An Ultra-high Throughput Mutational Spectrometer for Human Genetic Diagnostics

by

Craig Richard Forest

B.S.M.E., Georgia Institute of Technology (2001)
S.M.M.E., Massachusetts Institute of Technology (2003)

Submitted to the Department of Mechanical Engineering
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2007

Author .....
Department of Mechanical Engineering
May 20, 2007

Certified by .....
Ian W. Hunter
Hatsopoulos Professor of Mechanical Engineering
Thesis Supervisor

Accepted by .....
Lallit Anand
Chairman, Department Committee on Graduate Students

An Ultra-high Throughput Mutational Spectrometer for Human Genetic Diagnostics

by

Craig Richard Forest

Submitted to the Department of Mechanical Engineering
on May 20, 2007, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

Discovering the genetic causes of common diseases may require scanning for mutations in all of the genes in a million people, a significant undertaking. Such discoveries would revolutionize biotechnology, potentially enabling simple genetic tests for risk and targeted preventative or therapeutic strategies. An increase in throughput of genetic analysis instrumentation by several orders of magnitude is essential to undertake such an ambitious task.

In this thesis, progress will be presented towards the creation of such a "mutational spectrometer" instrument containing up to 10,000 capillary channels and enabled with subsystems for loading, separating, and detecting fluorescently-labeled DNA. Challenges include DNA manipulation, optical signal detection, macro/micro design integration, precision alignment and assembly, and thermal control. To manipulate DNA, we have utilized a bioMEMS design platform for interfacing to an array of separation channels that enables electrokinetic biomolecule loading, detection, and fraction collection in independent wells. Signal detection is accomplished by a sensitive (10^7 molecule limit-of-detection), scalable (to 10,000 independent channels), end-of-column fluorescence detection technology that accommodates tightly packed capillary arrays as required for ultra-high throughput electrophoretic separation. Capillary array assembly and constraint technologies have been developed for 2-D arrays containing as many as 10,000 replaceable capillaries. Thermal control requirements of 0.3 °C over the entire 10,000 channel array are met with a cross-flow water heat exchanger. Additional subsystems for forcing a viscous polymer matrix into the capillaries, and interfacing the capillary array to a fluid reservoir for electrophoresis have also been developed, as required. This work lays the foundation for the realization of a mutational spectrometer instrument that will enable population-wide pangenomic analyses to uncover the genetic causes of common diseases.

Thesis Supervisor: Ian W. Hunter
Title: Hatsopoulos Professor of Mechanical Engineering

Acknowledgments

From helpful suggestions by colleagues to heart-felt talks with friends and family, I owe a debt of gratitude to many for their kind assistance. Prof. Ian Hunter has been exceedingly generous and supportive–giving me the freedom to pursue any idea. The lab's facilities and equipment are unparalleled, and I consider myself extremely fortunate to have had the opportunity to study there. Thank you Ian for this opportunity. There have been many lab members that have offered personal and professional advice. Dr. Cathy Hogan taught me all of the biology that I know, and I'm deeply grateful, although I may have not always shown it. For all of the design, manufacturing, and analysis suggestions and discussion I owe much to my lab colleagues including Dr. Bryan Crane, Tim Fofonoff, Nate Ball, Bryan Schmid, Dr. Mike Garcia-Webb, Brian Hemond, Dr. Serge Lafontaine, Dr. Andrew Taberner, Brian Ruddy, Dr. Robert David, Ariel Hermann, Dr. Hsin-Ni Ho, Nasko Pavlov, Priam Pillai, Dr. Rachel Pytel, Nic Sabourin, Dr. Nate Vandesteeg, Dawn Wendell, Dr. Patrick Anquetil, and Dr. James Tangorra. Kate Melvin is also always so willing to lend a helping hand.

My committee members Profs. Bill Thilly and Peter So have contributed in fundamental ways to the direction of the project. Bill was kind enough to share his vision for DNA mutational spectrometry with me, and invite me onboard. Peter So's optics expertise helped me overcome the most challenging aspects of the project. Additionally, both individuals' personal advice, whether for job searching or proposal writing, has made all the difference.

Without a team of eager, intelligent, and independent undergraduate research assistants, this project would not have been possible. Thank you to Miguel Saez, Michael Beltran, Zachary Reynolds, Cynthia Lin, Maria Telleria, Lohith Kini, Nathan Ball, James Lin, Timothy Suen, Ian Collier, Darragh Buckley, Bryan Woodruff, Edward Hsieh, and Mike Del Zio. I would like to recognize the specific contributions to this thesis by these undergraduates, under my guidance and supervision. Their works and/or theses on the subject are are the following:

1. James J. Lin, Senior thesis student, MIT, Jan 2005-Jun 2005, "Manufacturing

Improvement and Thermal Property Characterization of the Frame Structure of an Ultra-High Throughput Mutational Spectrometer."

- Designed and improved manufacturing of the heat exchanger and its interfaces.

2. Timothy W. Suen, Senior thesis student, MIT, Jan 2005-Jun 2005, "Temperature Response of the Ultra-High Throughput Mutational Spectrometer."

   - Modeled the heat exchanger and conducted measurements to evaluate models.

3. Michael Del Zio Senior thesis student, MIT, Jan 2004-Jun 2004, "Design and Testing of the Thermal Properties of the Structure of an Ultra High-Throughput Mutational Spectrometer."

   - Manufactured the heat exchanger.

4. Nathan B. Ball, Senior thesis student, MIT, Jan 2004-Jun 2005, "Design and Characterization of a Gel Loading Mechanism for an Ultra-High Throughput Mutational Spectrometer."

   - Designed and manufactured LED array, UTMS frame, beam stop, and various other components of the system.
   - Designed and manufactured sub-10,000 capillary gel loaders and measured gel behavior in capillaries (e.g., loading time).

5. Miguel A. Saez, Senior thesis student, MIT, Jun 2006-Jun 2007, "Microforging technique for rapid, low-cost manufacture of lens array molds and its application in a biomedical instrument."

   - Performed numerous experiments to converge on optimal micro-forging process, as well as process modeling.

6. Michael J. Beltran, Senior thesis student, MIT, June 2006-Jun 2007, "One-hundred channel electrophoresis prototype for an ultra-high throughput mutational spectrometer."

- Designed and manufactured the 100-port micro-fluidic connector and versions of the buffer reservoirs for 100 capillary electrophoresis.

7. Lohith Kini, Undergraduate research student, MIT, Feb 2006-Jun 2006.

   - Investigated lenslet array manufacturing options.

8. Edward F. Hsieh, Undergraduate research student, MIT, Jul 2004-Dec 2004.

   - Measured physical properties of linear poly(acrylamide) gel (e.g., viscosity).

9. Ian Collier, Undergraduate research student, MIT, Jan 2005-Jun 2005.

   - Designed and manufactured prototype capillary array assembly instruments, with additional modeling of the assembly process.

10. Darragh Buckley, Undergraduate research student, MIT, Jun 2005-Sep 2005.

    - Designed and tested capillary array assembly techniques.

11. Bryan Woodruff, Undergraduate research student, MIT, Oct 2004-Jun 2005.

    - Designed and tested of capillary constraint devices.

12. Maria Telleria, Undergraduate research student, MIT, Feb 2006-Sep 2006.

    - Devised and built capillary array assembly and constraint support hardware.

13. Cynthia Lin, Undergraduate research student, MIT, Feb 2006-Jun 2006.

    - Designed and constructed capillary array support structures.

14. Zachary Reynolds, Undergraduate research student, MIT, Jan 2006-Jun 2006.

    - Assisted with 10,000 capillary array assembly.

I would very much like to thank my collaborators as well, including Dr. Xiao-Cheng Li-Sucholeiki, Wei-ming Zheng, Dr. Per Ekstrom, Karen-Marie Heintz, Leslie Kao, Joe Macomber, Dr. Amy Herr, and Dr. Ben Cook.

I have received support from so many friends and colleagues, whose valuable advice has altered my course. In particular, I am deeply indebted to Prof. John Hart, Alec Robertson, Dr. Mark Schattenburg, Prof. Alex Slocum, Dr. Eric Wade, Dr. Anton Thomas, Joseph Walish, Prof. Bill Singhose, Dr. Dan Ehrlich, Prof. Steven Soper, Prof. Richard Mathies, and Prof. David Gordon Wilson.

I could not have been more blessed with a supportive and loving family environment. My wife Laura, my parents Richard and Esther, and in-laws Don and April Major, have been there throughout. Thank you for your continual encouragement and support.

For funding, I would like to gratefully acknowledge the generous support of the National Science Foundation, Sandia National Laboratories, the MIT BioInstrumentation Laboratory, the MIT Graduate Student Council, and the MIT Biological Engineering Division.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation

The discovery of mutations in our DNA that confer risk for common diseases is one of the most important challenges in biotechnology research today. Imagine the revolution in human genetics that would ensue were such mutations known. Individuals could be cheaply and rapidly screened (e.g., [1]), and their genetic profiles could be compared with a database of markers for risk. If an individual's DNA contained mutations known to be associated with a disease, personalized medicine could be utilized from birth. The current suite of tests used during in-vitro fertilization, few of which are genetic, could be greatly expanded, providing prospective parents with invaluable information. Such discoveries would also permit the development of new drugs which target specific genetic pathways.

Clearly, the discovery of gene-disease associations for common diseases would revolutionize biotechnology. Is this a good thing? What is required to make such discoveries? More specifically, how long will it take, how much will it cost, and what instrumentation and technical approach are required? These are critical questions, and there is strong disagreement among the experts. Billions of public and private U.S. dollars have already been spent in this search (e.g., [2, 3, 4]) and, with a few exceptions that will be discussed, the genetic causes for common diseases remain unknown. Therefore, a great challenge and opportunity awaits biotechnology researchers

and the public in this century of genetic research.

This thesis focuses mainly on an instrumentation platform, termed an ultra-high throughput mutational spectrometer (UTMS), which could be a useful tool in searching through the vast amount of genetic information in the human population to discover gene-disease associations. The name UTMS is apt because of the instrument's required ability to sort through massive amounts of genetic information ("Ultra-high Throughput") to measure the frequencies of mutations in a populations of people, their "Mutational Spectrum."

Before describing the instrument in subsequent Chapters, we focus here on more fundamental questions concerning the nature of human mutations and their associations with diseases, strategies and technologies to apprehend these associations, costs, and time, in light of what is known about genetics today. We lastly introduce and justify the design for the UTMS.

1.2 Ethical implications of gene-disease associations

The ethical dilemmas inherent to a priori knowledge of genetic risk for disease are now being discussed in public forums [5]. What effects would such knowledge have on our society? Compiling a database of genetic markers that confer risk of common diseases does not imply that everyone must know their predisposition. But genetic screening for those who do want to know will be possible, cheap, and rapid–with today's current genotyping technologies.

For rare diseases for which the gene cause is known, such as Huntington's disease, an incurable brain disorder that can ravage the mind for decades, most young adults who know the disease runs in their family have avoided the DNA test that can tell whether they will get it, preferring the torture, and hope, of not knowing. On the other hand, those who have chosen to be tested say it will help them decide how to live their lives [5].

However, the choice may not be altogether personal. Life insurance companies, HMO's, even employers would have financial interest in such knowledge. Could individuals with or without the knowledge prevent these organizations from acquiring it? As with other bio-ethics issues (e.g., human embryonic stem cell research), governments would be called upon to establish guidelines that balance individual rights with societal protection.

Personal testing or not, knowing the genetic causes for common diseases, would certainly accelerate the search for cures, benefitting mankind as a whole. These cures could include genetic therapy, specifically modifying or inhibiting the causal genes, rather than their products–enzymes, proteins, RNA–as is commonly done today. In addition, individuals carrying risk-conferring genes could make lifestyle changes such as conception through in-vitro fertilization, wherein embryos or even gametes (sperm or eggs) could be pre-screened, enabling them to have healthy children. Most importantly, the ethics of personal screening should not impede the imperative search for the genetic causes of diseases.

1.3 What are diseases?

A disease is defined as a pathological condition resulting from various causes, either environmental (e.g., stresses, infections) or genetic defects, and characterized by an identifiable group of signs or symptoms [6]. Thus diseases can be caused by genetics or environment or combinations thereof. In this work, we focus exclusively on the genetic causes, under the important hypothesis that the diseases of interest have some genetic risk.

Diseases can be rare, which for this thesis we shall define as afflicting $<1\%$ of the global general population, or common, with prevalence of $\geq 1\%$. Some recognizable examples of these types of disease are shown in Table 1.1. Our definition is conservative, as others define "rare" as a prevalence of fewer than 5 per 10,000, or 0.05%, in the community [http://www.eurordis.org].

Rare genetic diseases are typically chronic and life-threatening. They affect indi-

Rare diseases 5000-8000 estimated	Common diseases >100 known
some cancers (e.g., kidney, testicular, endometrial, fallopian tube), Chrohn's disease, sickle cell anemia, Creutzfeldt-Jakob Disease (human mad-cow disease), aplastic anemia (bone marrow disease), Fabry disease, dwarfism, Tourette's syndrome, hemophilia A&B, Infantile myofibromatosis, leukemia, Duchenne muscular dystrophy, Amyotrophic lateral sclerosis (Lou Gehrigs disease)	most cancers (e.g., lung, stomach, colon, bladder, skin, breast, pancreatic), high blood pressure, hypertension, hyperlipidemia, coronary atherosclerosis, diabetes type II, Alzheimer's disease, Parkinson's disease, alcoholism, eating disorders, manic depression, schizophrenia, arthritis, asthma, alopecia (baldness), amblyopia (lazy eye)

Table 1.1: Sampling of rare and common human diseases of public importance that are known or suspected to have a genetic risk.

viduals' fecundity, or ability to reproduce. If they did not, they would be common. At least 2000 rare diseases have identified genetic origins [http://www.hgmd.org], and estimates for the number of rare diseases that have genetic origins range up to 80% of the 5000-8000 known rare diseases [http://www.eurordis.org]. Non-genetic rare diseases are the result of infections and allergies or due to degenerative and proliferative causes. Approximately 6-8% of the global population is afflicted with rare diseases [http://www.eurordis.org].

In contrast, common diseases generally do not affect fecundity. In the general population, lifetime risk for the sum of all cancers approaches 100%, for the sum of all cardiovascular diseases approaches 90%, and for diabetes approaches 6% [http://epidemiology.mit.edu]. Thus, although rare diseases are important to the individuals who have them, they are, by comparison, of limited significance in terms of public health. In contrast again to rare diseases, the genetic causes for common diseases are largely unknown, as will be presented.

1.4 Nature of human genetic mutations

A full discussion of population genetics is beyond the scope of this thesis, but in this Section we will strive to introduce some concepts pertinent to our discussion.

At the time of this writing, approximately 6.6×10^9 people inhabit the Earth [7]. We are decedents of a small aboriginal "founder" population some 10,000 generations or 250,000 years ago. Obviously, present-day homosapiens exhibits a wide variety of phenotypic, or physically expressed, traits. For some phenotypic traits, simple, single base differences at a single gene, or locus, account for observed variations. On the other hand, there are other individual traits (e.g., height) that are not attributable to different alleles, or mutant genes, at a single locus. In this case, many differenct genes, each with allelic variations, contribute to the total observed variability in a trait, with no particular gene having a singly large effect [8].

The existence of genetic risk for both rare and common diseases is generally accepted. Many of us are aware of diseases for which we are at risk because of affliction among our parents or siblings, for example. However, disagreements exist concerning the extent to which diseases are predominantly mono-allelic, and thus caused by a single mutation, or multi-allelic.

1.4.1 Mono-allelicism

The single allelic, or mono-allelic, effects are by their nature rare and inherited from a single progenitor. In 1964, Kimura and Crow [9] postulated that risk for diseases among genetically heterogeneous populations is wholly or predominantly represented by descended copies of a single mutational event [10, 11]. According to their argument, genes would be expected to carry one or fewer gene altering mutations in a small aboriginal population and that all individual point mutation rates were so low that the chance of the same mutational event occurring more than once in human genetic evolution was negligible [9, 12, 13]. Monogenic disorders are often rare because negative selection reduces the frequencies of variants that cause diseases characterized by early-onset morbidity and mortality [14]. Monogenic disorders are also required

to be highly penetrant, meaning that most or all of the individuals with a specific genotype show the disease phenotype. In addition, where multi-allelicism was once present, selective evolutionary pressures can skew risk toward a single risk-conferring allele [15, 16].

