirculating circuit assembly and priming

Organoids were loaded into the cell culture chamber by the same procedure previously described.

A fresh sterile syringe of media was drawn, and any bubbles were ejected. The outflow line from the cell culture chamber was removed from the pump and the syringe was attached in its place. The outflow line was positioned to drain to waste. The peristaltic pump was turned on, and the plunger on the syringe was slowly depressed at a rate such that fluid just began to flow in a pulsatile manner. This continued until the line was primed without air bubbles. At this point, the pump was stopped, the outflow line was clamped to full occlusion, the syringe was removed and discarded, and the outflow line was reattached to the pump.

The circuit was observed while it circulated to confirm that no bubbles were present in the lines or chambers. Small bubbles (discrete 2 mm bubbles in lines or chambers) were

allowed to persist; in the case of large bubbles causing segments of fluid-free line or bubbles of 4–5 mm or larger in diameter, the bubbles were allowed to circulate until the coalesced at the end of the outflow line, and then a small-volume re-priming according to the same procedure was used to flush them out by injection of additional sterile bubble-free media.

9.5.6 pH Monitoring

$\frac{1}{4}$"-28 UNF micro pH probe sets, pH sensor circuit, USB to serial converting circuit and calibration solutions were obtained from Atlas Scientific. pH probes were calibrated immediately before each experiment. Data was collected at 1 second intervals using a desktop computer running Termite 3.4 (CompuPhase.)

For the recirculating circuit, a flow-through pH probe was constructed by CNC milling of a cross-shaped 4-port $\frac{1}{4}$"-28 UNF threaded manifold. The two micro probes were inserted in opposing ports along one axis of the manifold, and inlet and outlet ports were inserted across the perpendicular axis. Threads were sealed with polytetrafluoroethylene tape and a thin outer backup layer of cyanoacrylate glue.

For the gas preconditioning device, a version of the device was designed having an additional two $\frac{1}{4}$"-28 UNF threaded channels into which the probes were inserted.

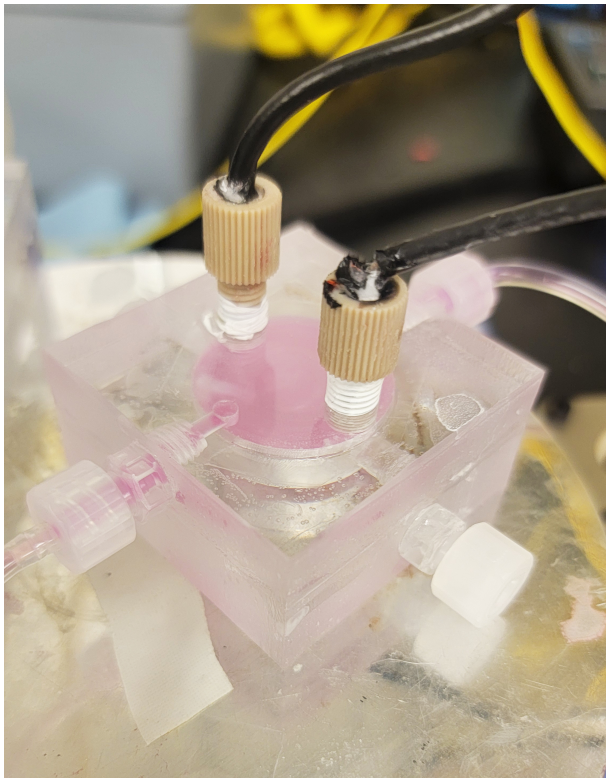


Figure 9.17: **Experimental monitoring setup for pH monitoring of gas preconditioning cell** Two additional $\frac{1}{4}$"-28 UNF threaded ports have been incorporated into one surface of the media chamber of the gas preconditioning device. This allows the tips of $\frac{1}{4}$"-28 UNF pH and reference electrodes to be fully immersed in the media chamber, directly measuring its pH while it operates.

9.5.7 Temperature Monitoring

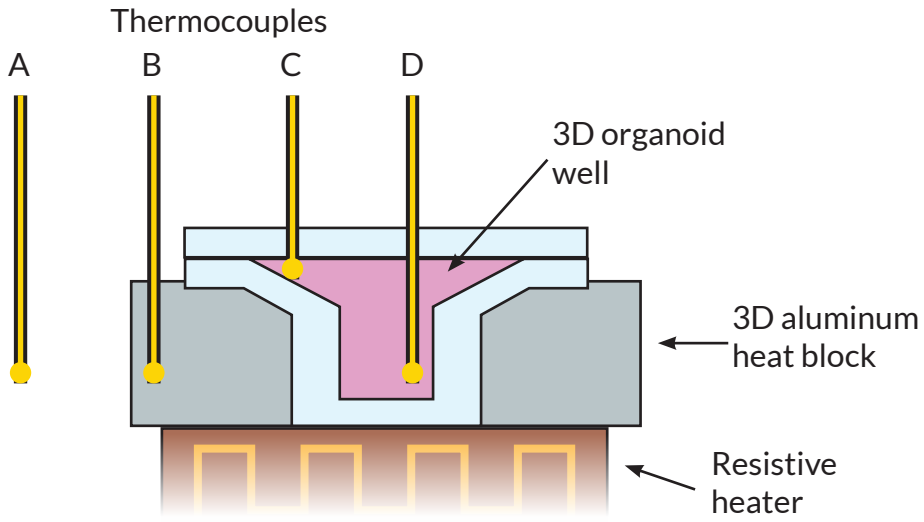
The upper access ports of the cell culture wells were opened and thermocouples were inserted to specified locations and depths. Parafilm plugs were constructed to seal the wells from evaporation and prevent the thermocouples from shifting. A 4-channel K type thermocouple

data logger (Gain Express) was used for temperature logging.

9.6 Results

9.6.1 Temperature regulation

Four thermocouples were placed as pictured in a pair of incubator-free 3D organoid wells. Fluid was not circulated. A temperature setpoint of 37.8 °C was programmed; this temperature was experimentally found to correspond to an internal well temperature of 37.0 °C at normal room temperature of 22 °C. The PID controller was tuned to undershoot this setpoint to avoid oscillations above physiological temperature. 1 inch of melamine foam insulation was placed below the heat block and above the roof of the cell culture devices; the sides of the heat block and cell culture devices were left open to room air.



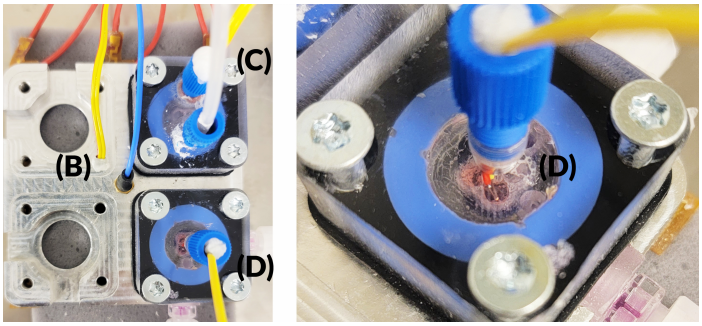


Figure 9.18: **Positioning of thermocouple probes** Previously used wells were cleaned and repurposed for this experiment. Top: cartoon of positioning of thermocouples. Thermocouple (B) is inserted into a screwhole of a heat block. Thermocouple (C) is inserted into the media volume at the top of the well, to measure the bottom to top temperature gradient. Thermocouple (D) is inserted into the media volume at the bottom of the well, where an organoid would sit. Bottom left: Positioning of thermocouples (B)-(D) (yellow wires). Bottom right: detail of positioning of thermocouple (D) within the well.

Temperatures of all four thermocouples were recorded over a period of 4.5 days, and the results are described below.

Ambient temperatures varied between 16.6 °C and 22.7 °C (50% of temperatures between 20.4 °C and 22.5 °C). Despite this, the bottom of the well, where the organoid sits, maintained a temperature between 36.3 °C and 37.6 °C, with a mean and median of 36.9 °C, 50% of temperatures lying between 36.8 °C and 37.0 °C and 99% of temperatures lying between 36.6 °C and 37.1 °C.

Temperatures of the heat block varied between 37.1 °C and 37.8 °C, with a mean and median of 37.6 °C and 50% of temperatures lying between 37.5 °C and 37.7 °C. The

temperature difference between the heat block and the media at the bottom of the cell culture well was between 0 °C and 1.2 °C, with a mean and median of 0.7 °C, 50% of differences between 0.7 °C and 0.8 °C and 99% of differences between 0.5 °C and 1 °C.

Temperatures of the media at the top of the well (distant from the organoid) varied between 34.6 °C and 35.9 °C, with a mean of 35.6 °C and a median of 35.8 °C. 50% of temperatures were between 35.4 °C and 35.9 °C, and 99% of temperatures were between 34.7 °C and 35.9 °C. The temperature difference between the bottom and top of the well was between 0.8 °C and 2.2 °C, with a mean of 1.3 °C and a median of 1.2 °C. 50% of differences were between 1.3 °C and 1.9 °C and 99% of differences were between 1.1 °C and 1.3 °C. This >1 C temperature gradient may be considered desirable since there is a potential benefit to having a small temperature gradient established within the media of the well: this encourages convection which in turn contributes to nutrient and waste circulation within the cell culture chamber.