1.4.2 Multi-allelicism

In contrast, the multi-allelic traits (alternatively, oligogenic or multigenic) are more complex, and could have arisen independently in human history, rather than inherited from a single founder population. In relation to disease, multi-allelicism implies that no single locus contains alleles that are necessary or sufficient to confer risk [11].

One popular hypothesis regarding the genetic cause of common diseases holds that the genetic factors underlying common diseases will be alleles that are themselves quite common in the general population [17, 18]. This is known as the common disease-common variant hypothesis (CD-CV). It predicts that the alleles responsible for common diseases will be found in all populations. There are locations in the genome where variations between individuals are common, known as single nucleotide polymorphisms (SNPs), independent of population. These SNPs are concentrated in coding and regulatory sequences in genes. According to CD-CV, complex polygenic diseases, which are evolutionary neutral (not affecting fecundity), are caused by these variations. Each variation in a complex disease will have a small effect on the disease phenotype resulting in additive or multiplicative effect of many susceptibility alleles.

A critical assumption about CD-CV is that there is little allelic heterogeneity within loci–that the mutations which cause the disease are identical for people in different families. However, if a gene contains low frequency mutations at many different locations, then the power of tests for the CD-CV hypothesis will be greatly reduced [19]. In other words, mutations that confer risk for a common disease in one family may only account for a small fraction of the total number of people afflicted in the general population. In fact, there is nothing in the basic process of evolution that requires gene-disease relationships to be strong or dominated by one or a small number of alleles or loci, and for complex traits we know the opposite is often true

[20].

Additionally, the selection of SNPs for study under this hypothesis become crucially important. In genome-wide studies, SNPs are chosen according to whether they are expected to effectively represent the total existing variation in the genome. This becomes problematic because such estimates of representation can underestimate the complexity of the problem [21]. The estimation is often based on extrapolation from very small samples of data with strong linkage to neighboring regions, or predictions based on theory, both of which do not fully capture the stochastic variance [22, 23, 24].

In fact, choosing two or three SNPs per gene as representatives of the underlying genome is inherently risky considering the pattern of mutations observed by Benzer and Freese in 1958 [25]. They demonstrated that reproducible, non-random mutations can be created at locations in genes termed "hotspots," by treating bacteriophage with mutagens. Figure 1-1 illustrates a representative spectra of mutations in a gene for a population of people. The plot shows the fraction of people in a population group, M_f, that have a mutation at various locations in a gene. Within a single gene for a large population (e.g., N=10,000 people), there will be several high frequency mutations and many lower frequency mutations, some present at frequencies as low as once in the population or M_f=10^{-4}. High frequency mutations, usually caused by deamination of the nucleotide base cytosine (C→U), can inactivate gene function, either with deleterious (reducing the number of offspring of the person with the mutation) or non-deleterious consequences. These results have since been confirmed in bacteria, human cells, and human populations [27, 28, 29]. Although not fully characterized in the literature, there are expected to be ∼10× more neutral mutations than gene-inactivating mutations. Specifically, current estimates are for 0.03±0.81 gene-inactivating mutations/exonic segment and 0.33±0.89 neutral mutations/exonic segment [26, 30, 31, 32, 33].

Were a gene-inactivating mutation non-deleterious, it would be expected to survive through human history. Late-onset diseases are only now becoming common because our ancestors didn't live long enough to get them–they were able to pass their genes to the next generation before the disease struck. As a result, complex traits whose

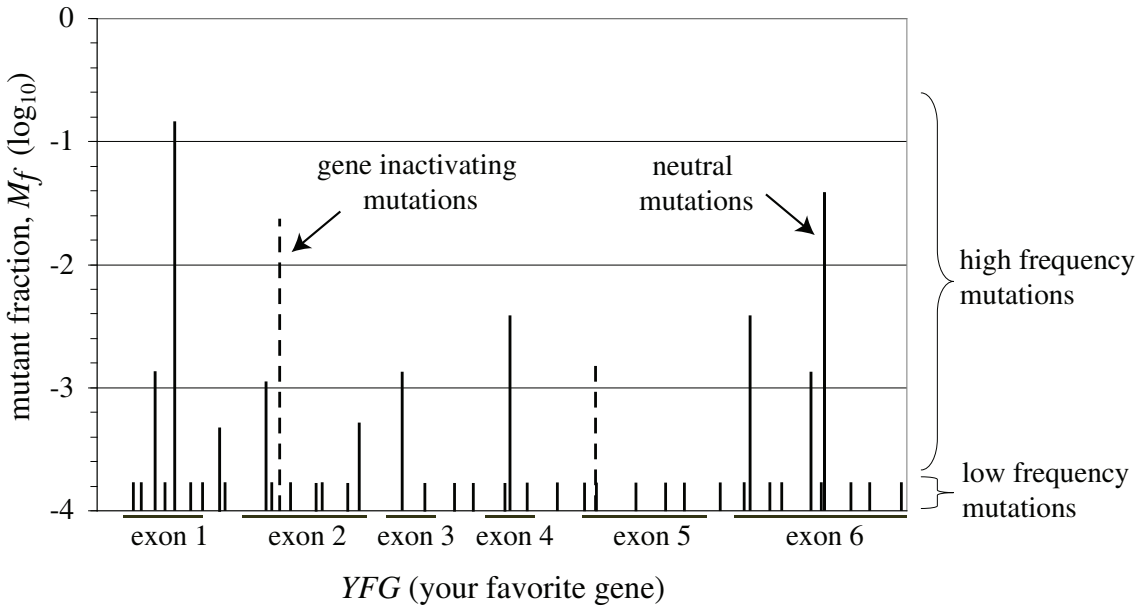

Figure 1-1: Mutation frequency versus location within a typical gene. Non-random sets of mutations occur in human genes at varying frequencies. Many different mutations exist at low frequencies in a population (e.g., 10^{-4}), while a few exist at higher frequencies. Common (high-frequency) or rare (low-frequency) mutations can be either gene-inactivating or neutral, though gene-inactivating mutations are expected, from scant observations, to be present at frequencies of $\sim$10$\times$ less than neutral mutations [26]. Mutations can exist in exons (coding regions) of the gene as well as in introns (regions between exons), but most (92% [http://www.hgmd.org]) deleterious, gene-inactivating mutations discovered are within exons.

genetic basis is non-deleterious are likely to have "noisy" genetic architectures [21]. This stochastic variance and the range of relevant mutation frequencies is, perhaps, an oversight within the SNP community. Additionally, the work of Benzer and Freese implies that many identical mutations that occurred independently would be expected to be found in the population today.

1.5 Strategies for apprehending gene-disease associations

By any method utilized for discovering the causes of *common* diseases, only two genes have been convincingly associated. They are the following:

- Melanocortin 1 receptor gene (*MC1R*), associated with skin cancers [34].
- Complement factor H gene (*HF1/CFH*), associated with macular degeneration [35, 36, 37].

Clearly, there is a paucity of data. In this Section we review the strategies for discovering such associations and their results. Lastly, we introduce a strategy devised by W.G. Thilly [26], which is the foundation for the UTMS instrumentation development.

1.5.1 Linkage disequilibrium

Where mono-allelism is expected to confer risk for disease, a powerful experimental technique known as linkage disequilibrium (LD) has been widely applied [10]. In this technique, genes are mapped by typing genetic markers in families to identify regions that are associated with disease more often than are expected by chance. Such linked regions are more likely to contain a causal mutation. For LD analyses to succeed, markers that flank the disease gene must segregate with the disease in families. Since monogenic markers are inherently rare and highly penetrant, the theory holds that each segregating disease allele will be found in the same 10,000-50,000 nucleotide base

chromosomal background in each family, and will co-segregate with disease status [38]. This background DNA segment is known as a haplotype. By sequencing the haplotype, the risk conferring allele and carrier gene can be discovered.

Using this technique, the genetic causes for over 2000 rare diseases have been discovered. This phenomenal success is attributable to the underlying monogenic, "Mendelian," nature of these diseases [39]. Some of the notable successes include the hemoglobin gene (*HBB*) base pair substitution that causes sickle-cell anemia among Africans carrying two mutant alleles but confers malarial resistance among heterozygotes carrying only one [15]; and a three base pair deletion in the *CFTR* gene that accounts for more than 80% of the alleles conferring risk for cystic fibrosis among northern Europeans but also confers apparent resistance to gastrointestinal infections among heterozygotes [16].

Due to its success with rare, monogenic diseases, this technique has also been applied to common diseases under the CV-CD hypothesis. Unfortunately, only about one gene has been discovered and convincingly linked to common disease using LD, that is the complement factor H gene (*HF1/CFH*), associated with macular degeneration [35, 36, 37]. Where other common disease-gene association have been claimed, as in the case of inflammatory bowel disease (IBD) [40, 41, 42, 43], schizophrenia [44], type I diabetes [45], and type II diabetes (*TCF7L2* association) [4], the genes discovered at most explain only a small fraction of the overall heritability of the disease, or have not be validated in subsequent experiments. For example, mutations known to confer risk of IBD together explain excess risk to siblings of about 2×, while in reality the total excess risk in excess of 30×, indicating that many other causal genes are yet to be discovered [46].

There have been variations and improvements on LD for systematic study of relatively large population sizes (∼1,000 persons), genome wide. These fall under the labels of cataloguing SNPs and performing association mapping [47, 48, 49], the identification of restriction-fragment length polymorphism markers [50], the identification of abundant highly polymorphic microsatellite (short tandemly repetitive DNA loci) [51, 52], and using staged approaches to increase statistical power [53].

However, LD has many limitations to the discovery of multi-allelic causes of common diseases such as low heritability of complex traits, low penetrance of alleles contributing to complex diseases, the inability of a standard set of haplotypes represented by SNPs, spaced 10-50 kb apart, to extract complete information about inheritance [54, 55, 56, 57, 58], assumptions that the allele being measured is the actual susceptibility allele, imprecise definition of phenotypes [59], assumptions of allelic heterogeneity within loci [60], and inadequately powered studies [8, 19]. Most importantly though, ignoring the stochastic variance inherent to genetic mutations in humans has confounded nearly all attempts for common disease association with multi-allelicism.

1.5.2 Proteomics

Another approach to identifying the genetic causes of common diseases involves directly characterizing the genes expressed in normal and diseased cohorts. Labeled "proteomics," or "functional genomics," this effort could circumvent the need to sort through complex genetic variation to find the small minority of important sites [61, 62, 63, 64]. However, important challenges in elucidating cause and effect relationships remain unsolved. For example, in a cancer cell expressing tens or hundreds of genes associated with abnormal growth, it is painstakingly difficult to determine which fraction are causal [21]. Furthermore, environmental factor are known to confound gene expression mapping (e.g., [65]).

1.5.3 Candidate-gene resequencing

DNA sequencing, or identifying every base in a continuous fragment of DNA, is the gold standard of mutation detection. One could consider sequencing suspect exons, genes, etc., and comparing the mutations in the disease cohort with a control cohort. Due to inherently low throughput, or alternatively high cost (by 3-4 orders of magnitude), study sizes are currently small and limited to a few genes, carefully predicted on the basis of biological hypothesis or within a previously determined

region of linkage. Thus statistical stringency is compromised. In addition, rare gene-inactivating mutations can present interpretation challenges [66].

1.5.4 Genome-wide association studies

Researchers generally agree that the method that has the highest likelihood of associating genetic mutations with common diseases is that of genome-wide association studies. Ideally, we would simply like to compare the entire genomes of a large group of people with a common disease to a large group without. As put summarily by Davis and colleagues [67], "comprehensive strategies that can collect the full spectrum of human genetic variation...[for]...large numbers of genes or the entire genome will enable the most powerful approaches to unlocking the power of human genetics."

By comparing frequencies of all point mutations across the genomes for case and control cohorts for each disease, statistically significant differences in mutation frequency could be recognized. More specifically, one would expect that a risk conferring gene would contain an excess of gene-inactivating mutations for the diseased cohort relative to the control cohort, while the two groups would be expected to have equal number of neutral mutations (see Figure 1-1). Were a gene not related to the disease, both groups would have identical spectra, with equal frequencies of gene-inactivating and neutral mutations. This type of comparison would enable the capture of the widest possible range of genetic possibilities, including mono-allelic risk and multi-allelic risk arising from surviving inherited or spontaneous mutations in more than one gene or from interaction of mutations in more than one gene including epistatic interactions [26].

In practice this a monumental effort and has not been attempted. Its magnitude is mind-boggling. To date, the entire genomes, comprising 3×10^9 bases, of only a few people (i.e., about two) have been thoroughly characterized, in this case by sequencing. A meaningful study would require on the order of 10^3-10^6 people. Are there ways to reduce the scope of the task sacrificing stringency for expediency and cost reduction? If so what is are the wisest trade-offs? And can alternative technologies be brought to bear?

Limiting the scope of the effort

The primary method for reducing the scope of the effort is deciding not to examine parts of the genome which are less suspicious or more difficult to examine. Here we discuss several methods and their justifications for doing so.

There are estimated to be at least 1.5×10^7 SNPs in the human genome at frequency of 1% or greater [68, 69]. All SNP databases combined contain less than 10^7 [71, 70, 2], so detecting and enumerating all of them would only account 67%. Technologies (discussed below in Section 1.7) and strategies which detect SNPs specifically can not practically achieve even this, so efforts have primarily been focused on limiting the search to variants of frequency 5% or more ($M_f = -1.3 = \log_{10}0.05$ in Figure 1-1). One such effort, known as the International HapMap project [3], is currently (as of early 2007) scanning $\sim3.9\times10^4$ SNPs per person in a total study of $\sim300\times10^6$ SNPs for all persons in a study of the common disease, type II diabetes. This approach is justified primarily based on assumptions about LD. Such an approach will, however, inevitably ignore more than 14.6×10^6, or 97%, of high frequency (>1%) mutations *per person*. Of course, 100% of mutations at frequencies less than 5% are ignored as well. This is to say nothing of the statistical power of the study, which may also be lacking if assumptions about the about the nature of mutations are not fulfilled.

It may therefore be more prudent to seek to detect *all* mutations to the lowest frequency theoretically possible (1 individual/population) in *all* exonic coding sequences, if it can be done technologically, within reasonable time and cost [29]. The choice to ignore introns, promoter regions, and other non-coding regions is not without risk, but is based on the knowledge from the Human Gene Mutation Database [http://www.hgmd.org] that 92% of deleterious, gene-inactivating or altering mutations that encode for risk of rare inherited diseases are found in exons and splice sites. Fortunately, exons and splice sites comprise only 1-2% of the genome.

From studies in mice [72], only about 10% of genes can carry non-deleterious, gene-inactivating mutations, so one might want to limit a genome-wide scan to these

genes. However, they haven't been identified in humans and we do not know enough about the genetics of common diseases to make such a choice [26].

1.6 Statistical significance

In the comparison between frequencies of mutations in case (diseased) and control cohorts, we must consider whether the differences that will be measured in any study are statistically significant, or stringent. These terms refer to the power of the test to reject the null hypothesis–that the difference in frequencies of mutations at a particular location between the case and control cohorts is *not* significant, and can be ascribed to chance alone. Hundreds of initial published findings of gene-disease associations have been subsequently invalidated by lack of significance [73].

Generally, one rejects the null hypothesis if the probability, p, of obtaining a result at least as great as the one measured is smaller than a threshold significance level, α, typically 5%. For example, if a coin tossed 20 times turns up heads 14 times, the probability of 14 or more heads or 6 or less tails is $0.0577\times2=0.115=p$. Since $p>0.05$, we can not reject the null hypothesis that the coin is fair. However, had we flipped 15 heads out of 20 tosses, $p=0.02<0.05$, and therefore we should reject the null hypothesis that the coin is fair and conclude that the coin is biased.

Without delving into great detail, one must ensure that the ratio of true positives to false positives, as the outputs of a genome-wide association study, is reasonable. As derived by Morgenthaler and Thilly [26], we should prudently assure that the difference between the number of mutations in the case and control cohorts in a gene fragment exceeds the threshold corresponding to 0.05 scaled by the variance in the mutation distribution observed in the case and control populations and corrected by the number of hypotheses tested. This will assure that the null hypothesis (the observed difference can be attributed to chance) can be confidently rejected. Tomita-Mitchell *et al.* [29] suggest, based on the spectra of SNPs observed in human genes from pooled populations, that sample sizes of 10,000 people would be required to statistically significantly identify multi-allelic, multigenic risks on the basis of summed

alleles in a single gene-disease association trial. This far exceeds the number of people used in almost all studies to date.

Given this exceedingly large sample size requirement, a range of techniques have been proposed to minimize this genotyping burden given a set of goals with respect to true and false discoveries of gene-disease associations. Analyzing their merits is beyond the scope of this text. Most fall under the category of multi-stage testing (e.g., [74, 53]), in which a two- or three- stage screening process is performed either on independent populations, or permutation of the same population. Each test can have a more modest threshold than 0.05, while retaining the overall stringency. The effectiveness of these strategies is not guaranteed, but depends on the distributions [38] and decreases as the proportion of the tests for which the null hypothesis is true increases [74]. These multi-stage approaches are currently being utilized by the genetics research community with varying success.

1.7 Technologies to discover gene-disease associations

A wide variety of mutation detection technologies exist that can be applied to the discovery of gene-disease associations [75]. In this section, we focus our attention on those few which have real potential to discover associations for common diseases by way of genome-wide association studies. In detail we discuss denaturing capillary electrophoresis, which has proven effectiveness in measuring high resolution mutation spectra in large populations.

1.7.1 Emerging high throughput sequencing technologies

As discussed briefly above in Section 1.5.3, sequencing of the entire genomes from 10^6 people would be ideal for our study, but is currently prohibitively expensive and time-consuming. By various methods which are beyond the scope of this document, sequencing technologies identify the specific bases in a continuous strand of DNA.

Emerging technologies, such as polony sequencing [76], are racing towards the goal of an entire human genome sequences for $1,000. Currently, these technologies have reduced the cost/kilobase $0.11/kb as compared to conventional, Sanger sequencing [77], at a cost of $1/kb. Thus for an entire genome with 10-fold coverage for accuracy would cost 3.3 million dollars. If a cost of $1,000/genome can be realized, 10 fold coverage of the 1% coding regions could cost merely $100/person. For that price, sequencing 10^6 people could be done for 100 million dollars. Clearly however, there is still much innovation required.

1.7.2 Gene chips for hybridization probes

In this method, arrays of short DNA fragments (oligonucleotides) are synthesized in fixed positions on a substrate. Target gene fragments are labeled, hybridized to the chip, and the chip is washed [78]. The pattern of labeling of up to hundreds-of-thousands of targets is detected and the sequence is then defined from the known position of the synthesized oligonucleotides.

This method is being widely employed in genomics studies, particularly in the HapMap project discussed above, for the enumeration of SNP markers in the digested genomic DNA fragments. Such a strategy is compelling, but for the shear number of persons required to be studied for statistical significance. As described above, only a small fraction of the SNPs for each individual in the study can be detected practically speaking, and this would need to be done for up to 10^6 persons. The HapMap project is showing the challenges associated with this technical approach based on the paucity of statistically significant results linking genes to common diseases.

1.7.3 Denaturing electrophoresis

Denaturing electrophoresis is a powerful mutation detection technique in which single-base mutations in short (100 bp-400 bp) double-stranded DNA fragments can be separated from otherwise identical, non-mutated (wildtype) DNA fragments based on their mobility differences while undergoing electrophoresis in a denaturing environ-

ment.

Electrophoresis is a common biological technique in which charged molecules are moved through a porous medium by application of an electric current in the medium. The molecules travel at a speed determined by a number of factors, including their charge and size, the viscosity of the medium, and the strength of the electric field generating the current. Thus, if a mixture of different sized or charged molecules are subjected to electrophoresis, they will stratify as they propagate towards the cathode, or positive potential.

Now, when a double-stranded fragment of DNA is treated with increasing temperature or chemical denaturant, the two strands come apart by breaking of hydrogen bonding between the base pairs. This behavior, or "melting," will vary depending on the specific sequence, as the bases adenine (A) and its complement thymine (T) are double-hydrogen bonded, while bases guanine (G) and cytosine (C) have triple-hydrogen bonding. Thus a sequence with more C-G pairings will melt under higher denaturant conditions. In addition, the arrangement of the bases, not simply their proportion, in the gene fragment, will affect melting behavior. Partial melting can occur when some, but not all of the hydrogen bonds have been overcome by the denaturant. In this state, a DNA molecule, partially melted, will have an effectively larger "size," and thus propagate slower during electrophoresis as compared to its fully double-stranded "unmelted" state.

By carefully choosing the denaturing conditions during electrophoresis, double-stranded DNA of differing sequences (e.g., mutant and wildtype gene fragments) can be separated from each other by this mechanism. Furthermore, we can chemically attach a "clamp," comprising about 20 bp of G-C, to one end of the double stranded fragments to prevent complete melting, or separation of the complementary strands. We can also chemically attach fluorescent markers or radioactive labels to the strands to make them easily detectable.

In practice, an additional procedure is performed to aid in the separation of the sequences. Before separation in a porous medium, they are completely melted and then allowed to reform. Were there only two original double-stranded sequences

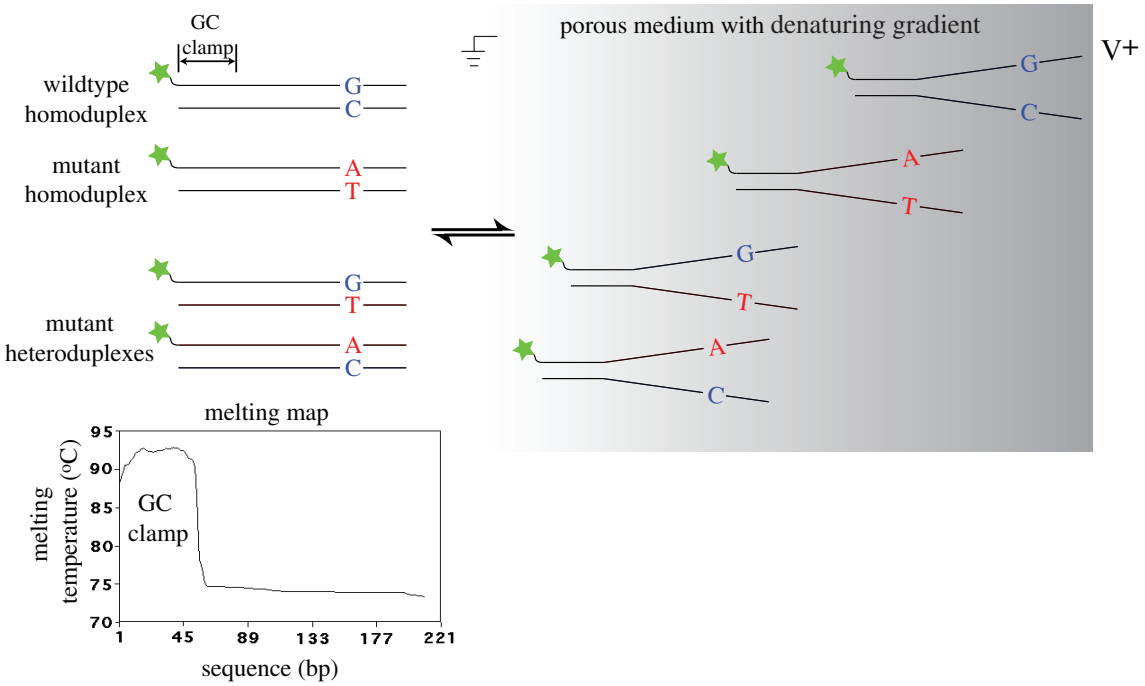

Figure 1-2: Principle of denaturing gradient gel electrophoresis. Four different double-stranded DNA fragments can be separated on a porous medium by the application of a denaturing gradient and electric field, because of the different conditions under which they melt.

differing by one base pair (i.e., a mutant and a wildtype gene fragment), one strand of one sequence will occasionally bond, not with its original complement, but with strand two of sequence two upon reforming. Likewise strand two of sequence one can bond with strand one of sequence two. In this way, one can generate four different double-stranded DNA fragments, where there were originally only two. Two will be the original complementary "homoduplexes," and there will be two mismatched "heteroduplexes." This is advantageous because the heteroduplexes will melt under much different conditions than the homoduplexes allowing easy separation because of their much different propagation speeds. In fact, even if the mutant homoduplex is not well separated from the wildtype homoduplex, the fraction of the original sample that was mutant, M_f, can be gleaned from information about the heteroduplexes relative to the homoduplexes.

This technique, originally denaturing gradient gel electrophoresis (DGGE), was developed by Fisher and Lerman in 1983 [79], and is illustrated in Figure 1-2. Lerman

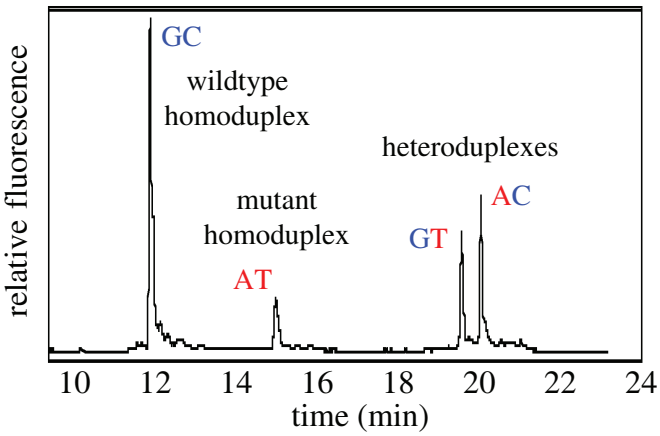

Figure 1-3: Electropherogram of a separation run using denaturing capillary electrophoresis. The proportion of the original DNA sample that was mutant homoduplex, called the mutant fraction, M_f, can be determined from the ratio of various peaks.

et al. further improved the practicality of the technique by creating algorithms [80], which are based on the theory of Poland, Fixman, and Freire [81, 82, 83], to accurately predict the so-called "melting map," or denaturing temperature as a function of base pair position. This maps, as shown in Figure 1-2 by way of example, represent the equilibrium temperature for equal probability of the helical and melted (denatured) states.

Monitoring DNA concentration in a DGGE separation, typically by fluorescence emission, at a fixed distance from the electrodes, yields a plot such as shown in Figure 1-3. The mutant fraction, M_f, in the original sample can be determined from this electropherogram. This measure is the critical output of denaturing electrophoresis, in the study of the frequency of mutations in populations, as in Figure 1-1 above. More specifically, M_f can be determined from the electropherogram as

$$M_f = \frac{A_{\text{mutant homoduplex}} + \frac{1}{2}A_{\text{heteroduplexes}}}{A_{\text{wildtype homoduplex}} + A_{\text{mutant homoduplex}} + A_{\text{heteroduplexes}}}, \quad (1.1)$$

where $A_{\text{heteroduplexes}}$ is the area under the two neighboring heteroduplex peaks and $A_{\text{wildtype homoduplex}}$ and $A_{\text{mutant homoduplex}}$ are the areas under the respective wildtype and mutant homoduplex peaks.

The porous medium through which the DNA propagates can be formed into a sheet, or slab, or it can be loaded under high pressure into a long, narrow capillary tube or channel. The application of DGGE to a capillary tube was pioneered by Khrapko *et al.* in 1994 [84], to enable denaturing capillary electrophoresis (DCE). Furthermore, it was shown that, rather than a denaturing gradient as shown in Figure 1-2, a zone of constant denaturant can be used to separate DNA fragments according to sequence. This variant is called constant denaturing capillary electrophoresis (CDCE). Still further, temporal [85, 86] or spatial [87] denaturing gradients, typically applied with temperature ranges of several degrees may be advantageously applied to ease the denaturant accuracy restriction. As long as the gradient spans the required temperature, high efficiency separations can still be achieved.

Mutation scanning example using denaturing capillary electrophoresis

In this section we introduce an example of the power of DCE to detection mutations in genes of pooled population samples. Li-Sucholeiki *et al.* [88] have used this technique to study mutations in the gene β-globin from two groups of people: 5000 African-Americans and 5000 Han Chinese. The results are shown in Figure 1-4. From this result, we can see, for example, that a small fraction of the African-American population carries a mutant allele conferring risk for sickle-cell anemia, while the Han Chinese population does not. From this small study, one can grasp the potential for mutation scanning across all genes in a million people as a way to compile high resolution spectra of mutations corresponding to different diseases, in the search for statistically significant differences in mutation frequency between case and control cohorts, as in this example for sickle-cell anemia, a mono-allelic disease.

To obtain such data, blood samples from the 10,000 people were obtained. Blood was pooled into group sizes suitable for DCE (e.g., 100 persons per pool). DNA was then purified from the blood's white blood cells. An amplification technique called polymerase chain reaction (PCR) was used to effectively excise and amplify each exonic gene fragment, while simultaneously adding the G-C clamp region and fluorescent marker to the gene fragment. These pooled, amplified, labeled gene fragments

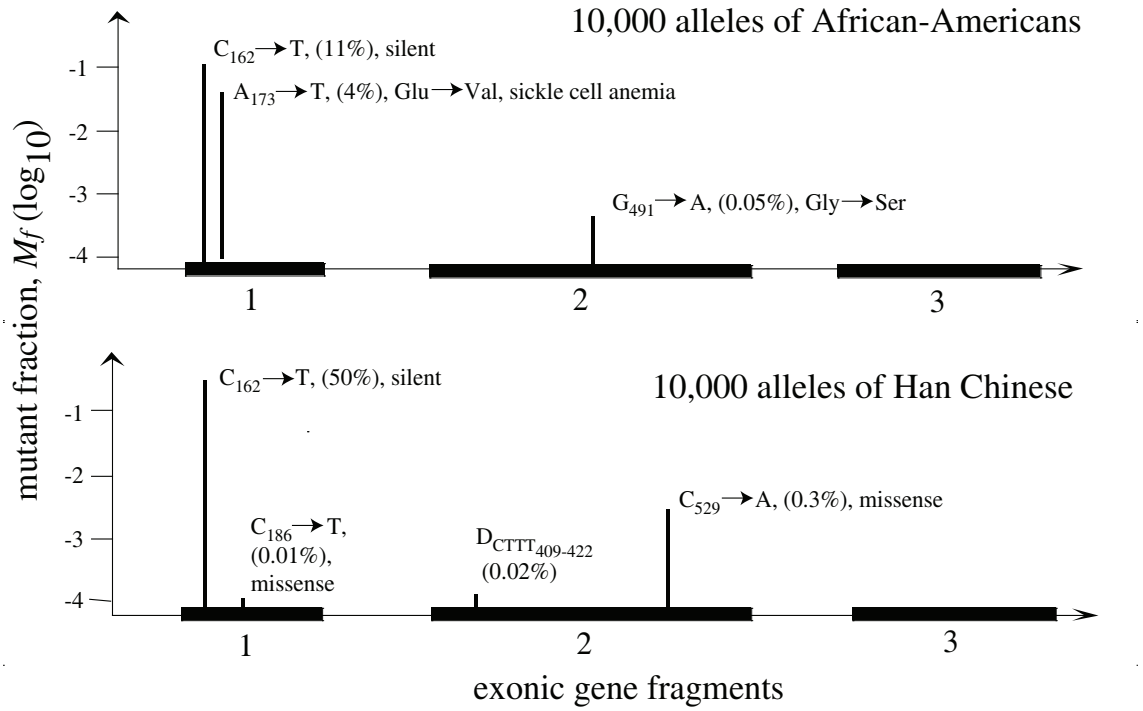

Figure 1-4: Mutation scan in all three exons of the human β-globin gene in two population groups: 5000 African-Americans and 5000 Han Chinese. These were the first high-resolution spectra of human populations. (adapted from [88])

were then run on a DCE instrument at the appropriate denaturing conditions, as determined using Lerman's algorithm, in a porous medium of linear poly(acrylamide).