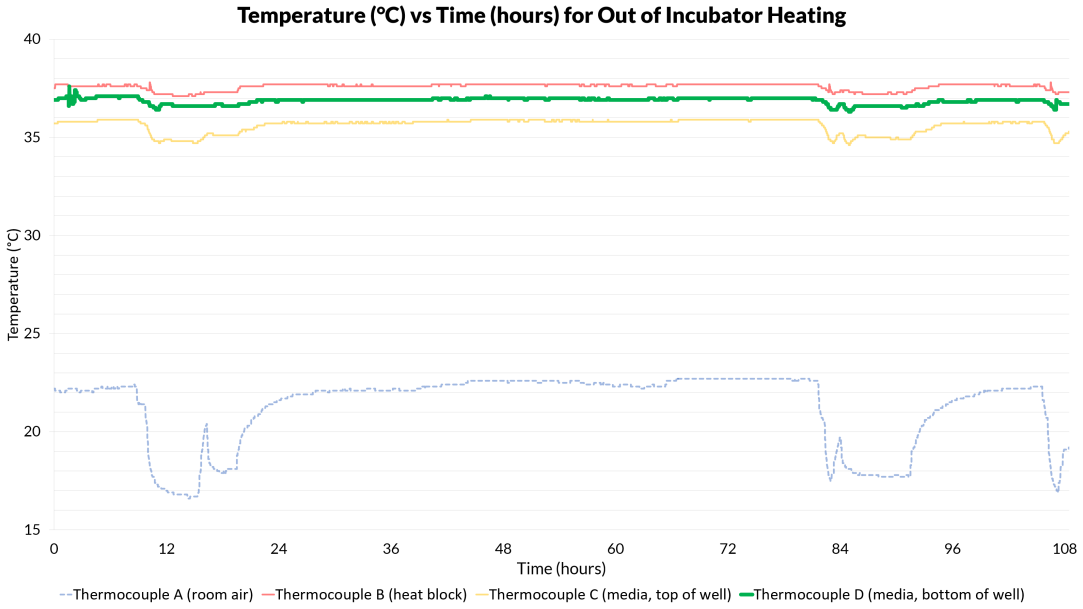


Figure 9.19: **Temperature stability of 3D heat block temperature control for incubatorless organoid culture** Measurements were taken over 4.5 days. Temperature traces correspond to thermocouple placements described and pictured above. Thermocouple A (room air) was suspended above the lab bench and did not contact any solid surface. This experiment was run over a weekend; on weekdays the laboratory thermostat decreases room temperature from approximately 22 °C to approximately 17–18 °C overnight. Note muted drops in temperature of interior of well relative to drops in room temperature.

Inspection of a subinterval of measurements during which ambient temperature was stable within 0.1 °C of 22.7 °C reveals stable in-well and heat block temperatures, with up to 0.1 C deviation from a median temperatures of 37.7 °C for the heat block, 37.0 °C for the media at the bottom of the well and 35.9 °C for the media at the top of the well. The differences between the temperature measurements also remained stable in this interval. See figure below.

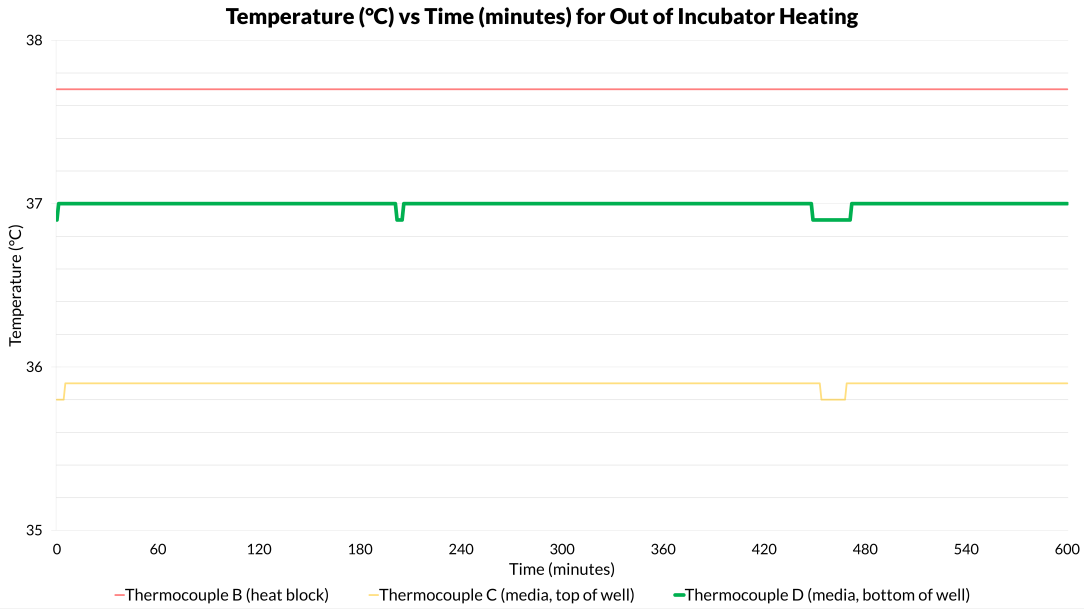


Figure 9.20: **Temperature stability of 3D heat block temperature control for incubator-less organoid culture during subinterval with constant ambient temperature** A 10 hour subinterval of the previously shown data wherein ambient temperature was uniform at 22.7 °C is pictured, showing a uniform bottom-of-well temperature of 37.0 °C, with only intermittent, brief dips to 36.9 °C and similar temperature uniformity in the range of 35.9 °C for top-of-well temperatures.

The data obtained displays a semblance of a monotonic relationship between the deviation of the bottom-of-well temperature (and heat block and top-of-well temperature) from target and the deviation of the room temperature from its daytime average. This suggests that the temperature gradient between these locations is dependent on ambient temperature; this is expected since heat dissipation from the surfaces of the heat block and device should increase with drop in ambient temperature. Further computational simulation and modelling is needed to develop a predictive model for this difference. We defer this until after further optimization of the geometry, thermocouple location and insulation of the device.

9.6.2 Peristaltic pump flow velocity and backflow

Our peristaltic pump design was validated by flow direction and velocity measurement. We expect, given proper occlusion of the peristaltic pump tubing, that the pump should exhibit no backflow as well as peak and average flow velocities proportional to the angular velocity of the motor which drives it. Fluctuations in fluid velocity are expected as a characteristic inherent to all peristaltic pumps; with our three-roller pump design, we can expect an oscillatory

waveform with three peaks per revolution.

Flow rate was measured using a Sensirion SLF3S-1300F bidirectional flow meter, using the Sensirion USB interface and Control Center software. Flow rate was tested at angular velocities of 3.75 to 60 rpm, in 3.75 rpm increments. The flow rate waveform at angular velocity of 11.25 revolutions per minute are shown below. We observe no backflow at any tested angular velocity. The observed waveform (and all other waveforms below 60 rpm) has a major component with the three-peak oscillatory waveform expected. We note different peak velocities and waveforms between the three peaks. This is most likely attributable to small amounts of play and/or lack of centering due to oversized tolerances in the fit of the rotor to the driveshaft. We expect that reducing these tolerances would cause a more even waveform.

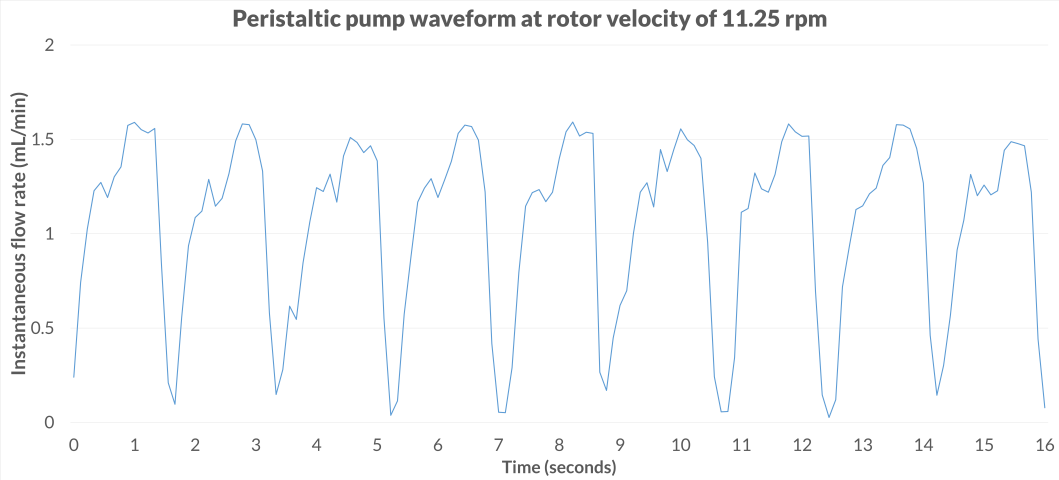


Figure 9.21: **Fixed rotor pump flow rate waveform, 11.25 rpm**

The relationship between pump rotor angular velocity and average pump flow rate is shown below. The expected linear relationship is seen between tested velocities of 3.75 and 60 rotations per minute. It became difficult to measure flow rate above 60 rpm on account of

the limited sampling rate of the Sensirion SLF3S sensor used: the nominal sampling rate of 10 Hz and the rotor pulse velocity of 3 Hz (3 pulses per rotation) are close enough in frequency to generate interference patterns in the sampled data, which is apparent by breakdown of the triphasic waveform geometry expected based for a peristaltic pump. Higher sampling rates flow sensors will be needed for characterization of our peristaltic pump at high angular velocities.