1.8 Cohort allelic sums test–CAST

In the preceding sections, we have discussed the value of identifying all mutations in all exonic gene segments of both a case and control cohort for the purpose of identifying associations between mutations and common diseases. In this Section we lay out specific implementation of a genome-wide association study, as introduced by Morgenthaler and Thilly [26], which serves as the motivation for the development of the 10,000 capillary UTMS instrumentation.

In 1998, Tomita-Mitchell *et al.* [29] suggested, based on the spectra of SNPs observed in human genes from pooled populations, that sample sizes of 10,000 people would be required to detect multi-allelic, multigenic risks on the basis of summed alleles in a single gene-disease association trial. In Morgenthaler and Thilly [26], this proposal is extended to a study of 100 diseases in parallel, for a total of 100 diseases $\times$ 10,000 people/diseasee = 10^6 people. The benefit of such a parallel approach is that the control cohort, then comprising 99 $\times$ 10,000 = 990,000 people, can be used to robustly define the mean and distribution of the sum of mutations carried in the general population for each gene.

For each person, the proposed approach involves scanning the exonic coding regions of each gene for mutations. There are 25,000 genes/genome and 2 autosomal genome copies per person. Using DCE mutation scanning technology, the coding regions of each gene can conveniently be represented by about 20 sections of 100 bp of DNA. So the total number of gene fragments per person is 25,000 genes/genome $\times$ 2 genomes/person $\times$ 20 gene fragments/gene = 10^6 gene fragments/person.

Thus, in the proposed cohort allelic sums test, or CAST, as described, mutation scanning in 10^6 gene fragments/person $\times$ 10^6 people = 10^{12} gene fragments is required. The task can be somewhat reduced by pooling samples in DCE. Grouped into pools of 100 persons, merely 10^{10} capillary separation runs would be required.

Were 400 96-capillary instruments used for this task, roughly the throughput of existing genome sequencing centers with billion-dollar infrastructures, this task would require approximately 15 years. On the other hand, 50 10,000-capillary instruments could accomplish study in less than three years, for an estimated cost of 500 million U.S. dollars. To emphasize this point, a facility with 50 10,000-capillary denaturing capillary electrophoresis instruments running 24 hrs/day could systematically identify the frequencies of mutations in the coding regions of the entire genomes of one million people grouped into 10,000 persons/disease, at dramatically reduced time and cost as compared to other technologies. Such a study could revolutionize biotechnology, permitting the identification of gene-disease associations for the 100 most common diseases and is the fundamental motivation for the work described in this thesis.

Thus we set out to design and manufacture a 10,000 capillary denaturing capillary electrophoresis instrument. This is a massive undertaking. For rough comparison, the development of a prototype 768 channel DNA sequencer [89] at MIT's Whitehead Institute required 10 years, $10M, and 20 full-time post-doctoral staff and graduate students [Personal communication with Dan Ehrlich, Whitehead Institute, September 20, 2006]. In this work, we designed, built, and tested all of the key subsystems for the 10,000 capillary instrument as proof-of-concept, which lays the foundation for the realization of a mutational spectrometer instrument which will enable population-wide pangenomic analyses to uncover the genetic causes of common diseases.

1.9 Ultra-high throughput denaturing capillary electrophoresis

A schematic of the ultra-high throughput mutational spectrometer (UTMS) instrument is shown in Figure 1-5. The instrument enables parallel loading, separation, detection, and sequestration of DNA mutations from pools of 100 persons per channel for 10,000 parallel channels simultaneously and continuously, utilizing denaturing capillary electrophoresis.

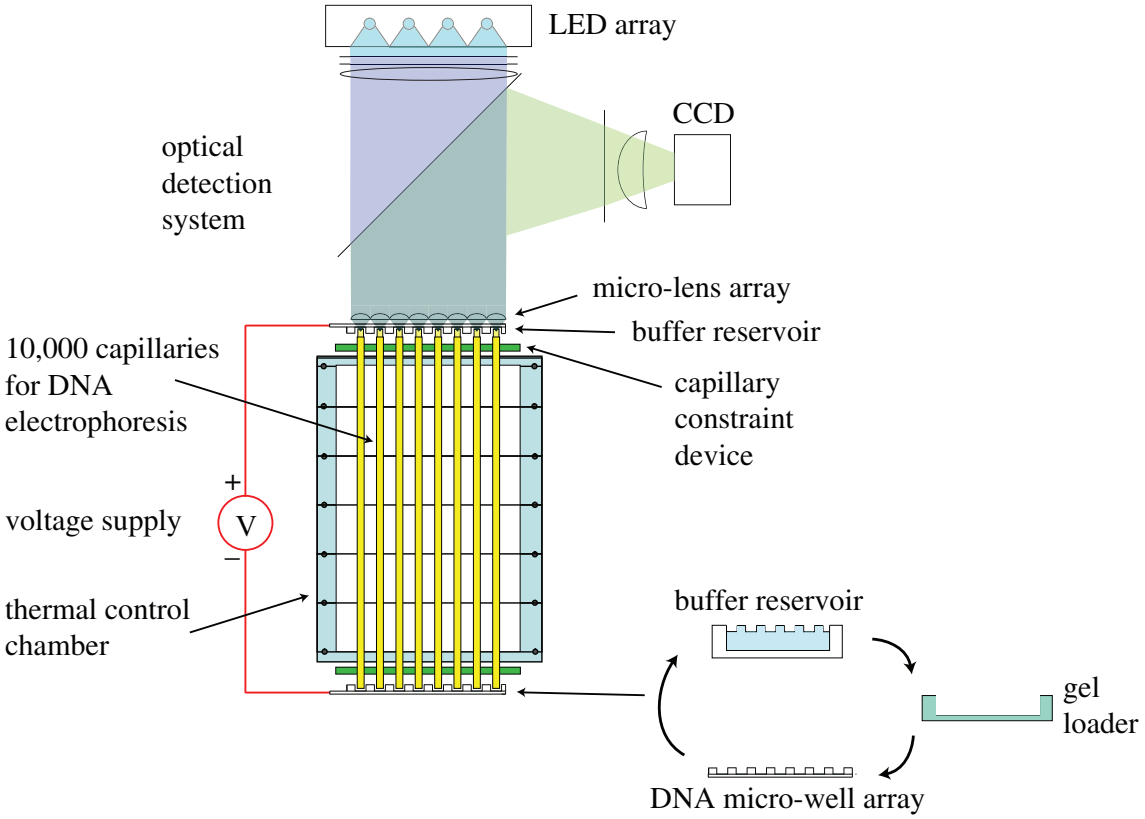

Figure 1-5: Schematic of the ultra-high throughput mutational spectrometer.

On the lower half of the schematic is a bundle of 10,000 capillaries arranged in a rectangular array of 100×100, spaced 1 mm apart, as positioned by a pair of capillary constraint devices. The entire array can be thermally and electrophoretically controlled for separating mutant from wildtype DNA. A variety of devices can interface with the bottom of the capillary array. In use, we first mate a gel loading device to fill all of the capillaries with a viscous gel, or sieving matrix. Second, the capillary array is loaded with DNA using a micro-well array which literally plugs into the capillary array to electrokinetically inject 10,000 parallel samples from corresponding 250 nL reservoirs. Subsequently, the micro-well array is replaced with a reservoir of buffer solution for the electrophoresis run. At the upper end of the capillary array a second buffer reservoir is placed after the gel is loaded. In the upper half of the schematic is the optical detection system. An LED (light-emitting-diode) array is used for excitation. The light from the LED array passes through some optical elements and is then focused by a lenslet array. Each lenslet, one per capillary, focuses a portion of

the incident radiation onto the tip of each capillary. If fluorescently-labeled DNA is present in the tip of the capillary, fluorescence is excited. The subsequent emission is captured by the lenslet array and relayed to a CCD detector. The entire instrument stands about a meter tall, with 300 mm long capillaries.

1.9.1 Design justification

In the proposed UTMS design, we have opted for a combination of macro/micro manufacturing processes to achieve 10,000 capillary electrophoresis capability. There have remarkable innovations in electrophoresis instrumentation recent years, in which the separations are performed, not in "macro" capillary tubes, but in "micro" etched channels (e.g., [89, 90]). This design change has greatly facilitated rapid, integrated analysis of few samples. As these designs have substantially impacted the field, we should consider, why not use 10,000 etched channels for our instrument? In short, our fundamental requirement is for extremely high throughput, where the scalability of the etched channel approach is limited.

There have been demonstrations of 384 80 mm radially arranged etched channels on an 200 mm circular glass wafer [90] and 384 450 mm long, parallel, etched channels on a 250 mm $\times$ 500 mm glass plate [91]. In both of these cases, the channel density was maximized on the available real estate. Shimadzu Corp. (Japan), in cooperation with Network Biosystems (Woburn, MA), has commercialized a 768 channel sequencer which uses 2 glass plates of the latter design, with a single high power laser scanner. An instrument with 10,000 parallel channel of this design would require no less than 13 separate scanning laser assemblies, gel loaders, DNA sample loaders, high voltage power supplies for electrophoresis, etc. For the optical detection system, for example, the scanning laser confocal microscope design serially interrogates the channels as it rotates. The duty cycle per capillary approaches signal integration limits in the Shimadzu Corp. design, and is thus impractical for more than 1,000 capillaries. In contrast, our novel detection approach requires no moving parts, a single 1 megapixel detector, a low-cost LED array, and a lost-cost lenslet array. This detection system is lower in cost by an order of magnitude than the scanning confocal design, is much

simpler to implement, and more robust, lacking the requirement for complex scanning control hardware and algorithms.

In addition, the manufacturing costs for 10,000/384 = 26 glass channeled wafers are comparable to the cost of a single 10,000 channel capillary array, when considering the addition of the detection systems, gel loaders, sample loaders, heaters, power supplies, etc. The 10,000 capillary array costs are lower when considering our ability to replace individual capillaries at a cost of approximately $2, instead of replacing entire etched assemblies of glass plates. There are additional advantages to the use of capillary arrays over etched channels in regards to more simple thermal control, mitigation of band broadening, gel loading, collection of eluted fractions of DNA, and independent electrophoretic control.

1.9.2 Functional requirements

Detailed functional requirements for each of the subsystems required for the UTMS are presented in each Chapter. Here we introduce general requirements as motivation for the design. In the work presented in this thesis, we sought to address each of the functional requirements and integrate the resulting multi-disciplinary subsystems into a working instrument. The UTMS subsystems must enable:

- DNA manipulation.
  - Load DNA into capillary array and collect separation fractions in independent fluid reservoirs while permitting detection.
- Optical signal detection.
  - Measure 10^7-10^{10} biomolecule fluorescence from 10,000 independent channels simultaneously sampled.
- Mechanical alignment, constraint, assembly.
  - Assemble and constrain 10,000 capillaries accurately and repeatably, permitting replacement, fluidic sealing.

  - Alignment of detection, manipulation, separation elements.

- Thermal control.
  - Control the temperature of a 250 mm section of 10,000 capillaries uniformly to 0.4 °C between 50-90 °C while array outputs 200 W.

- Signal processing.
  - Signal-to-noise ratio (SNR)>3 for 10^7 molecule fluorescence, with sufficient temporal and spatial resolution for identification of single-base mutations.

- Electrophoretic control.
  - Enable electrophoresis in 10,000 capillaries at 5 kV, 10 μA/capillary.

- Separation matrix handling.
  - Load gel, clean 10,000 capillaries simultaneously.

This thesis is organized into Chapters for each subsystem: thermal control, gel loading, micro-well array, buffer reservoir, capillary array, and detection system to address these functional requirements. In addition, Chapter 8 presents DNA electrophoresis experiments with the integrated subsystems, including mutation detection using denaturing capillary electrophoresis. Chapter 9 summarizes the contributions of this work, presents futures work and alternative applications of this instrument.

Chapter 2

Thermal control

2.1 Introduction

To separate mutant duplex DNA from wildtype using denaturing capillary electrophoresis, the DNA must migrate through a denaturing environment. As the UTMS is intended for use with any of the exonic gene fragments in the human genome, we desire simple control of the strength of this denaturing environment. This is most easily implemented by externally controlling the temperature of the sieving matrix inside of the capillary array.

2.1.1 Functional requirements

As shown in Figure 2-1, DNA denaturing temperatures range from 50-90 °C, assuming no other denaturants are present.* According to Figure 2-1, approximately 4% of genes (e.g.,*CYP 2D6*) contain exons that require separation temperatures above 80 °C. This requirement can be problematic because even DNA sequence clamps with melting temperatures above 100 °C will dissociate in the electric field at these high temperatures. As our collaborators work to overcome this challenge with a biological solution, we restrict the instrument temperature range required as 50-80 °C.

The tolerance, or accuracy required, on the temperature is approximately 0.4 °C,

*The denaturant urea can reduce the temperature required for separation by 3 °C/M/L of sieving matrix.

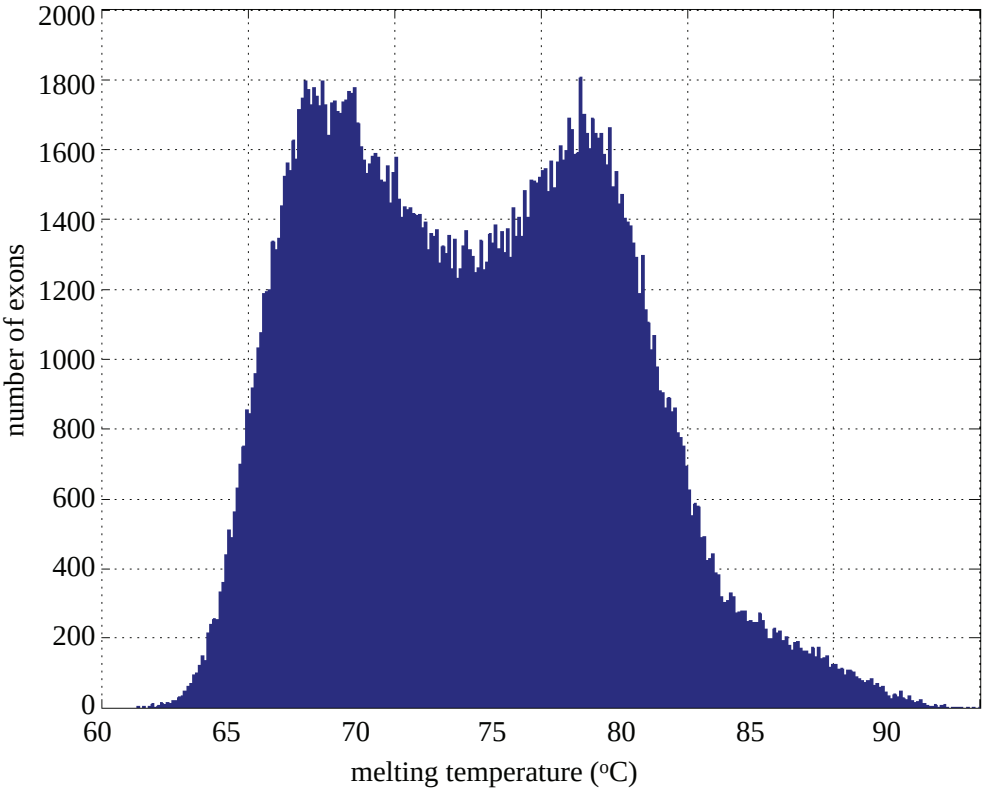

Figure 2-1: Distribution of exon melting temperatures in the human genome. Melting temperatures of all 236,039 exons from a set of 27,561 genes predicted for the human genome by Ensembl, a joint project of EMBL-EBI and the Sanger Institute. Resolution is 0.1 °C. Figure provided by PubGene, Inc., Oslo, Norway.

as empirically determined. This accuracy represents the required *uniformity* of the sieving matrix temperature within all capillaries in the instrument at a given time. For CDCE, this uniformity accuracy requirement applies axially, over a capillary length of 100-250 mm, as well as laterally across the array.

Temperature control is complicated by Joule heating, a consequence of electrophoresis. Joule heating refers to the increase in temperature of a conductor (i.e., sieving matrix) as a result of resistance to an electrical current flowing through it. Experimentally, a typical electric field strength of V=10 V/mm applied to a L=300 mm long capillary (inner diameter, d_i=75 μm, outer diameter, d_o=360 μm) results in a current of 9 μA. Since P=IV, the total power output, P, of 10,000 capillaries would be 270 W or a heat flux, q'', of

$$q'' = \frac{\text{power}}{\text{surface area}} = \frac{IV}{\pi d_o L} = 79.5 \text{ W/m}^2, \tag{2.1}$$

where I is the current, V is the voltage. Thus, efforts to externally control the temperature of the sieving matrix must overcome the heat flux due to Joule heating.

There are several variants of denaturing capillary electrophoresis (DCE) which may be implemented as discussed in Chapter 1, Section 1.7.3, that have been shown to yield single base mismatch mutation detection sensitivity. Constant DCE refers to a zone of constant temperature through which the analytes pass, this temperature being slightly less than the highest denaturing temperature required to melt a DNA duplex in the system. Instead of constant DCE, temporal [85, 86] or spatial [87] gradients with ranges of several degrees may be advantageously applied to ease the thermal accuracy restriction. As long as the gradient spans the required temperature, high efficiency separations can still be achieved.

2.1.2 Capillary array geometry assumptions

For the design of a thermal control system, the capillaries are assumed to be in a 2-D square array of 100×100 capillaries spaced 1 mm radially apart. This packing geometry was selected considering the field of view of the end-of-column fluorescence detec-

tion system, array manufacturability concerns, and space between the free-standing capillaries to permit flow of a heat transfer fluid.

Such a geometry is amenable to thermal control with a cross-flowing heat transfer fluid. A diagram of this heat exchanger configuration is shown in Figure 2-2.

2.1.3 Air versus water as the heat transfer fluid

The choice of heat transfer fluid in the thermal control system is not straightforward. Water has a heat capacity (ρc_p) about 5000× that of air, as well as a convection coefficient, h, about 100× that of air. The implications are that flowing water across the array will have minimal spatial gradients and it will be able to change the capillary temperature very rapidly (<1 s). More specifically, in contrast to the steady-state spatial thermal gradient results for water (0.2 °C at 20 L/min), air would need to flow at 15051 L/min (0.25 m^3/s) for ΔT=1 °C. Regarding rapid capillary temperature changes, lumped parameter modeling has shown that air would require at least 7.5× longer to reach steady state. In addition, while vortex shedding is not a concern with water, the Reynolds number for air flow to maintain the desired spatial thermal gradient would be approximately 100, with a velocity of 0.25 m/s. This would excite frequencies around 4000 Hz, which could excite higher vibration modes in the capillaries since their natural frequencies are 10-23 Hz depending on constraint type. The equations used for this analysis will be presented in the course of this Chapter.

These are compelling advantages. However, water poses additional challenges such as sealing, pumping, and the inability to rapidly change (e.g., cycle) its temperature. Air is utilized in all commercial multi-capillary instruments (e.g., Applied Biosystems, Foster City, CA; Amersham Biosciences Corp., Piscataway, NJ; SpectruMedix, LLC, State College, PA). But these instrument cannot heat above 70 °C and do not have such large Joule heating to overcome from the massive array of 10,000 capillaries. Based on this argument, water was selected as the heat transfer fluid in the thermal control system.

In addition, a solid-state heater jacket was observed at People's Genetics (Woburn, MA), based on the design by Karger *et al.* [92, 93]. This design features a thermally

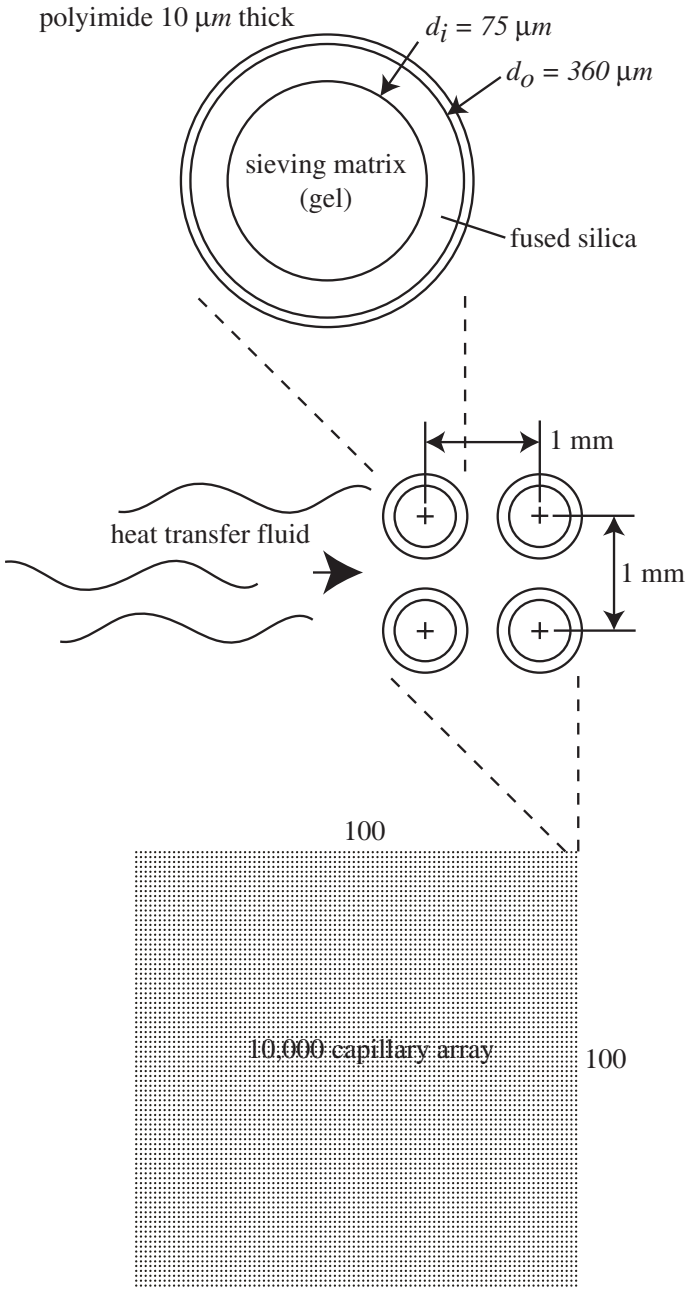

Figure 2-2: Capillary dimensions and their arrangement in the array. A cross-flowing heat transfer fluid can be used to control the temperature of the sieving matrix for optimal DNA separation.

conductive jacket wrapped around a bundle of capillaries. The jacket is heated by a Peltier device or resistors, while the temperature is sensed by thermistors or thermopiles and heat is dissipated by convection to ambient air. While the temperature control is very stable and precise (±0.013 °C [94]), simple to implement (current control), and has sufficient range, this design concept was not selected due to challenges with array temperature uniformity, accuracy, and the inability to rapidly cool. Complexity of integration with a 10,000 capillary array was also formidable, as the reported configuration [94] uses only four capillaries per solid-state heater.

2.2 Modeling

Many key design questions regarding thermal control of 10,000 capillary electrophoresis can be addressed using heat transfer and fluid dynamics theory before experimental efforts are undertaken. The results of modeling efforts will be summarized here as they predict performance related to the functional requirements described above.

In our modeling efforts, we first sought to predict the expected steady-state spatial gradients both across the array and radially through the capillaries. Next, we modeled the transient response of the system to step inputs from the either Joule heating of the capillary array or a step temperature change in the heat transfer fluid. Lastly, we concerned ourselves with prediction of vortex shedding frequencies and the pressure drop across the array.

2.2.1 Steady-state spatial thermal gradients

The steady-state operating conditions of the capillary array in the heat exchanger were modeled to determine the spatial thermal gradients in the system. Conservation of energy modeling and finite element analysis were performed to estimate the spatial gradients across the capillary array–to understand the total temperature change from the inlet to the outlet in the heat transfer fluid. To model the steady-state temperature difference between the gel sieving matrix and the heat transfer fluid within a single capillary, a heat exchanger analysis and finite element analysis were

undertaken.

Capillary array (Inter-capillary) gradients

At steady-state, conservation of energy requires that the heat transfer fluid absorbs all heat dissipated by the capillaries due to Joule heating. This can be stated as

$$\dot{E} = 0 = NAq'' - \rho Q c_p (T_{\text{out}} - T_{\text{in}}), \tag{2.2}$$

where NAq'' is the heat output of the capillary array and $\rho Q c_p \Delta T$ is the fluid temperature change. N is the number of capillaries, A is the outer surface area of each capillary, q'' is the heat flux per capillary from Equation 2.1, ρ is the density of the heat transfer fluid, Q is its volumetric flow rate, c_p is its heat capacity, and T is its temperature at the respective inlet and outlet to the capillary array. Using this equation, water flowing across the array at 20 L/min (3.3×10^{-4} m^3/s) will have ΔT=0.2 °C.[†]

A finite element analysis was performed using ANSYS (ANSYS, Inc., Canonsburg, PA) as well to more accurately model the array and fluid interaction. Only five rows were used in the simulation to minimize computational complexity. In the model, water enters the capillary array at 50 °C and 3.3×10^{-4} m^3/s, which corresponds to fluid velocity of 0.11 m/s. The capillaries are fused silica, the polyimide coating has been ignored, and the gel inside the capillaries is assumed to have the same properties as water. The results of the simulation are shown in Figure 2-3. From the results, we see that the water exits the array with ΔT=0.237 °C. This change is relatively linear along the row, increasing because the water is heated as it flows over the capillaries. Thus, the two modeling approaches to understanding capillary array thermal gradients are consistent and within the 0.4 °C functional requirement.

[†]Properties of water were evaluated at 55 °C, a nominal operating temperature.

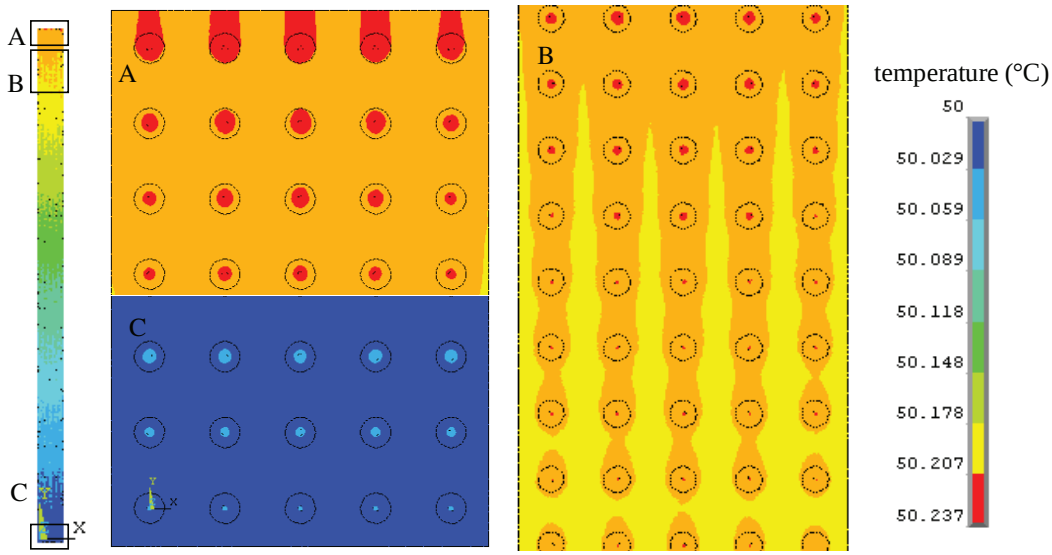

Figure 2-3: Simulated steady-state temperature profile across the capillary array. The full model is on the left, several sections have been magnified in the center, and the contour scale is on the right.

Intra-capillary gradients

Since the DNA separations occur within the gel sieving matrix, the intra-capillary thermal gradients can be detrimental. Gradients will exist because the heat generated from electrophoresis must be conducted to the surface of the capillary before it can undergo convection to the heat transfer fluid. Using the same simulation as described in the previous Section, the results for the intra-capillary thermal gradient are shown in Figure 2-4. In this result, the average temperature of the gel sieving matrix is 0.02 °C warmer than the temperature of the water at the capillary wall and 0.04 °C warmer than the at the midpoint between two adjacent capillaries. Using a second analysis, closed-form analytical modeling of the heat exchanger as a tube bank, the temperature difference between the gel and water is negligible. This is because the heat transfer rate from the gel to the fluid is one to two orders of magnitude larger than the heat flux from electrophoresis. For both analyses then, these gradients appears to be negligible.

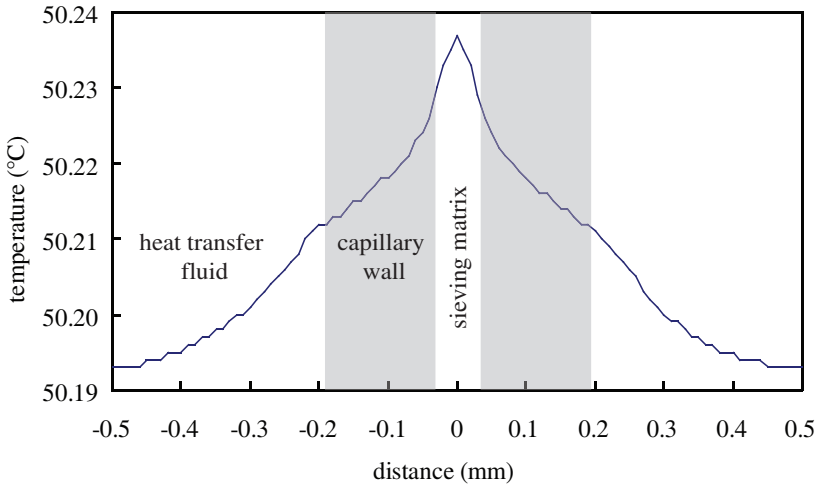

Figure 2-4: Simulated temperature profile across a capillary and surrounding heat transfer fluid.

2.2.2 Transient response

Modeling was also performed to study the transient thermal responses of the capillary array to system step inputs. As electrophoresis in the capillary array is initiated, there will be a step input in Joule heating, with an associated thermal response. In addition, if the heat transfer fluid temperature is stepped (e.g., from 50 °C to 60 °C), as may be desirable during operation, another transient response will occur.

A lumped parameter model utilizing control volumes was derived to understand the expected behavior of the capillary array to these changes. Two modes of heat transfer were included: convection between the capillaries and the fluid and constant heat generation in the capillaries. This modeling neglects the conduction to the fluid conduit, which proved substantial in experimental results described below. Assumptions about water flowing at 20 L/min (3.3×10^{-4} m^3/s) were again imposed. Results are shown in Figure 2-5. For both scenarios, the settling time for the capillaries to reach steady state is about 1 s. They increase in temperature until the rate of convection matches the internal heat flux. As in Figure 2-3, the capillary's temperature is roughly proportional to its row number.

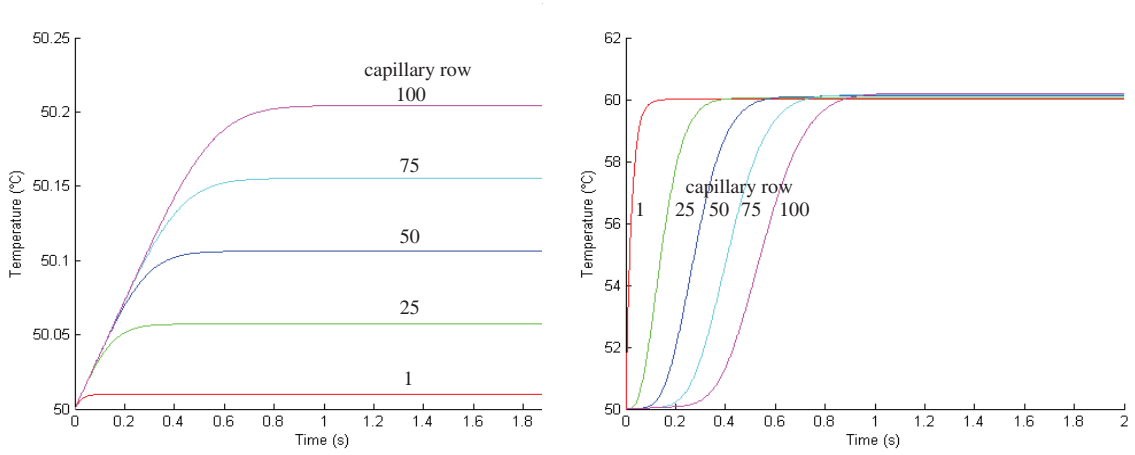

Figure 2-5: Modeled step responses to (*left*) electrophoresis with the fluid and capillaries initially at 50 °C and (*right*) change in fluid inlet temperature from 50 °C to 60 °C.