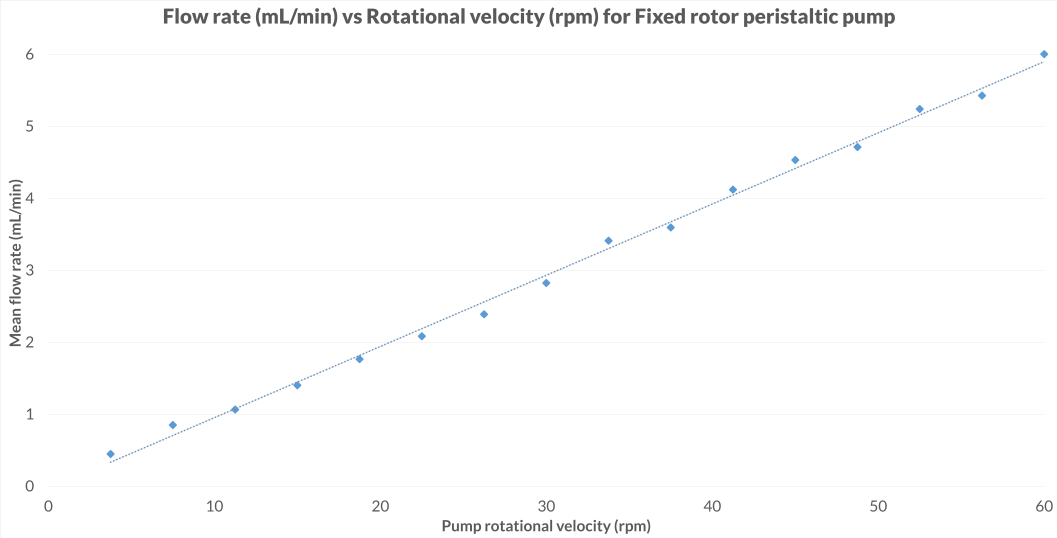


Figure 9.22: **Fixed rotor pump flow rate versus angular velocity**

9.6.3 pH regulation of cell culture media via gas preconditioning chamber in non-recirculating loop

A gas preconditioning device incorporating a pH monitoring apparatus was assembled as described above. A microcontroller controlling a peristaltic pump actuator of the type previously described was set to replace the fluid in the media chamber of the device on four-hour intervals. Media provided was Gibco DMEM supplemented with Primocin whose dissolved carbon dioxide concentration had been preequilibrated with room air by vigorous magnetic stir-

ring in a conical tube at 500 rpm for one hour. The media was kept at room temperature prior to loading into the gas preconditioning device to avoid excess dissolution of carbon dioxide. The gas control fluid chamber of the device was filled with freshly prepared gas control fluid consisting of 0.8 M sodium bicarbonate and 60 mM sodium carbonate. The device was maintained at approximately 37 C with 100 rpm magnetic stirring using an IKA RCT magnetic hotplate. The pH was measured at 1 second intervals using a Atlas Scientific Micro pH probe and reference sensors.

Initial media pH following media change exceeds 8, as expected on account of low carbon dioxide concentration of the media following equilibration with room air. Immediately following media change, we observe a monotonic decrease in pH as gas is exchanged with the gas control fluid chamber across the polymethylpentene film. The pH asymptotically approaches a value of 7.2. Approximately steady-state pH is established at approximately 180 minutes and remains steady to 240 minutes at which time the media in the gas preconditioning chamber is replaced.

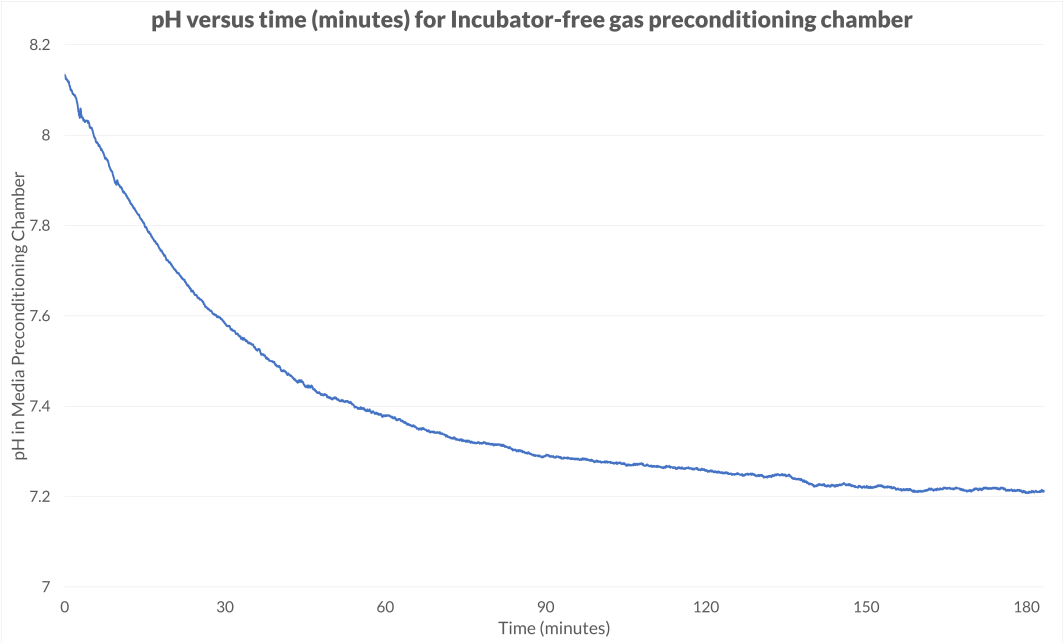


Figure 9.23: **Plot of pH inside of gas preconditioning chamber versus time since refilling of the chamber with fresh (equilibrated with room air, high pH) media** After approximately 180 minutes, the pH of the media in the preconditioning chamber reaches a steady state value of 7.21.

9.6.4 pH regulation of cell culture media via membrane respiration in recirculating fluidic loop

An inline pH meter apparatus was constructed as described previously and inserted in the recirculating fluidic loop described previously. The peristaltic pump, controlled by a desktop computer using a Pololu Tic T825 USB motor driver was driven at a constant rate of 96 rotations per minute. The recirculating gas exchange device was inserted into the circuit, with the media path flowing through the high aspect ratio serpentine chamber such that it was

spread out in a thin layer in contact with the polymethylpentene film. The gas control fluid chamber of the recirculating gas exchange device was filled with a freshly prepared mixture of 0.8 M sodium bicarbonate and 60 mM sodium carbonate. The circuit was primed with Gibco DMEM supplemented with Primocin whose dissolved carbon dioxide concentration had been preequilibrated with room air by vigorous magnetic stirring in a conical tube at 500 rpm for one hour. The device was maintained at approximately 37 C with 100 rpm magnetic stirring using an IKA RCT magnetic hotplate. The pH was measured at 1 second intervals.

Similarly to the pH response in the static gas preconditioning chamber, the recirculating gas exchange device showed a monotonic decrease in pH which started immediately after exposure of the media to the gas control fluid across the polymethylpentene film and asymptotically approached the target pH of 7.2. The pH equilibration kinetics and steady state pH were also similar, with a steady state pH of approximately 7.2 reached by approximately 180 minutes and thereafter maintained for the remaining 2 hours of the experiment.

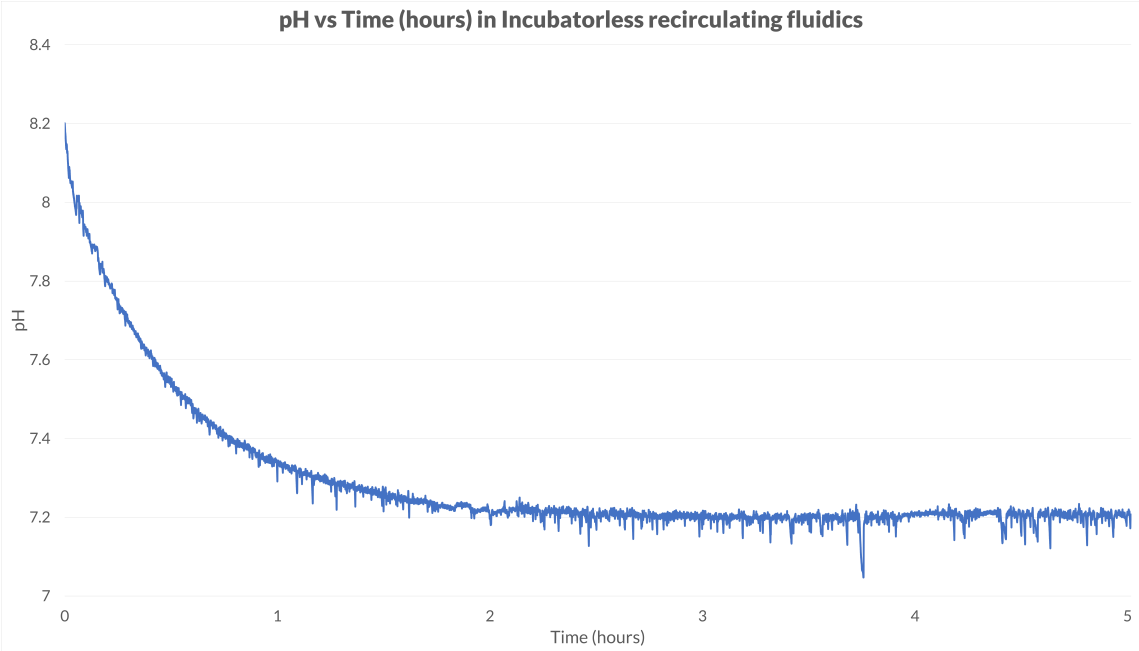


Figure 9.24: **Plot of pH in recirculating fluidic loop with inline membrane respiration device** pH is measured starting from circuit priming, and measured at 1 second intervals for the subsequent 5 hours.

9.6.5 Cell culture and imaging

Organoids were seeded in wells of the cell culture device as previously described. They were maintained with recirculating gas exchange for a period of 7 days; the gas exchange fluid was changed on the 4th day. pH was observed qualitatively by media color and based off of this color, pH was maintained in an acceptable range between 7.0 and 7.5. Further characterization using quantitative pH measurement will be needed in future work. Microscopy showed that organoids maintained their morphology, except for some spreading of cells onto the polycarbonate surface. An illustrative microscope image is shown below. Further or more formal

characterization of cell culture results is deferred to future work.

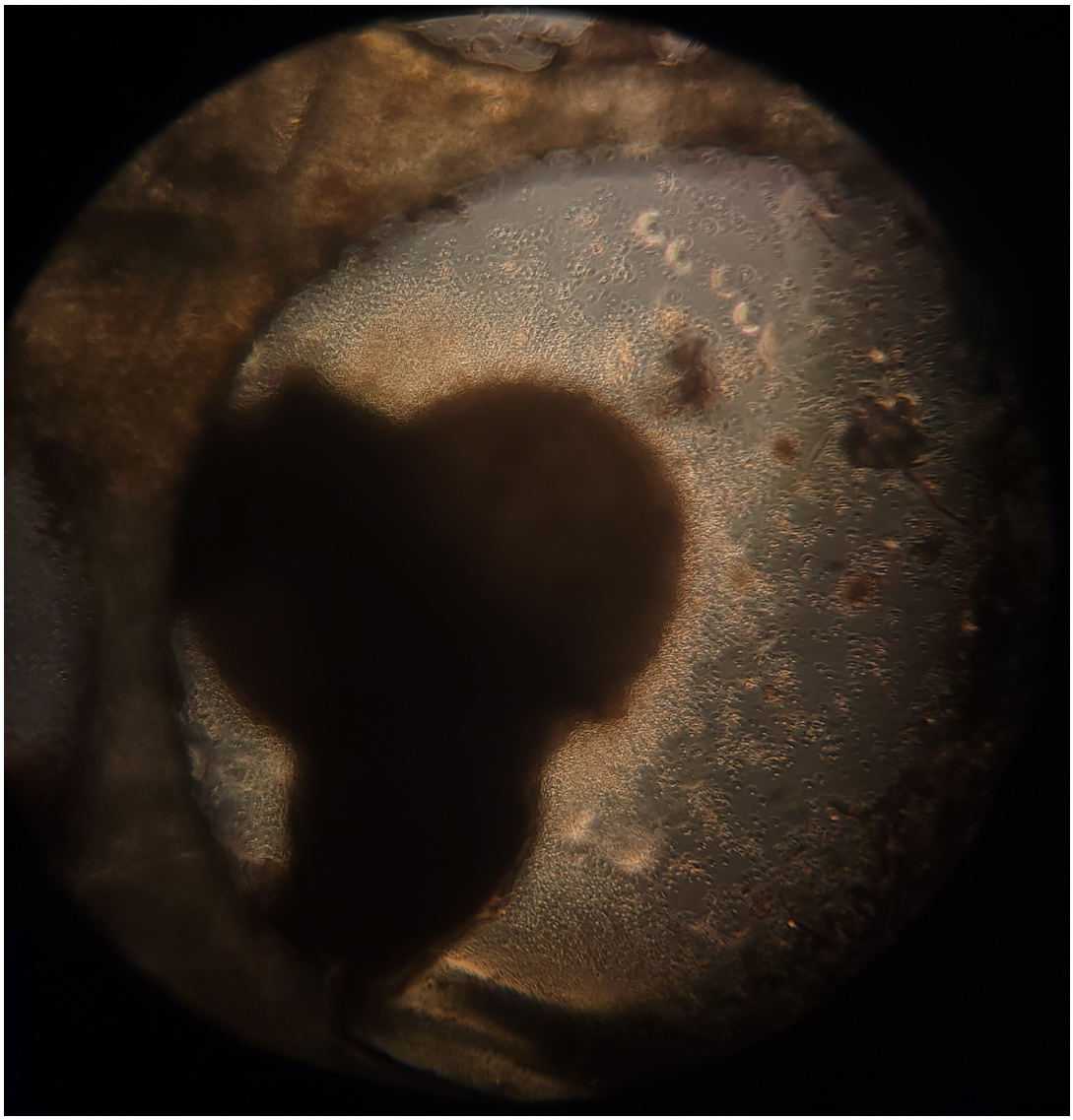


Figure 9.25: **Murine cortical organoid in incubator-free cell culture well viewed through microscopy window** The photograph was acquired on Day 3 of culture in the cell culture device, using recirculating perfusion culture supplied by an inline membrane oxygenation device. The sample was imaged using a Zeiss Axiovert 25 microscope. Digital photography through the eyepiece of the microscope was used to acquire the photo. We note gross integrity of the organoid as well as spreading of cells onto the (cell adhesive) imaging surface of the well.

Discussion

In the preceding part of this dissertation, I have described a vision for expanding the approach that biologists take to gathering data for their laboratory experiments by incorporating a "new bioinformatics" which adds data gathering instruments which provide high dimensional data (such as microscopy or electrode array data), continuously acquired timeseries data or a combination of both approaches. This approach could be transformative to biology, giving deeper insight into biological questions and also providing sufficient quantity and quality of data to apply modern computational statistical methods to biological discovery.

However, as discussed, improving the data gathering instrumentation of biological experiments will likely require that the standard procedures by which we do laboratory biology adapt to these tools. First, in a world where biological experiment units are always being observed, potentially by bulky or delicate equipment or even by a large mass of different instruments from all sides, the standard approach of using a simple petri dish or functionally equivalent flask and storing it in a bulky, limiting incubator which controls all environmental factors by manipulating an airspace is no longer the most appropriate method of maintaining cell cultures. Incubator space is limited real estate in the laboratory, and it is not practical to fill most of it with instrumentation as a standard practice. But, with current protocols, it is also

not possible to perform many, possibly longitudinal, data gathering operations with cell cultures outside of these incubators since this would require exposing the cell cultures under observation to long and repeated periods where their environment is not regulated. Furthermore, doing this sort of observation across many replicates also disrupts the shared incubator environment for all experimental units contained within. These difficulties alone necessitate exploring approaches to cell culture which do not require cell culture incubators for regulation of gas concentrations, temperature and osmolarity.

Second, I hypothesize that more widespread adoption of biological observation modalities which are aimed at observing functional or metabolic state of cells, for example live biochemical assays or elecrophysiological assays, will also require that the environmental fidelity of cell culture systems more accurately represent biological reality. I hypothesize that for many types of experiments, the relative lack of control over media osmolarity or pH seen in standard "air gap" incubator-based cell culture will no longer be acceptable due to how these factors directly impact the function and chemistry of the cell.

I have described potential methods to shift away from cell culture methods which are dependent on conventional incubators for gas control, osmolarity and temperature. Central to these methods is the idea of performing gas exchange through a nonporous, non water vapor permeable but oxygen and carbon dioxide permeable film window. The simplest devices built around this principle, which I call simply air gap-free, still use a regulated gas environment for gas concentration control but no longer rely on a humidified environment. At a middle level of sophistication, I describe self-enclosed incubator-free cell culture devices. These are standalone cell culture vessels which incorporate self-contained second chambers of fluid which regulate

the gas concentration in the cell culture chamber of the vessel thereby providing control over both gas concentration and osmolarity in the media, without requiring any external equipment at all other than for heating. The last, most sophisticated class of devices are devices which use similar fluid based, incubator free gas control mechanisms to the self-contained incubator-free vessels, but which have physically separated the gas control interface for the cell culture media from the actual chamber where the cell culture is performed. These devices rely on fluidic automation to function, adding complexity but in return they free up the real estate on all sides of the cell culture device for instrumentation. This is a very important step forward for our concept of highly instrumented cell culture because, using these sorts of cell culture devices, the only limitation on the scale and variety of instrument used to observe biology is that which is defined by the footprint of the biological entity under study itself.