2.2.3 Pressure drop

As the heat transfer fluid passes through the array of capillaries, its pressure will drop. This pressure differential will cause the capillaries to deflect. Deflection will induce stress in the capillaries, which could potentially shear them from their constraints if they are not well supported. Using a variety of methods, this pressure was estimated and the resultant shear stress and the capillary deflection were computed.

The pressure drop in a tube bank can be calculated according to the derivation by Zhukauskas and Ulinsakas [95] as

$$\Delta P = N\chi\left(\frac{\rho V_{\max}^2}{2}\right) f, \tag{2.3}$$

where N is the number of rows of capillaries, χ is the correction factor, ρ is the fluid density, f is the friction factor, and $V_{\max}$ is the maximum velocity of the fluid, as given by

$$\frac{V_{\max}}{V} = \frac{S_T}{S_T - d_o}. \tag{2.4}$$

V is the mean fluid velocity, d_o is the capillary outer diameter, and S_T is the intercapillary spacing, 1 mm. The friction factor, f, is extrapolated as a function of the Reynolds number and longitudinal pitch, to be 2.48. The factor χ is 1 for an

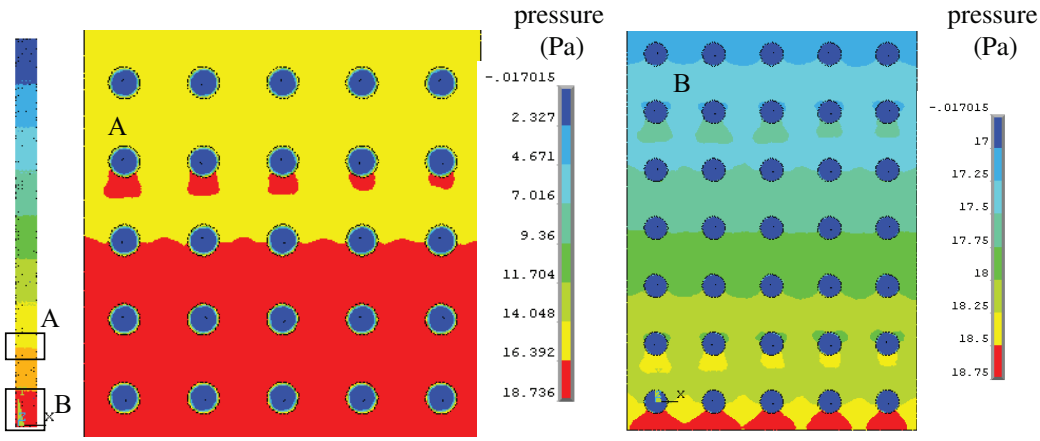

Figure 2-6: Simulated steady-state fluid pressure profile across the capillary array. The full model on the left and Section A are shown with the same contour scale. Contours in Section B are more closely spaced.

aligned tube bank with equal transverse and longitudinal spacing. For V=0.11 m/s corresponding to 20 L/min flow rate, $\Delta P = 26$ Pa over the entire capillary array.

Alternatively, one can calculate the pressure drop from the drag force on the capillary as the fluid moves past it. Accordingly,

$$C_D = \frac{\boldsymbol{\mathcal{D}}}{\frac{1}{2}\rho V^2 A}, \tag{2.5}$$

where C_D is a dimensionless parameter that is a function of the Reynolds number among other variables. The quantity $\frac{\boldsymbol{\mathcal{D}}}{A}$ is effectively the pressure on the capillary cross-section. The Reynolds number is given by

$$\mathrm{Re} = \frac{\rho V d_o}{\mu}, \tag{2.6}$$

where μ is the viscosity. For our V=0.11 m/s, Re=9, thus the flow is entirely laminar. At Re=9 for flow past a cylinder, C_D=3.5 [96]. Solving Equation 2.5 for pressure $\frac{\boldsymbol{\mathcal{D}}}{A}$ gives 0.21 Pa per capillary, or 21 Pa for the entire array.

Finite element analysis was also used to estimate the pressure drop across the capillary array. The results are shown in Figure 2-6. The pressure gradient for the entire array is 18.55 Pa according to this analysis.

So the results are consistent, 26 Pa for the tube bank analysis, 21 Pa for the drag equation, and 18.6 Pa for the finite element analysis. From this pressure we can calculate the maximum shear force on the capillary due to the cross-flow. This is

$$R = \frac{1}{2} P L d_o \tag{2.7}$$

where L is the length of the capillary. Using the median pressure value of $P = 21$ Pa, $R = 11\ \mu N$.

This shear force induces a maximum shear stress, τ, which is given by

$$\tau = \frac{2R}{A}, \tag{2.8}$$

where A is the cross sectional area given by

$$A = \pi \left(d_o^2 - d_i^2 \right). \tag{2.9}$$

Using Equation 2.8, $\tau = 225$ Pa, which is far below the yield stress of fused silica which is 5.88×10^9 Pa [97]. So there is absolutely no risk of the capillaries begin sheared off by drag forces from the heat transfer fluid flow. The maximum deflection of the capillaries under this flow, as simply supported beams, is given by

$$\delta_{\max} = \frac{5 P L^5}{384 E I}, \tag{2.10}$$

where E is the Young's modulus equal to 66×10^9 Pa and I is the moment of inertia given by

$$I = \pi r^3 t, \tag{2.11}$$

where $r = \frac{d_o + d_i}{4}$ and $t = \frac{d_o - d_i}{2}$. Under a pressure of 0.21 Pa per capillary, the maximum deflection δ is 200 μm. Since the capillaries are spaced on a 1,000 μm pitch, and have diameter, d_o=360 μm, there is no chance that they will touch each other.

2.2.4 Vortex shedding

Vortex shedding is a phenomenon in which a bluff body obstructing fluid flow leaves eddies in its wake [98]. If the capillaries have natural frequencies matched to the vortex shedding natural frequency, resonance could lead to fracture.

According to previous studies [98, 99, 100], vortex shedding can only occur above critical Reynolds numbers of 49, 40, and 47, respectively. Below such critical Reynolds numbers, the wake of a bluff body doe not develop into periodic instabilities [98]. Instead, the flow is symmetric with steady recirculation regions. In our modeling thus far, a volumetric flow rate of 20 L/min (3.3×10^{-4} m^3/s) results in a Reynolds number of 8.6 (see Equation 2.6. Thus, there is no possibility of vortex shedding. Were the flow rate to be increased to 117 L/min (0.0019 m^3/s) however, the capillaries would generate vortices at f_v=20 Hz according to

$$f_v = \frac{St \cdot V}{d_o}, \tag{2.12}$$

where St is the Strouhal number, given by Williamson and Brown [101] as

$$St = A + \frac{B}{\sqrt{\mathrm{Re}}} + \frac{C}{\mathrm{Re}}, \tag{2.13}$$

with constants A, B, and C from the literature [101]. This excitation frequency is near the capillaries' fundamental frequencies of 10 Hz and 23 Hz for pinned-pinned and clamped-clamped end conditions, as derived by Beards [102]. If extremely high volumetric flow rates are used, this should be considered.

Finite element analysis was also performed to study the fluid velocities around the capillaries. The results are shown in Figure 2-7. The fluid moves around the capillaries slowly; there is no vortex shedding or recirculation region in the wake of the capillaries.

To further reassure ourselves about the pressure drop and vortex shedding phenomena which could be disastrous to the costly and fragile 10,000 capillary array, a test rig was built in which to observe the effects of a diffuse cross-flow of water on a

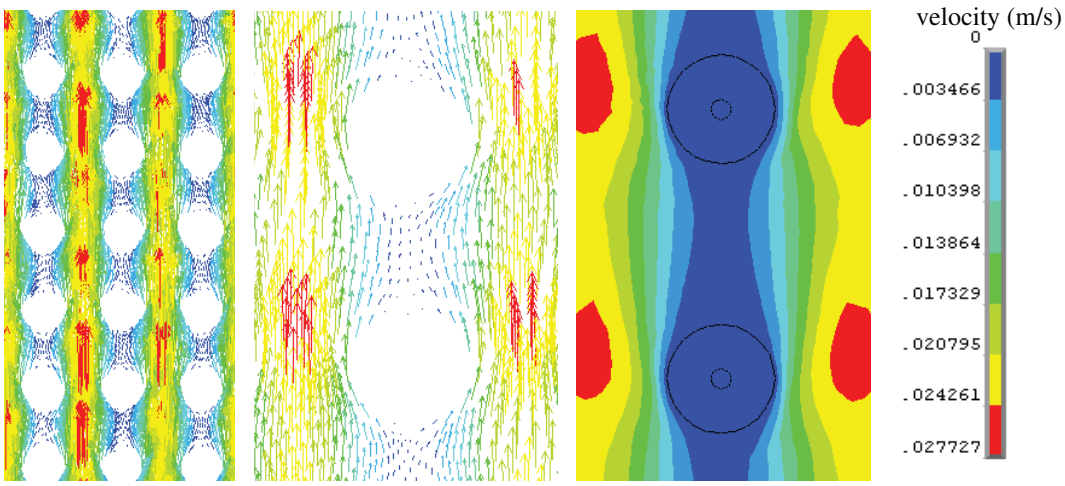

Figure 2-7: Simulated steady-state velocity profile of the heat transfer fluid, water, around the capillary array. Vector plots (*left*) and contour plot (*right*).

row of capillaries. The test rig under flow conditions is shown in Figure 2-8. A row of ten capillaries is installed and tensioned for testing. A hose delivers water from left to right. Under the approximately 0.1 m/s flow velocity, no deflection or vibration was observed. Notably, the capillaries do bundle together when drying after testing due to surface tension. Additionally, the capillaries for this experiment were tensioned as is expected in the ultra-high throughput mutational spectrometer.

2.3 Design and manufacturing

The modeling work culminated in the design of a cross-flow, water-cooled heat exchanger. The device is shown in Figure 2-9. Six modular aluminum segments, anodized for corrosion resistance and sealed with gaskets, are stacked to form a flow enclosure. The device has three chambers, separated by PMMA flow diffusers. In operation, water flows into the inlet chamber, laterally through the flow diffuser and across the capillary array, and into the outlet chamber. The device was designed for flexibility. In addition to the immersion heater (Omega, RI Series) shown installed in Figure 2-9, it can accommodate capillary arrays ranging from 1-10,000 capillaries. All adapters, interfaces, and gasket seals have been designed and manufactured. The 100 capillary array is shown installed in the heat exchanger and sealed in Figure 2-10.

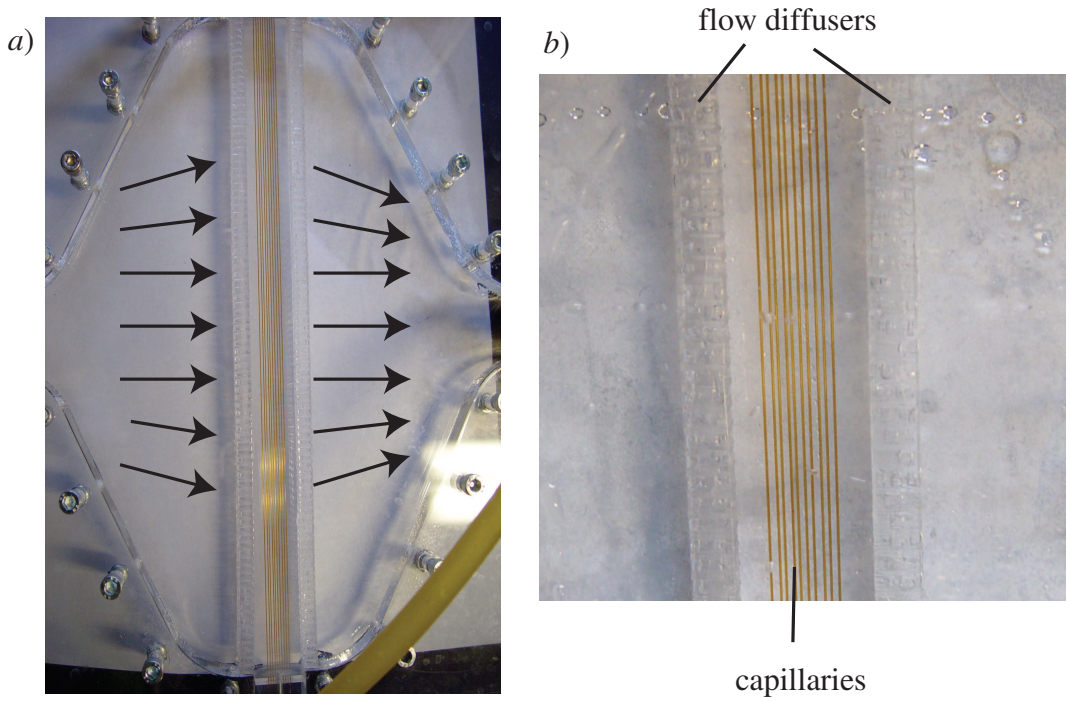

Figure 2-8: (*a*) Apparatus for observing effects of diffuse cross-flowing water on a row of capillaries. (*b*) Close-up of the apparatus and row of capillaries experiencing water flow velocity of approximately 0.1 m/s. No deflection or vibration is visible.

We subsequently replaced the aluminum lid of the enclosure with the polymer Delrin to electrically isolated the upper end of the capillary array. More details about the design of the capillary array itself are in Chapter 6.

An external water recirculator (PolyScience, Model 9601) controls the temperature (−25 °C to 150 °C) and flow rate (0.28×10^{-4}–3.3×10^{-4} m^3/s). An array of seven resistive temperature devices (RTDs) were installed in ports in the stacked segments to provide axial and lateral thermal sensing capability.

2.4 Experimental results and discussion

A suite of experiments were performed on the thermal control system to verify its ability to meet the functional requirements described in Section 2.1.1.

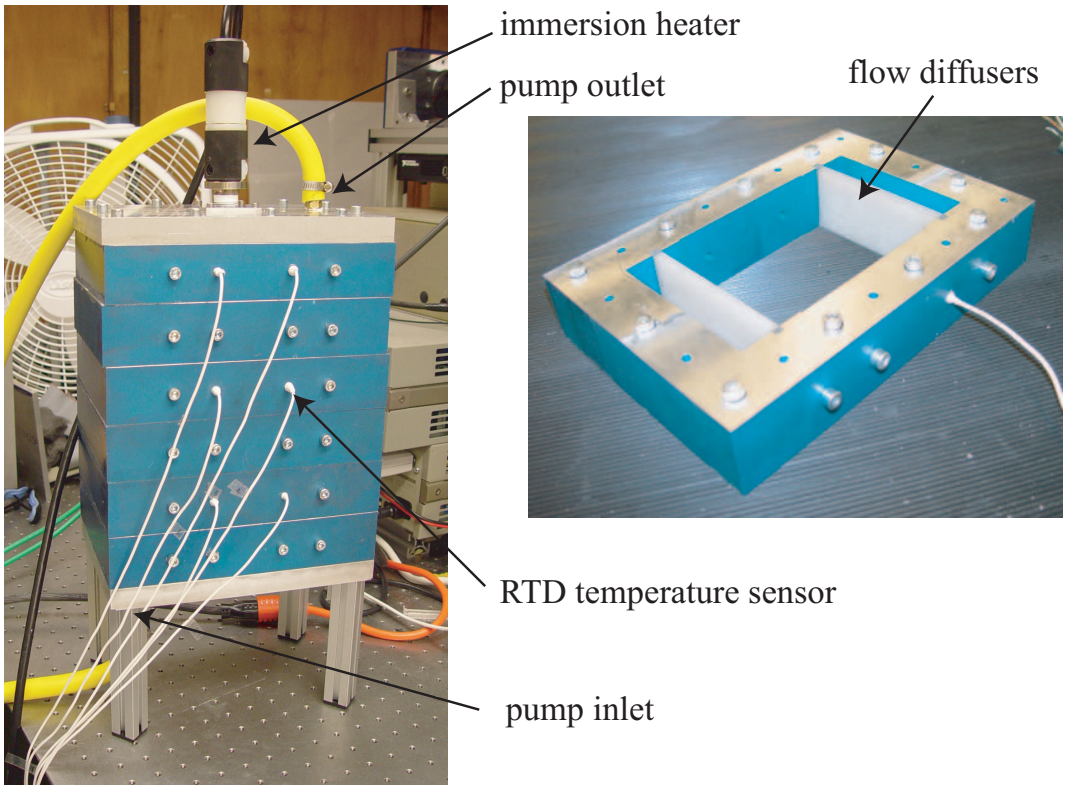

Figure 2-9: Heat exchanger for the capillary array. Six 50 mm thick segments (one shown alone) are stacked to form the device. An immersion heater is shown installed instead of the capillary array for testing.