Both of the approaches to incubator free cell culture will also be important advances when it comes to the biological fidelity of cell culture. These technologies will reduce the problems with evaporation inherent to any humidified chamber - air gap system and will also reduce the problems with shared incubators and their difficulties in maintaining temperature and gas concentration in the setting of high use. Beyond these things, these devices also present exciting opportunities to use experiment specific gas blends and even hydrostatic pressures. It is convention to use media at gas concentrations, pH and osmolarity approximating serum and atmospheric however it is well known that all of these variables differ considerably between different cell types. Ability to control these parameters on a per replicate basis “comes free” with the incubator-free technology implementations which I have invented.

In addition to describing the engineering and chemical principles which theoretically

allow for the three types of air gap-free and incubator-free cell culture devices described above, I have also shown functional prototype designs and methods for manufacturing them inside of a university engineering laboratory. These include a suite of CNC milling, laser cutting and solvent welding approaches which are, in their own right, a standalone contribution to the field of laboratory "homebrew" equipment manufacturing. The full implications of these techniques for laboratory biology plasticware and other devices are beyond the scope of this thesis but will be the focus of future publications. CNC milled and solvent-welded thermoplastic devices offer considerable cytocompatibility and gas permeability advantages over the standard paradigm of PDMS fluidic devices for biological applications. They are also arguably more straightforward to manufacture without specialist expertise or equipment. Furthermore, they can also be manufactured at scales down to the microfluidic realm and up to the size of petri dishes or larger, something which is not readily doable with PDMS. Similarly, the laser cut metal and CNC milled approaches to laboratory equipment manufacturing is a valuable tool in the maker biology armamentarium since it produces devices with better durability, stability and dimensional accuracy than conventional approaches such as 3D printing, the quality of the products of which has likely impaired widespread adoption of the "maker" concept in biological laboratories.

I have performed preliminary proof of concept experiments whose results show that these air gap-free and incubator-free devices do function as intended in terms of temperature and carbon dioxide concentration control. The pH and carbon dioxide control results for air gap-free and simple, self-contained incubator free devices support the notion that these devices are indeed exchanging gases freely across the gas permeable films contained within. It remains

to perform additional experiments to strengthen these data with increased time point measurements and replication. These will add kinetic information and statistical certainty to the proof of concept results shown in this thesis. We were not able to perform oxygen control experiments beyond simple, sparse time point observational studies; it remains to perform experiments using continuous oxygen level monitoring with an array of target oxygen concentrations.

Similarly, cell viability studies are outside the scope of this thesis but will be important to fully characterizing these devices. We plan to perform and publish results measuring the growth rates and cell viabilities in all described cell culture devices, using both two dimensional cultures and organoids.

Lastly, I have succeeded in demonstrating active osmolarity issues with conventional air gap, incubator based cell culture methods, but osmolarity results for our air gap free appear to show too much technical error in the measurements to be able to make any conclusions on the basis of the limited data we have. In order to make conclusions regarding whether we have actually fixed the osmolarity control problem in cell culture will require experiments with either more replication or more precise instrumentation. These experiments are outside the scope of this thesis however we will perform them for a future publication. When these experiments are performed, we intend to use continuous sodium or conductivity monitoring as a proxy for osmolarity in order to give increased statistical power as well as measurement of dynamics of evaporation.

Bibliography

[1] Mobile Industry Processor Interface Alliance. Camera serial interface 2 (MIPI CSI-2). https://www.mipi.org/specifications/csi-2. [Accessed 10 Aug 2022].

[2] Patrícia Alves-Oliveira, Matthew Bavier, Samrudha Malandkar, Ryan Eldridge, Julie Sayigh, Elin A Björling, and Maya Cakmak. Flexi: A robust and flexible social robot embodiment kit. In *Designing Interactive Systems Conference*, pages 1177–1191, 2022.

[3] National Electrical Manufacturers Association. NEMA ICS 16-2001 (Motion/position control motors, controls, and feedback devices), 2001.

[4] Tom Baden, Andre Maia Chagas, Greg Gage, Timothy Marzullo, Lucia L Prieto-Godino, and Thomas Euler. Open labware: 3-d printing your own lab equipment. *PLoS biology*, 13(3):e1002086, 2015.

[5] Chris Bakal, John Aach, George Church, and Norbert Perrimon. Quantitative morphological signatures define local signaling networks regulating cell morphology. *science*, 316(5832):1753–1756, 2007.

[6] L Basini, G D'Angelo, M Gobbi, GC Sarti, and C Gostoli. A desalination process through sweeping gas membrane distillation. *Desalination*, 64:245–257, 1987.

[7] Barry D Bavister. A minichamber device for maintaining a constant carbon dioxide in air atmosphere during prolonged culture of cells on the stage of an inverted microscope. *In vitro cellular & developmental biology*, 24(8):759–763, 1988.

[8] Michael R Behrens, Haley C Fuller, Emily R Swist, Jingwen Wu, Md Islam, Zhicheng Long, Warren C Ruder, Robert Steward, et al. Open-source, 3D-printed peristaltic pumps for small volume point-of-care liquid handling. *Scientific reports*, 10(1):1–10, 2020.

[9] Wolfgang Beyer. Sequence tube maps. https://github.com/wolfib/sequenceTubeMap, 2016.

[10] Axion Biosystems. Maestro pro — Axion biosystems — axionbiosystems.com. https://www.axionbiosystems.com/products/systems/maestro-pro. [Accessed 10 Aug 2022].

[11] Maxwell Biosystems. High-Density MEA Technology - MaxWell Biosystems — mxwbio.com. https://www.mxwbio.com/technology/. [Accessed 10 Aug 2022].

[12] Maxwell Biosystems. MaxTwo Multiwell Microelectrode Array – MaxWell Biosystems — mxwbio.com. https://www.mxwbio.com/products/maxtwo-multiwell-microelectrode-array/. [Accessed 10 Aug 2022].

[13] Ljiljana Brankovic, Costas S. Iliopoulos, Ritu Kundu, Manal Mohamed, Solon P. Pissis, and Fatima Vayani. Linear-time superbubble identification algorithm for genome assembly. *Theoretical Computer Science*, 609, Part 2:374 – 383, 2016.

[14] Gary R Bright, Gregory W Fisher, Jadwiga Rogowska, and D Lansing Taylor. Fluorescence ratio imaging microscopy: temporal and spatial measurements of cytoplasmic pH. *The Journal of cell biology*, 104(4):1019–1033, 1987.

[15] Brian L Browning and Sharon R Browning. A unified approach to genotype imputation and haplotype-phase inference for large data sets of trios and unrelated individuals. *The American Journal of Human Genetics*, 84(2):210–223, 2009.

[16] Darren J Burgess. Spatial transcriptomics coming of age. *Nature Reviews Genetics*, 20(6):317–317, 2019.

[17] Fady T Charbel, William E Hoffman, Mukesh Misra, Kelly Hannigan, and James I Ausman. Cerebral interstitial tissue oxygen tension, pH, HCO_3, CO_2. *Surgical neurology*, 48(4):414–417, 1997.

[18] Ketki Chawla, Sebastian C Bürgel, Gregor W Schmidt, Hans-Michael Kaltenbach, Fabian Rudolf, Olivier Frey, and Andreas Hierlemann. Integrating impedance-based growth-rate monitoring into a microfluidic cell culture platform for live-cell microscopy. *Microsystems & nanoengineering*, 4(1):1–12, 2018.

[19] Soong-Won Cho, Dong-Ku Kang, Jae-Bum Choo, Andrew J Demllo, and Soo-Ik Chang. Recent advances in microfluidic technologies for biochemistry and molecular biology. *BMB reports*, 44(11):705–712, 2011.

[20] E Cimetta, M Flaibani, M Mella, E Serena, L Boldrin, P De Coppi, and N Elvassore. Enhancement of viability of muscle precursor cells on 3D scaffold in a perfusion bioreactor. *The International journal of artificial organs*, 30(5):415–428, 2007.

[21] Martin J Cline. Gas permeable culture flask and method for culturing mammalian cells, September 1991. US5047347.

[22] International Electrotechnical Commission. Universal serial bus interfaces for data and power - part 3-1: Universal serial bus 3.1 specification, 2017. IEC 62680-3-1:2017.

[23] 1000 Genomes Project Consortium et al. A global reference for human genetic variation. *Nature*, 526(7571):68–74, 2015.

[24] Petr Danecek, Adam Auton, Goncalo Abecasis, Cornelis A Albers, Eric Banks, Mark A DePristo, Robert E Handsaker, Gerton Lunter, Gabor T Marth, Stephen T Sherry, et al. The variant call format and VCFtools. *Bioinformatics*, 27(15):2156–2158, 2011.

[25] Olivier Delaneau, Jean-Francois Zagury, and Jonathan Marchini. Improved whole-chromosome phasing for disease and population genetic studies. *Nature methods*, 10(1):5, 2013.

[26] Peter Donnelly and Stephen Leslie. The coalescent and its descendants. *arXiv preprint arXiv:1006.1514*, 2010.

[27] R.J. Duffin. Topology of series-parallel networks. *Journal of Mathematical Analysis and Applications*, 10(2):303 – 318, 1965.

[28] S Dufresne, A Hewitt, and S Robitaille. Ozone sterilization: another option for healthcare in the 21st century. *American journal of infection control*, 32(3):E26–E27, 2004.

[29] Richard Durbin. Efficient haplotype matching and storage using the positional burrows–wheeler transform (PBWT). *Bioinformatics*, 30(9):1266–1272, 2014.

[30] Alessandro D'Ausilio. Arduino: A low-cost multipurpose lab equipment. *Behavior research methods*, 44(2):305–313, 2012.

[31] Peter Esser, Lenia Weitzmann, et al. Evaporation from cell culture plates. *Thermo Scientific*, 2011. TILSPNUNCBU02 0111.

[32] AK Evseev, SV Zhuravel, A Yu Alentiev, IV Goroncharovskaya, and SS Petrikov. Membranes in extracorporeal blood oxygenation technology. *Membranes and Membrane Technologies*, 1(4):201–211, 2019.

[33] Guiqian Fang, Hao Wang, Zhancun Bian, Jie Sun, Aiqin Liu, Hao Fang, Bo Liu, Qingqiang Yao, and Zhongyu Wu. Recent development of boronic acid-based fluorescent sensors. *RSC advances*, 8(51):29400–29427, 2018.

[34] USB Implementers Forum. Universal serial bus device class definition for video devices, revision 1.5, 2012.

[35] Erik Garrison, Jouni Sirén, Adam M Novak, Glenn Hickey, Jordan M Eizenga, Eric T Dawson, William Jones, Shilpa Garg, Charles Markello, Michael F Lin, et al. Variation graph toolkit improves read mapping by representing genetic variation in the reference. *Nature biotechnology*, 36(9):875–879, 2018.

[36] Alex S Genshaft, Carly GK Ziegler, Constantine N Tzouanas, Benjamin E Mead, Alex M Jaeger, Andrew W Navia, Ryan P King, Miyeko D Mana, Siyi Huang, Vanessa Mitsialis, et al. Live cell tagging tracking and isolation for spatial transcriptomics using photoactivatable cell dyes. *Nature communications*, 12(1):1–15, 2021.

[37] Troy A. Giambernardi and Robert J. Klebe. Gas permeable pouches for growing cultures, April 1998. US5736398.

[38] Presens Precision Sensing GmbH. Product: OxoDishAE; OD24 — presens.de. `https://www.presens.de/products/detail/oxodish-od24`. [Accessed 10 Aug 2022].

[39] Elizabeth Godynyuk, Maya N Bluitt, Jessica R Tooley, Alexxai V Kravitz, and Meaghan C Creed. An open-source, automated home-cage sipper device for monitoring liquid ingestive behavior in rodents. *Eneuro*, 6(5), 2019.

[40] Andrew Gordon, Alejandro Colman-Lerner, Tina E Chin, Kirsten R Benjamin, Richard C Yu, and Roger Brent. Single-cell quantification of molecules and rates using open-source microscope-based cytometry. *Nature methods*, 4(2):175–181, 2007.

[41] Makiko Goto, Kiichi Sato, Atsushi Murakami, Manabu Tokeshi, and Takehiko Kitamori. Development of a microchip-based bioassay system using cultured cells. *Analytical chemistry*, 77(7):2125–2131, 2005.

[42] Christine Grienberger and Arthur Konnerth. Imaging calcium in neurons. *Neuron*, 73(5):862–885, 2012.

[43] Ihsan A Haddad and Alvin R Arsenault. Cell culture assembly, April 1976. US3948732.

[44] L Lee Hamm, Nazih Nakhoul, and Kathleen S Hering-Smith. Acid-base homeostasis. *Clinical Journal of the American Society of Nephrology*, 10(12):2232–2242, 2015.

[45] U Hassan, NN Watkins, C Edwards, and R Bashir. Flow metering characterization within an electrical cell counting microfluidic device. *Lab on a Chip*, 14(8):1469–1476, 2014.

[46] Mark R Hatch. *The maker revolution: Building a future on creativity and innovation in an exponential world*. John Wiley & Sons, 2017.

[47] Gabriel Helmlinger, Fan Yuan, Marc Dellian, and Rakesh K Jain. Interstitial ph and po2 gradients in solid tumors in vivo: high-resolution measurements reveal a lack of correlation. *Nature medicine*, 3(2):177–182, 1997.

[48] Yun Seok Heo, Lourdes M Cabrera, Jonathan W Song, Nobuyuki Futai, Yi-Chung Tung, Gary D Smith, and Shuichi Takayama. Characterization and resolution of evaporation-mediated osmolality shifts that constrain microfluidic cell culture in poly(dimethylsiloxane) devices. *Analytical chemistry*, 79(3):1126–1134, 2007.

[49] Chung-Liang Ho, Tun-Yi Mou, Pei-Shuan Chiang, Chu-Li Weng, and Nan-Haw Chow. Mini chamber system for long-term maintenance and observation of cultured cells. *Biotechniques*, 38(2):267–273, 2005.

[50] Moritz Hofer and Matthias P Lutolf. Engineering organoids. *Nature Reviews Materials*, 6(5):402–420, 2021.

[51] innoME GmbH. Live cell imaging – zenCELL owl – Microscopy for the incubator – zencellowl.com. `https://zencellowl.com/live-cell-imaging/`. [Accessed 10 Aug 2022].

[52] Cruzio Internet. Cruzio enterprise internet: Business-grade connections. `https://cruzio.com/services/broadband/enterprise/`, Sep 2020. [Accessed August 10 2022].

[53] Mohammad Manzurul Islam, Sarwar Morshed, and Parijat Goswami. Cloud computing: A survey on its limitations and potential solutions. *International Journal of Computer Science Issues (IJCSI)*, 10(4):159, 2013.

[54] Sasan Jalili-Firoozinezhad, Francesca S Gazzaniga, Elizabeth L Calamari, Diogo M Camacho, Cicely W Fadel, Amir Bein, Ben Swenor, Bret Nestor, Michael J Cronce, Alessio Tovaglieri, et al. A complex human gut microbiome cultured in an anaerobic intestine-on-a-chip. *Nature biomedical engineering*, 3(7):520–531, 2019.

[55] R Daniel Johnson and Leonidas G Bachas. Ionophore-based ion-selective potentiometric and optical sensors. *Analytical and bioanalytical chemistry*, 376(3):328–341, 2003.

[56] Jolle W Jolles. Broad-scale applications of the raspberry pi: A review and guide for biologists. *Methods in Ecology and Evolution*, 12(9):1562–1579, 2021.

[57] Antreas Kantaros, Olaf Diegel, Dimitrios Piromalis, Georgios Tsaramirsis, Alaa Omar Khadidos, Adil Omar Khadidos, Fazal Qudus Khan, and Sadeeq Jan. 3d printing: Making an innovative technology widely accessible through makerspaces and outsourced services. *Materials Today: Proceedings*, 49:2712–2723, 2022.

[58] Alon Keinan and Andrew G Clark. Recent explosive human population growth has resulted in an excess of rare genetic variants. *science*, 336(6082):740–743, 2012.

[59] Stacy M Kenyon, Michelle M Meighan, and Mark A Hayes. Recent developments in electrophoretic separations on microfluidic devices. *Electrophoresis*, 32(5):482–493, 2011.

[60] John Frank Charles Kingman. The coalescent. *Stochastic processes and their applications*, 13(3):235–248, 1982.

[61] Theodor Kolobow and Robert L Bowman. Construction and evaluation of an alveolar membrane artificial heart-lung. *ASAIO Journal*, 9(1):238–243, 1963.

[62] Martin U Kopp, Andrew J de Mello, and Andreas Manz. Chemical amplification: continuous-flow PCR on a chip. *Science*, 280(5366):1046–1048, 1998.

[63] Klaus Koren and Silvia E Zieger. Optode based chemical imaging—possibilities, challenges, and new avenues in multidimensional optical sensing. *ACS sensors*, 6(5):1671–1680, 2021.

[64] Matteo Laganà and Manuela T Raimondi. A miniaturized, optically accessible bioreactor for systematic 3D tissue engineering research. *Biomedical microdevices*, 14(1):225–234, 2012.

[65] Frederic Leens. An introduction to i2c and spi protocols. ieee instrumentation measurement magazine 12, no. 1 (february 2009): 8-13, 2009.

[66] Na Li and Matthew Stephens. Modeling linkage disequilibrium and identifying recombination hotspots using single-nucleotide polymorphism data. *Genetics*, 165(4):2213–2233, 2003.

[67] Yun Li, Cristen J Willer, Jun Ding, Paul Scheet, and Gonçalo R Abecasis. Mach: using sequence and genotype data to estimate haplotypes and unobserved genotypes. *Genetic epidemiology*, 34(8):816–834, 2010.

[68] John Liner. Multi-dish laboratory unit, August 1971. US3597326.