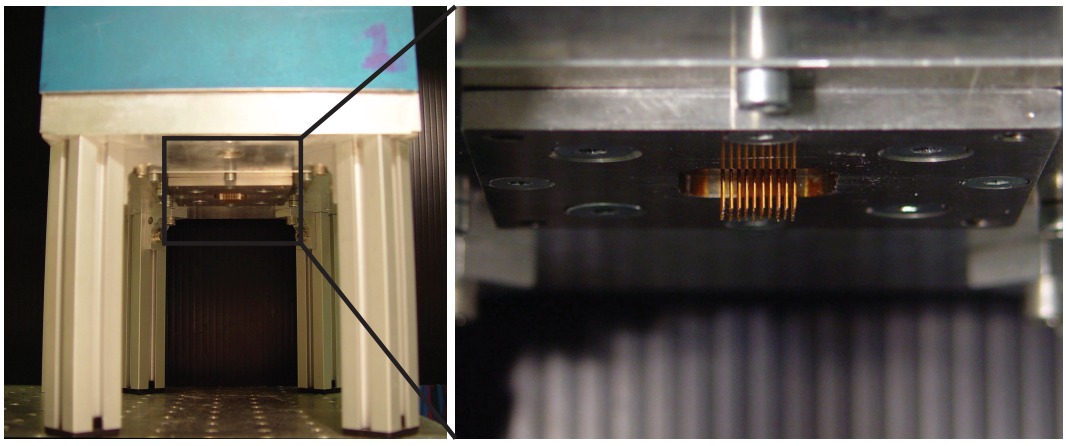

Figure 2-10: The 100 capillary array is shown installed in the heat exchanger (*left*) (close-up (*right*)) and sealed from water pressure inside.

2.4.1 Experimental setup

The experimental setup for these experiments is shown in Figure 2-9. The immersion heater was used to input 255 W at 70 V DC to the heat exchanger to simulate the heat output of the 10,000 capillary array.

The RTDs, sampling at ∼2 Hz, were calibrated by taking average measurements on all seven RTDs at various temperatures while they were immersed in the external water recirculator bath. The temperature on the digital display of the recirculator was not used in these experiments, except as a rough guide for initially setting the temperature. For the experiments, one calibrated RTD was placed in the recirculator's bath while the other six were installed in the heat exchanger's ports.

Temperatures were varied from 50-60 °C and flow velocities were varied from 0.67×10^{-4}–3.3×10^{-4} m/s (flow rate of 4-20 L/min). These temperatures are useful for DNA separation in accordance with Figure 2-1 and the flow rates were chosen in accordance with the models to meet functional requirements.

2.4.2 Steady-state temperature offset

For a recirculator water temperature of 50 °C, the water in the heat exchanger was 0.2 °C cooler. This steady-state offset is caused by convection from the water to the ambient air as it traverses the system. No axial thermal gradients were observed in the measurements.

2.4.3 Step response and steady-state thermal gradient

With the immersion heater activated to simulate Joule heating, the water in the heat exchanger rises to 0.24 °C above the recirculator temperature at a flow rate of 20 L/min (3.3×10^{-4} m/s). This response is shown in Figure 2-11. The rise time is approximately 350 s. This experiment also illustrates the expected steady-state thermal gradients in the water flowing across the capillary array. For a set temperature of 50 °C (in the recirculator), the water immediately upstream and downstream of the array will be 49.80 °C, and 50.24 °C, respectively. Thus, the gradient is 0.44 °C,

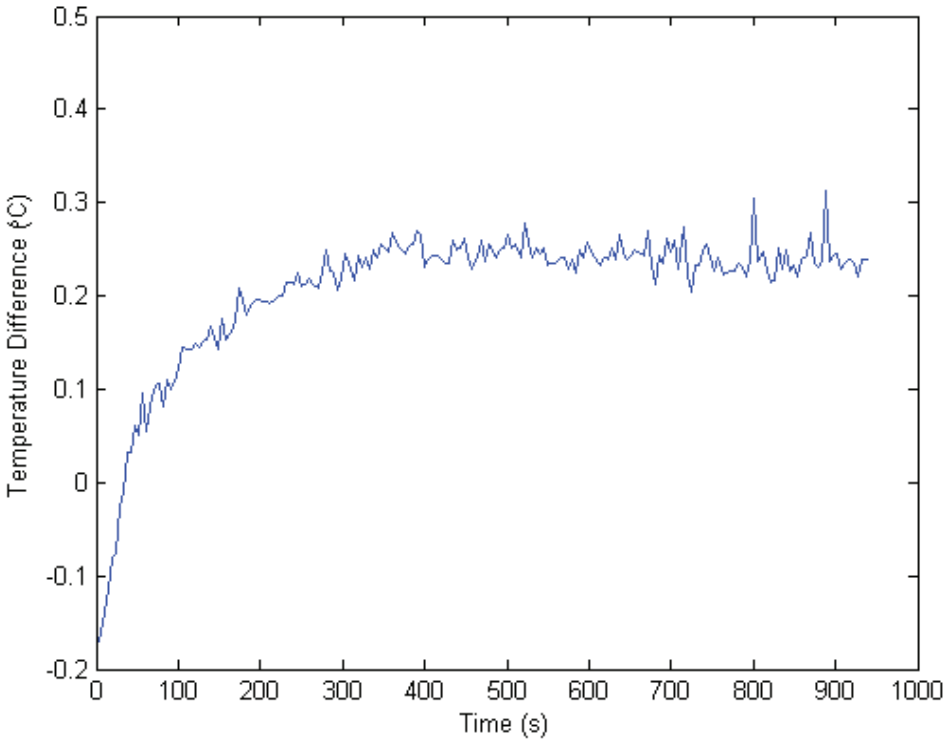

Figure 2-11: Experimental mean temperature response to a 255 W heat input in the heat exchanger, at a flow rate of 20 L/min. Data is shown relative to the temperature of the water in the recirculator bath, at 50 °C.

which nearly matches the functional requirement.

At slower flow rates, the gradient is larger: 0.63 °C at 7.5 L/min (1.3×10^{-4} m/s), and 0.85 °C at 4 L/min (0.67×10^{-4} m/s). Measured axial thermal gradients were negligible for all flow rates with the Joule heating simulation activated.

2.4.4 Trajectory following

The maximum heating and cooling rate of the recirculator was determined empirically by directly connecting the its inlet to outlet. These rates are 0.0083 °C/s heating and 0.0108 °C/s cooling, as governed by the recirculator's PID controller and heat transfer to the ambient environment. The recirculator required approximately 1200 s to change from 50 °C to 60 °C, with a 300 s settling time.

Simulation without Joule heating

While the inlet water temperature is increasing at the maximum rate, the heat exchanger water temperature is 0.4 °C offset lower than the reservoir temperature. When decreasing at the maximum rate, it is 0.34 °C offset higher than the reservoir temperature.. These offsets are achieved after a transient of 300 s.

Simulation with Joule heating

With the immersion heater outputting 255 W and the inlet water temperature increasing at the maximum rate, the heat exchanger water temperature is 0.06 °C offset lower than the reservoir temperature. When decreasing at the maximum rate, it is 0.67 °C offset higher than the reservoir temperature. These offsets are achieved after a transient of 300 s.

2.5 Discussion

We can compare the results of the modeling and experimental measurements with the stated functional requirements to evaluate the design. The external water recirculator

appears to have sufficient temperature range (−25 °C to 150 °C) and volumetric flow rates (up to 20 L/min) to meet the functional requirements. The steady-state temperature offset of the water in the heat exchanger of 0.2 °C relative to the water in the recirculator was not revealed in the modeling presented, but can be explained by convection to the ambient air as it traverses the system. This systematic error can be compensated for with a bias, so is not of a concern.

Experiments with the immersion heater to simulate capillary array Joule heating at 255 W can reveal much about the system step response and steady-state inter-capillary thermal gradients. After a 300 s rise time, the water in the heat exchanger reaches a steady-state gradient with a range of 0.44 °C at a flow rate of 20 L/min. This rise time suggests that DNA separations should not be performed during the first five minutes of electrophoresis in the capillary array. A capillary entry length of 50-100 mm would suffice for this purpose, as the typical DNA speed are 10-20 mm/min at 8 μA electrophoretic current. Secondly, the steady-state thermal gradient is marginally acceptable. Separations in the last row of the array would be possible, but would be slightly worse than those in the first row, assuming that the first row is set at the ideal, theoretical, melting temperature. This could be ameliorated by setting the middle row at the ideal temperature, or alternatively sweeping the temperature by a range of 0.44 °C during the run. The work of Minarik *et al.* [85, 86] further suggests that such temporal gradients can be practically implemented and even advantageous.

The intra-capillary radial gradients were modeled to range by, at most, 0.04 °C. Complexity precluded experimental verification, but the magnitude of these gradients is negligible. If they are, experimentally, an order of magnitude larger, then a bias compensation can be easily implemented.

Minarik *et al.* [85, 86] have shown that sinusoidal variations in temperature can improve the efficiency of separations by reducing the required accuracy and mitigating spatial thermal gradient concerns. From the trajectory following data presented, the minimum period for 2 °C amplitude temporal temperature oscillations would be 1,000 s. For 1 °C amplitude oscillations, the minimum period would be 600 s. Thus, during

a 20-40 minute run, 1.2-4 cycles can be performed. This is somewhat less than the typical sinusoidal temperature profiles used of 1.5 °C amplitude at 1/120 Hz. If the water supply for the heat exchanger were to be provided by two different recirculators operating at two different temperatures, electromechanical valves could easily switch the flow, and thus temperatures, to the heat exchanger. The modeling of the capillary array response to such a step input in temperature reveals that the total period for such cycling could be 5-10 s. Thus, several hundred temperature oscillations would be possible.

Vortex shedding and the pressure drop induced shear forces on the capillary array do not appear to be of a concern, as shown with extensive modeling and the experimental test rig. As a result, a free-standing capillary array, simply supported at its ends, will be implemented to maximize heat transfer to the water and simplify design and manufacturing of the capillary array.

Chapter 3

Gel loading device

Ultra-high throughput capillary array electrophoresis requires the integration of a suite of subsystems. In this work, we have developed a pneumatic device that can be interfaced to a capillary array for the purpose of filling the capillaries with a porous medium, a viscous gel matrix, for DNA mutation separation using electrophoresis, or washing the capillaries, or for flushing them with N_2. Indeed, we have loaded capillary arrays containing 1-100 capillaries with gels such as 3-5% linear poly(acrylamide) (LPA) and 6% polyvinylpyrrolidone (PVP) in several minutes at pressures of hundreds of kPa, and subsequently flushed and dried them. The design has been scaled and manufactured for 10,000 capillary array capability as well. Throughout, modeling has been performed, such as for the viscous polymer flow in the capillaries and the statistical distribution of capillary filling times. We also explored the potential for lowering the gel viscosity to reduce injection times and/or pressures by applying time varying shear stress or heat. This effective and rapid device for capillary array gel loading, washing, and drying can help enable ultra-high throughput capillary array electrophoresis instrumentation.

3.1 Introduction

Electrophoresis requires that the analytes (e.g., DNA) migrate through an (ideally) stationary medium. Where high resolution separations are required, as in DNA se-

quencing or mutation detection, that medium is commonly a polymerized gel matrix. Separations in slab gels are performed using a cross-linked matrix, while those in capillaries or channels use a replaceable non-cross-linked matrix.

Replaceable linear poly(acrylamide) (LPA) was introduced by Karger *et al.* in 1990 [103] and subsequently improved [104]. These media have been widely employed for DNA sequencing and other applications. For Karger's chemistry as well as another gel based on linear poly(ethylene oxide), capillary coatings that covalently link chains of the polymer to the silano groups of the silica are required before use [105, 106] to suppress the detrimental effects of electroosmotic flow [107]. Other gel chemistries, such as those based on polyvinylpyrrolidone (PVP) [108], and poly(dimethylacrylamide) [109], are "dynamically coating," suppressing electroosmotic flow by shielding [110, 111].

These polymerized gels are commonly replaced after each capillary run. In some instances, the capillaries are washed in between runs using an aqueous surfactant solution. In addition, capillaries can be dried by flowing gases through them. Detailed listings of the consumables (gel, buffer, wash) utilized in this research are included in Appendix A.

Following Karger's advancement in polymer gel chemistry, numerous methods of loading the solution into capillaries were devised. The simplest method is syringe injection. This is performed by filling a typically 100 μL syringe (SGE, Inc., Victoria, Australia) with the gel, connecting it to capillary with outer diameter d_o=360 μm using a d_o=350 μm teflon tube (Bodman Industries, Aston, PA), and applying pressure to the syringe by hand. In our experience, application of 2.5 N (0.5 lb) force to the syringe plunger generates 1.4 MPa (200 psi) for gel injection. This method is reliable, though cumbersome, for single capillaries, and is impractical for multi-capillary arrays.

Multi-capillary array instruments use either a hydraulic or pneumatic system to apply pressure to a reservoir of gel to force it into the capillary array. In the hydraulic system used by Spectrumedix, Inc. (State College, PA), a mechanical piston is driven into a syringe body by a lead screw to generate up to 14 MPa (2000 psi). In another

commercialized system with 96 capillaries, by Amersham (Piscataway, NJ), a N_2 tank pressurizes chambers where the ends of sets of 16 capillaries are dipped into gel matrix reservoirs.

Arrays of etched channels for electrophoresis have been filled with a similar system, for example by Scherer *et al.* [112], with several minor modifications such as a central port interface and use of helium for its low gas solubility. This system was rated at 7 MPa (1,000 psi).

The pneumatic gel loader concept appears adaptable to large arrays containing thousands of capillaries, but has never been implemented. Additionally, modeling can enable the prediction of gel loading performance, which has not be undertaken in previous research. Thus we set out to model and design gel loaders capable of 1-10,000 capillary capacities.

3.2 Design

3.2.1 Modeling

Fluid dynamics theory can be used to predict the time required to fill the capillaries with the viscous polymer matrix as a function of pressure and capillary length. The gel matrix is very viscous, like honey, so its flow through the capillary pipe is governed by a balance between viscous and pressure forces. As derived by J. Poiseuille, and restated by Munson *et al.* [96], the fully developed viscous flow with parabolic flow profile will be governed by

$$Q = \frac{\pi d_i^4 \Delta p}{128 \mu l}, \tag{3.1}$$

where Q is the volumetric flow rate, d_i is the capillary inner diameter, Δp is the pressure relative to atmospheric at one end of the capillary, μ is the gel viscosity, and l is the capillary length. The time, t required for a gel molecule to travel the length of the capillary will be simply v/Q, where v is the capillary volume, or equivalently

$$t = \frac{32 \mu l^2}{d_i^2 \Delta p}. \tag{3.2}$$

Equation 3.1 applies only to fully developed flow, so it is prudent to check that $l \gg l_e$, where l_e is the dimensionless entrance length given for laminar flow by

$$l_e = 0.06\text{Re } d_i. \tag{3.3}$$

Re is the Reynolds number, in this case given by

$$\text{Re} = \frac{\rho V l}{\mu} = \frac{4\rho Q l}{\pi d_i^2 \mu}, \tag{3.4}$$

where V is the average velocity in the capillary.