[69] Po-Ru Loh, Petr Danecek, Pier Francesco Palamara, Christian Fuchsberger, Yakir A Reshef, Hilary K Finucane, Sebastian Schoenherr, Lukas Forer, Shane McCarthy, Goncalo R Abecasis, et al. Reference-based phasing using the haplotype reference consortium panel. *Nature genetics*, 48(11):1443, 2016.

[70] Leonardo C Lopez, Garth L Wilkes, Phil M Stricklen, and Scott A White. Synthesis, structure, and properties of poly(4-methyl-1-pentene). *Journal of Macromolecular Science, Part C: Polymer Reviews*, 32(3-4):301–406, 1992.

[71] GS Luka, E Nowak, J Kawchuk, M Hoorfar, and H Najjaran. Portable device for the detection of colorimetric assays. *Royal Society open science*, 4(11):171025, 2017.

[72] Gerton Lunter. Fast haplotype matching in very large cohorts using the Li and Stephens model. *bioRxiv*, 2016.

[73] Victoria T Ly, Pierre V Baudin, Pattawong Pansodtee, Erik A Jung, Kateryna Voitiuk, Yohei M Rosen, Helen Rankin Willsey, Gary L Mantalas, Spencer T Seiler, John A Selberg, et al. Picroscope: low-cost system for simultaneous longitudinal biological imaging. *Communications biology*, 4(1):1–11, 2021.

[74] Haiyang Ma, Jia Hao, and Roger Zimmermann. Access point centric scheduling for dash streaming in multirate 802.11 wireless network. In *2014 IEEE International Conference on Multimedia and Expo (ICME)*, pages 1–6. IEEE, 2014.

[75] André Maia Chagas. Haves and have nots must find a better way: The case for open scientific hardware. *PLoS biology*, 16(9):e3000014, 2018.

[76] M Malpighi. De pulmonibus. *Philosophical Transactions of the Royal Society*, 1661.

[77] Gregory R. Martin and Allison J. Tanner. Multilayered cell culture apparatus, June 2010. US7745209.

[78] G Marx. An apparatus for transferring a gas between two liquids, December 1974. US3856475.

[79] Uwe Maskos, Karima Kissa, Cécile St. Cloment, and Philippe Brûlet. Retrograde trans-synaptic transfer of green fluorescent protein allows the genetic mapping of neuronal circuits in transgenic mice. *Proceedings of the National Academy of Sciences*, 99(15):10120–10125, 2002.

[80] David Mayhew and Venkata Krishnan. Pci express and advanced switching: evolutionary path to building next generation interconnects. In *11th Symposium on High Performance Interconnects, 2003. Proceedings.*, pages 21–29. IEEE, 2003.

[81] Turney J McKee and Svetlana V Komarova. Is it time to reinvent basic cell culture medium?, 2017.

[82] Paul Medvedev and Michael Brudno. Maximum likelihood genome assembly. *Journal of computational Biology*, 16(8):1101–1116, 2009.

[83] Michael L Metzker. Sequencing technologies—the next generation. *Nature reviews genetics*, 11(1):31–46, 2010.

[84] WW Minuth, S Kloth, J Aigner, M Sittinger, and W Röckl. Approach to an organotypical environment for cultured cells and tissues. *Biotechniques*, 20(3):498–501, 1996.

[85] Thomas A Moore, Peter Brodersen, and Edmond WK Young. Multiple myeloma cell drug responses differ in thermoplastic vs PDMS microfluidic devices. *Analytical chemistry*, 89(21):11391–11398, 2017.

[86] Ali Mortazavi, Brian A Williams, Kenneth McCue, Lorian Schaeffer, and Barbara Wold. Mapping and quantifying mammalian transcriptomes by rna-seq. *Nature Methods*, 5(7):621–628, 2008.

[87] Jan Müller, Marco Ballini, Paolo Livi, Yihui Chen, Milos Radivojevic, Amir Shadmani, Vijay Viswam, Ian L Jones, Michele Fiscella, Roland Diggelmann, et al. High-resolution CMOS MEA platform to study neurons at subcellular, cellular, and network levels. *Lab on a Chip*, 15(13):2767–2780, 2015.

[88] Alphonsus HC Ng, Uvaraj Uddayasankar, and Aaron R Wheeler. Immunoassays in microfluidic systems. *Analytical and bioanalytical chemistry*, 397(3):991–1007, 2010.

[89] Adam M Novak, Erik Garrison, and Benedict Paten. A graph extension of the positional burrows–wheeler transform and its applications. *Algorithms for Molecular Biology*, 12(1):1–12, 2017.

[90] Adam M Novak, Glenn Hickey, Erik Garrison, Sean Blum, Abram Connelly, Alexander Dilthey, Jordan Eizenga, M. A. Saleh Elmohamed, Sally Guthrie, André Kahles, Stephen Keenan, Jerome Kelleher, Deniz Kural, Heng Li, Michael F Lin, Karen Miga, Nancy Ouyang, Goran Rakocevic, Maciek Smuga-Otto, Alexander Wait Zaranek, Richard Durbin, Gil McVean, David Haussler, and Benedict Paten. Genome graphs. *bioRxiv*, 2017.

[91] Jared O'Connell, Kevin Sharp, Nick Shrine, Louise Wain, Ian Hall, Martin Tobin, Jean-Francois Zagury, Olivier Delaneau, and Jonathan Marchini. Haplotype estimation for biobank-scale data sets. *Nature genetics*, 48(7):817, 2016.

[92] Omnexus. HDT @0.46 MPa (67 psi). https://omnexus.specialchem.com/polymer-properties/properties/hdt-0-46-mpa-67-psi?src=prop-cnx.

[93] OmniVision Technologies, Inc. *OV5640 color CMOS QSXGA (5 megapixel) image sensor with OmniBSI technology*, 2011. Rev. 2.03.

[94] Taku Onodera, Kunihiko Sadakane, and Tetsuo Shibuya. *Detecting Superbubbles in Assembly Graphs*, pages 338–348. Springer Berlin Heidelberg, Berlin, Heidelberg, 2013.

[95] N Opitz and Dietrich W Luebbers. Theory and development of fluorescence-based optochemical oxygen sensors: oxygen optodes. *International anesthesiology clinics*, 25(3):177–197, 1987.

[96] T Panneerselvam, S Raghuraman, and N Vamsi Krishnan. Investigating mechanical properties of 3d-printed polyethylene terephthalate glycol material under fused deposition modeling. *Journal of The Institution of Engineers (India): Series C*, 102(2):375–387, 2021.

[97] Benedict Paten, Adam Novak, and David Haussler. Mapping to a reference genome structure. *arXiv preprint arXiv:1404.5010*, 2014.

[98] Benedict Paten, Adam M Novak, Erik Garrison, and Glenn Hickey. Superbubbles, ultrabubbles and cacti. *bioRxiv*, 2017.

[99] J. Thomas Patterson. A microscope-stage incubator. *The Biological Bulletin*, 14(6):324–327, 1908.

[100] Hanchuan Peng. Bioimage informatics: a new area of engineering biology. *Bioinformatics*, 24(17):1827–1836, 2008.

[101] Howard Latimer Penman. Natural evaporation from open water, bare soil and grass. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 193(1032):120–145, 1948.

[102] Daniel Perlman. Cell culture plate with oxygen and carbon dioxide-permeable waterproof sealing membrane, January 1999. US5858770.

[103] Ingo Pinnau and Lora G Toy. Gas and vapor transport properties of amorphous perfluorinated copolymer membranes based on 2,2-bistrifluoromethyl-4,5-difluoro-1,3-dioxole/tetrafluoroethylene. *Journal of Membrane Science*, 109(1):125–133, 1996.

[104] Steve M Potter and Thomas B DeMarse. A new approach to neural cell culture for long-term studies. *Journal of neuroscience methods*, 110(1-2):17–24, 2001.

[105] Keil J Regehr, Maribella Domenech, Justin T Koepsel, Kristopher C Carver, Stephanie J Ellison-Zelski, William L Murphy, Linda A Schuler, Elaine T Alarid, and David J Beebe. Biological implications of polydimethylsiloxane-based microfluidic cell culture. *Lab on a Chip*, 9(15):2132–2139, 2009.

[106] Valentin Romanov, Raheel Samuel, Marzieh Chaharlang, Alexander R Jafek, Adam Frost, and Bruce K Gale. Fdm 3d printing of high-pressure, heat-resistant, transparent microfluidic devices. *Analytical chemistry*, 90(17):10450–10456, 2018.

[107] Yohei Rosen, Jordan Eizenga, and Benedict Paten. Describing the local structure of sequence graphs. In *International Conference on Algorithms for Computational Biology*, pages 24–46. Springer, 2017.

[108] Yohei Rosen, Jordan Eizenga, and Benedict Paten. Modelling haplotypes with respect to reference cohort variation graphs. *Bioinformatics*, 33(14):i118–i123, 2017.

[109] Yohei M Rosen and Benedict J Paten. An average-case sublinear forward algorithm for the haploid li and stephens model. *Algorithms for Molecular Biology*, 14(1):1–12, 2019.

[110] Nassim Rousset, Rubén López Sandoval, Mario Matteo Modena, Andreas Hierlemann, and Patrick M Misun. Modeling and measuring glucose diffusion and consumption by colorectal cancer spheroids in hanging drops using integrated biosensors. *Microsystems & nanoengineering*, 8(1):1–18, 2022.

[111] Dirk Saalfrank, Anil Krishna Konduri, Shahrzad Latifi, Rouhollah Habibey, Asiyeh Golabchi, Aurel Vasile Martiniuc, Alois Knoll, Sven Ingebrandt, and Axel Blau. Incubator-independent cell-culture perfusion platform for continuous long-term microelectrode array electrophysiology and time-lapse imaging. *Royal Society open science*, 2(6):150031, 2015.

[112] David Sabourin, Peder Skafte-Pedersen, Martin Jensen Søe, Mette Hemmingsen, Massimo Alberti, Vasile Coman, Jesper Petersen, Jenny Emnéus, Jörg P Kutter, Detlef Snakenborg, et al. The main stream component platform: a holistic approach to microfluidic system design. *Journal of Laboratory Automation*, 18(3):212–228, 2013.

[113] Fash Safdari and Anatoliy Gorbenko. Experimental evaluation of performance anomaly in mixed data rate ieee802. 11ac wireless networks. In *2019 10th International Conference on Dependable Systems, Services and Technologies (DESSERT)*, pages 82–87. IEEE, 2019.

[114] Lucid Scientific. Lucid Scientific — lucidsci.com. `https://lucidsci.com/#resipher`. [Accessed 10 Aug 2022].

[115] Thermo Scientific. Improving assay consistency using Nunc Edge 2.0 96-well plates. *Thermo Scientific*, 2011. COL22027 0917.

[116] Scoperta Life Sciences LLC. Celloger Mini – Scoperta life sciences, LLC. `https://scopertals.com/product/celloger-mini/`. [Accessed 10 Aug 2022].

[117] Violetta Sessi, Bergoi Ibarlucea, Florent Seichepine, Stephanie Klinghammer, Imad Ibrahim, André Heinzig, Nadine Szabo, Thomas Mikolajick, Andreas Hierlemann, Urs Frey, et al. Multisite dopamine sensing with femtomolar resolution using a CMOS enabled aptasensor chip. *Frontiers in neuroscience*, 16, 2022.

[118] Takatsugu Shimono, Yu Shomura, Iwao Hioki, Akira Shimamoto, Hironori Tenpaku, Yasumi Maze, Koji Onoda, Motoshi Takao, Hideto Shimpo, and Isao Yada. Silicone-coated polypropylene hollow-fiber oxygenator: experimental evaluation and preliminary clinical use. *The Annals of thoracic surgery*, 63(6):1730–1736, 1997.

[119] Xiaojing Su, Edmond WK Young, Heather AS Underkofler, Timothy J Kamp, Craig T January, and David J Beebe. Microfluidic cell culture and its application in high-throughput drug screening: cardiotoxicity assay for hERG channels. *Journal of biomolecular screening*, 16(1):101–111, 2011.

[120] Peter H Sudmant, Tobias Rausch, Eugene J Gardner, Robert E Handsaker, Alexej Abyzov, John Huddleston, Yan Zhang, Kai Ye, Goo Jun, Markus Hsi-Yang Fritz, et al. An integrated map of structural variation in 2,504 human genomes. *Nature*, 526(7571):75–81, 2015.

[121] Kozo Suma, Takayuki Tsuji, Yasuo Takeuchi, Kenji Inoue, Kenji Shiroma, Tetsuo Yoshikawa, and Jun Narumi. Clinical performance of microporous polypropylene hollow-fiber oxygenator. *The Annals of thoracic surgery*, 32(6):558–562, 1981.

[122] Wing-Kin Sung, Kunihiko Sadakane, Tetsuo Shibuya, Abha Belorkar, and Iana Pyrogova. An O(M log M)-time algorithm for detecting superbubbles. *IEEE/ACM Trans. Comput. Biol. Bioinformatics*, 12(4):770–777, July 2015.

[123] Jason R Swedlow, Ilya G Goldberg, Kevin W Eliceiri, OME consortium, et al. Bioimage informatics for experimental biology. *Annual review of biophysics*, 38:327, 2009.

[124] Yutaka Tahara and Katsuya Obara. A novel shell-less culture system for chick embryos using a plastic film as culture vessels. *The Journal of Poultry Science*, page 0130043, 2014.

[125] Atsushi Takano, Masato Tanaka, and Nobuyuki Futai. On-chip CO_2 incubation for pocket-sized microfluidic cell culture. *Microfluidics and nanofluidics*, 12(6):907–915, 2012.

[126] Atsushi Takano, Masato Tanaka, and Nobuyuki Futai. On-chip multi-gas incubation for microfluidic cell cultures under hypoxia. *Biomicrofluidics*, 8(6):061101, 2014.

[127] Masahiro Tanaka, Raynard Bateman, Daniel Rauh, Eugeni Vaisberg, Shyam Ramachandani, Chao Zhang, Kirk C Hansen, Alma L Burlingame, Jay K Trautman, Kevan M Shokat, et al. An unbiased cell morphology–based screen for new, biologically active small molecules. *PLoS biology*, 3(5):e128, 2005.

[128] Mansour Tayebi-khorami, Nahid Chegeni, Maryam Tahmasebi Birgani, Amir Danyaei, Reza Fardid, Jaber Zafari, et al. Construction a co2 incubator for cell culture with capability of transmitting microwave radiation. *Journal of Medical Signals & Sensors*, 12(2):127, 2022.

[129] John M Toomasian, Robert J Schreiner, David E Meyer, Monica E Schmidt, Sarah E Hagan, Grant W Griffith, Robert H Bartlett, and Keith E Cook. A polymethylpentene fiber gas exchanger for long-term extracorporeal life support. *ASAIO journal*, 51(4):390–397, 2005.

[130] Liam Tung. Raspberry pi: Why they are so hard to buy right now, and what you can do about it. https://www.zdnet.com/article/raspberry-pi-why-they-are-so-hard-to-buy-right-now-and-what-you-can-do-about-it/. [Accessed August 10 2022].

[131] Jacobo Valdes, Robert E. Tarjan, and Eugene L. Lawler. The recognition of series parallel digraphs. *SIAM Journal on Computing*, 11(2):298–313, 1982.

[132] Andreas Weltin, Kinga Slotwinski, Jochen Kieninger, Isabella Moser, Gerhard Jobst, Marcus Wego, Ralf Ehret, and Gerald A Urban. Cell culture monitoring for drug screening and cancer research: a transparent, microfluidic, multi-sensor microsystem. *Lab on a Chip*, 14(1):138–146, 2014.

[133] Sachini Wickramasinghe, Truong Do, and Phuong Tran. Fdm-based 3d printing of polymer and associated composite: A review on mechanical properties, defects and treatments. *Polymers*, 12(7):1529, 2020.

[134] Vincent Wiegmann, Cristina Bernal Martinez, and Frank Baganz. A simple method to determine evaporation and compensate for liquid losses in small-scale cell culture systems. *Biotechnology letters*, 40(7):1029–1036, 2018.

[135] Brian T. Wilhelm and Josette-Renée Landry. RNA-seq—quantitative measurement of expression through massively parallel RNA-sequencing. *Methods*, 48(3):249–257, 2009.

[136] Amy L Williams, Nick Patterson, Joseph Glessner, Hakon Hakonarson, and David Reich. Phasing of many thousands of genotyped samples. *The American Journal of Human Genetics*, 91(2):238–251, 2012.

[137] David F Wilson, William MF Lee, Sosina Makonnen, Olga Finikova, Sofia Apreleva, and Sergei A Vinogradov. Oxygen pressures in the interstitial space and their relationship to those in the blood plasma in resting skeletal muscle. *Journal of applied physiology*, 101(6):1648–1656, 2006.

[138] John R. Wilson. Gas permeable cell culture device and method of use, August 2013. US8518692.

[139] John R. Wilson. Highly efficient gas permeable devices and methods for culturing cells, August 2014. US8809044.

[140] AAPM Wimalarathne. A study of whether extension tubes can substitute for macro photography. 2016.

[141] Yongxian Xu, Peng Zou, and Adam E Cohen. Voltage imaging with genetically encoded indicators. *Current opinion in chemical biology*, 39:1–10, 2017.

[142] Helene Zirath, Mario Rothbauer, Sarah Spitz, Barbara Bachmann, Christian Jordan, Bernhard Müller, Josef Ehgartner, Eleni Priglinger, Severin Mühleder, Heinz Redl, et al. Every breath you take: non-invasive real-time oxygen biosensing in two-and three-dimensional microfluidic cell models. *Frontiers in physiology*, 9:815, 2018.

[143] Kinga Zor, Arto Heiskanen, Claudia Caviglia, M Vergani, E Landini, Fozia Shah, Marco Carminati, A Martínez-Serrano, T Ramos Moreno, Merab Kokaia, et al. A compact multifunctional microfluidic platform for exploring cellular dynamics in real-time using electrochemical detection. *Rsc Advances*, 4(109):63761–63771, 2014.