In our experiments, the capillaries are typically d_i=75 μm and 300<l (mm)<500. Applied pressures are in the range of 150-250 kPa. Gel viscosity is typically 8000 Pa·s, as described in the next Section. For these values, Re≪1 (laminar flow), so Equation 3.3 is valid and l_e<0.1 mm. Thus, the flow can be considered fully developed for the entire capillary length. Under these assumptions, t from Equation 3.2 is on the order of tens of seconds. This time, t, is inversely proportional to the applied pressure and varies with the capillary length squared.

3.2.2 Gel viscosity

The gel sieving matrix used for these experiments is an aqueous solution of ∼5% linear poly(acrylamide) (LPA) (Spectrumedix PCR-QC or provided by collaborator Weiming Zhang, see Appendix A), which has ρ ≈1,000 kg/m^3. Using a rheometer (TA Instruments, AR2000), we measured the gel viscosity as a function of shear rate (see Figure 3-1) and shear stress (see Figure 3-2) for several temperatures in the range of 25-35 °C.

As can be seen in Figure 3-1, the apparent viscosity decreases with increasing shear rate. This type of fluid is known as shear thinning, and is non-Newtonian behavior. Most likely, as the polymer chains align under high shear rates, the fluid imposes less resistance to shear. From our data, four orders of magnitude reduction in the viscosity were observed at 1,000 Hz shear rate as compared to the viscosity near the

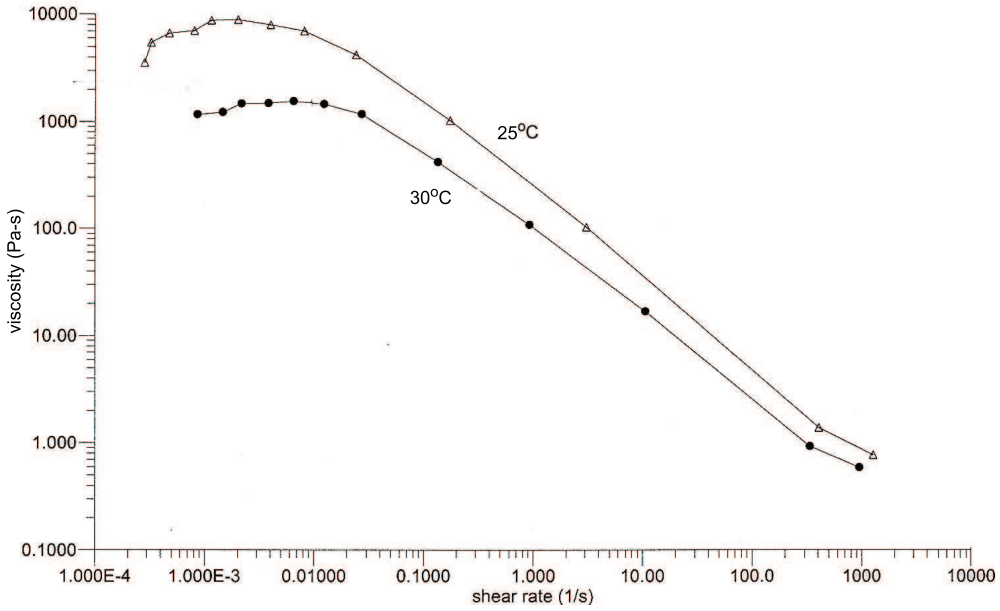

Figure 3-1: Viscosity of ∼5% linear polyacrylamide (LPA) gel as a function of shear rate. Shown for 25 °C and 30 °C.

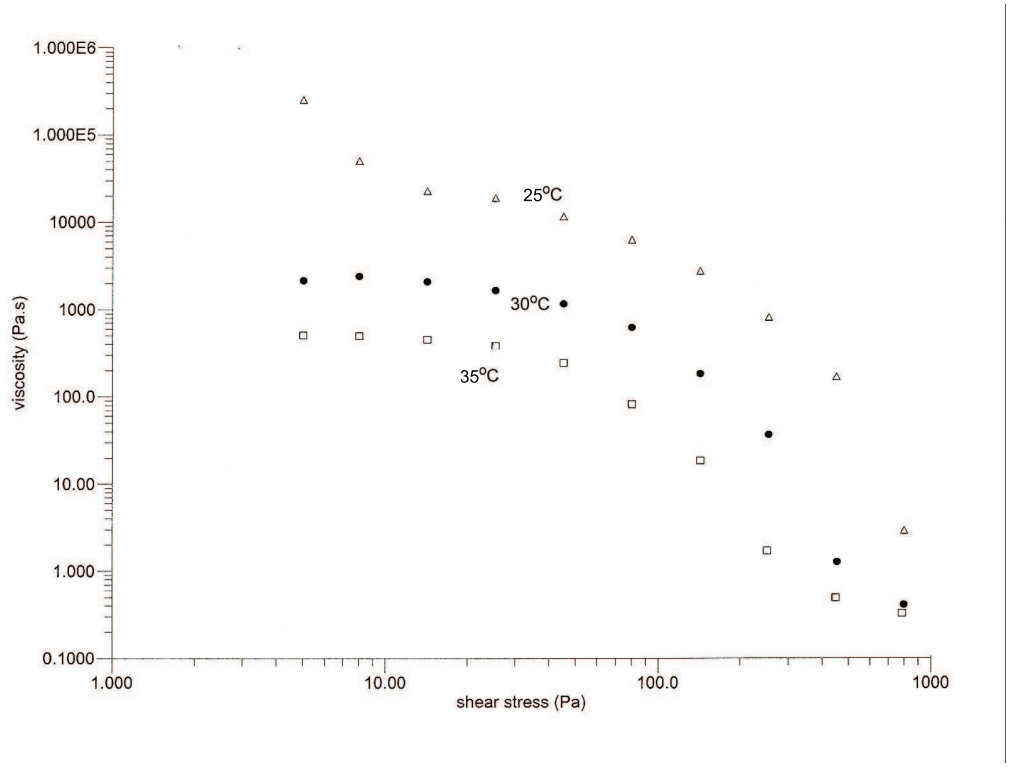

Figure 3-2: Viscosity of ∼5% linear polyacrylamide (LPA) gel as a function of shear stress. Shown for 25 °C-35 °C.

quasi-static shear rate. Additionally, the viscosity is highly sensitive to temperature (see Figure 3-2). A 10 °C increase in temperature can lower the viscosity by two orders of magnitude. The viscosity at room temperature and low shear rate, from Figure 3-1, is about 8000 Pa·s. These results are consistent with other reported values for viscosity versus shear rate [113, 114] and temperature [114, 115].

This data begs some interesting questions about the design of a gel loader. Since the time to fill the capillary is proportional to the viscosity, this time could potentially be lowered by heating the gel or shearing it during loading.

High frequency gel loader

We designed and constructed bench-level experiment to attempt to induce high shear rate viscosity reduction in the gel. In the design, shown in Figure 3-3, an acoustic speaker is coupled to a water chamber to apply driving frequencies of 50-1,000 Hz to the capillary's immediate environment. The gel is loaded into single capillary with a syringe driven by a linear motor while the speaker is resonating. A force sensor between the syringe and linear motor enable monitoring of the pressure required to load the gel into the capillary. The capillary, mounted to the syringe, passes through the water chamber.

Unfortunately, no reduction in required syringe force was observed for the range of frequencies tested. The speed sound in water is s=1500 m/s. The wavelength of the acoustic wave at this f=1,000 Hz is given by

$$\lambda = \frac{s}{f}. \tag{3.5}$$

So λ=1.5 m$\gg d_o$. A frequency of 10 MHz would generate wavelengths of $2d_i$, and may be more effective. Also the capillary is a thick walled cylinder, with d_o=360 μm. A wall thickness of 10 μm would be far more compressible than the 137.5 μm used in this experiment. Temperature seems to be a far simpler way of reducing the viscosity of the gel, so provisions were made in the design of the gel loading device to apply isothermal heating.

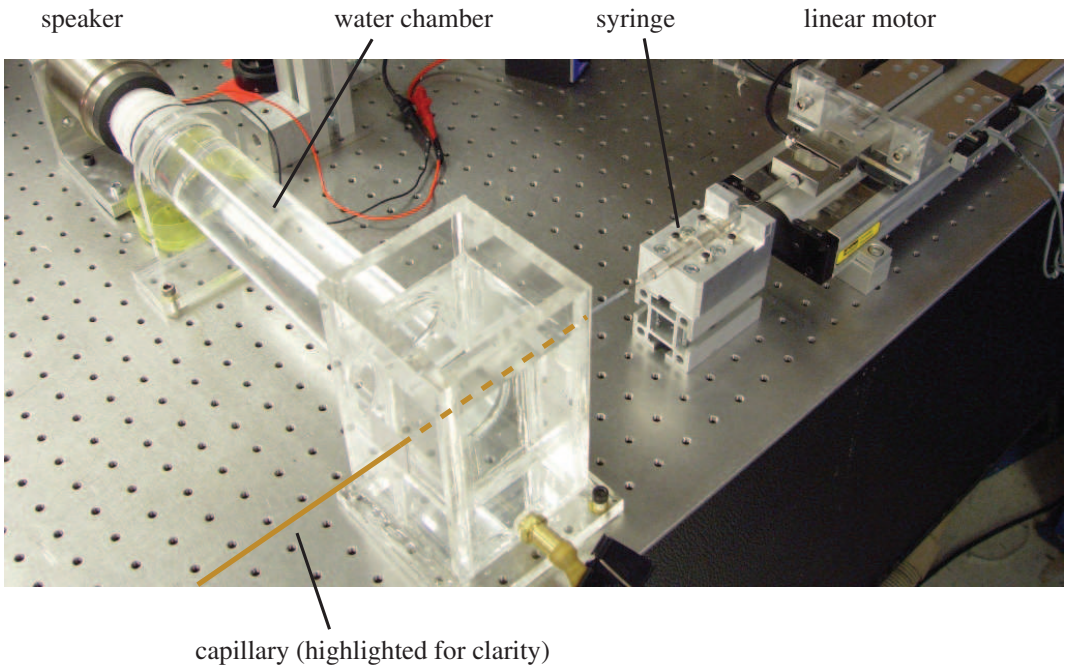

Figure 3-3: Experimental setup for 50-1,000 Hz vibration of the capillary in an aqueous environment to induce shear thinning viscosity reduction.

3.2.3 Device hardware

Numerous gel loaders were designed and manufactured. They are all based on the same design concept, illustrated in Figure 3-4. In the design, a gel-filled container–either an Eppendorf tube or an aluminum tray–is inserted into an aluminum chamber. A lid is then fastened onto the chamber and sealed with either an o-ring or a silicone gasket. The lid contains a number of holes through which capillaries can be inserted either before lid installation or after. A valve is then opened to admit pressurized N_2, ranging from 150 kPa (22 psi) to 4200 kPa (600 psi) through NPT fittings. The pressure forces the gel up through the capillary (array). After the valve is closed, a relief valve can be opened.

To wash or dry the capillaries, the Eppendorf tubes or trays can be alternatively filled with a wash solution (e.g., mild surfactant in water), or N_2 to dry the capillaries.

We have manufactured and tested a variety of gel loaders. Shown in Figs. 3-5 though 3-8 are gel loaders with respective capacities of 1, 80, 100, and 10,000 capillaries.

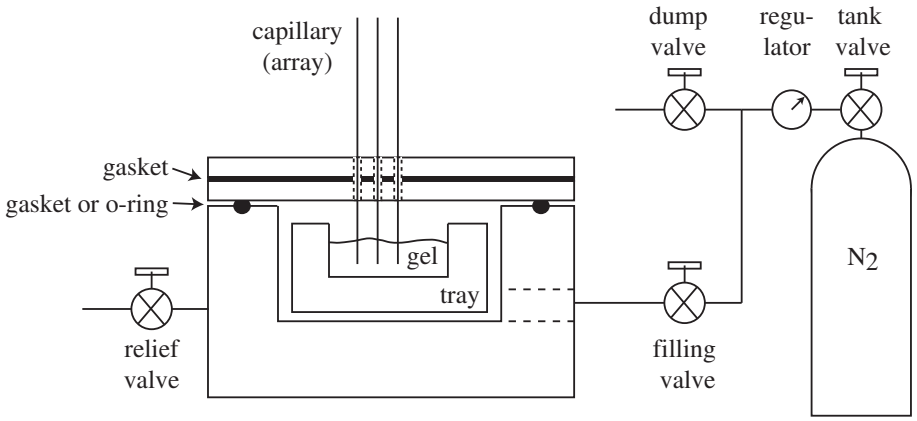

Figure 3-4: Schematic of the gel loading device. A tank deliveres regulated N_2 to an enclosed chamber, where a tray pre-filled with gel sieving matrix is located. Capillaries can enter the chamber through a silicone gasket. This design has been tested in a variety of conditions, from 1 capillary capability up to 4200 kPa (600 psi), to 10,000 capillary capability up to 35 kPa (50 psi).

The designs feature a large aluminum body that can be used to apply isothermal control of the gel to lower viscosity. The tray is filled with gel amply so that their tips remain immersed after gel loading. Details of the respective designs are included in Figure captions.

3.2.4 Gel filling time measurement apparatus

To evaluate the performance of the gel loaders and enable comparison between empirical gel loading and theory, we designed and built an instrument for measuring the time required for gel to traverse the capillaries under pressure, as shown in Figure 3-9. In the instrument, gel is force up through the capillaries by the gel loader. As it reaches the capillary exit, the gel completes an electrical circuit. BNC cables on the circuit board (see Figure 3-9 *right*) enable autonomous electronic readout and control of a solenoid filling valve on the gel loader.

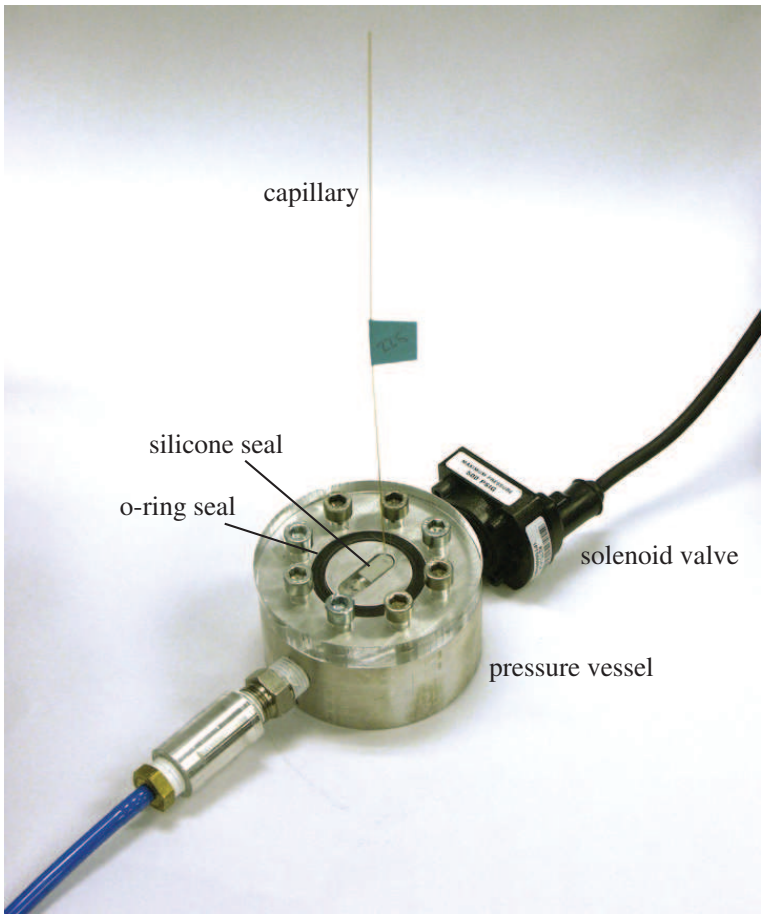

Figure 3-5: Single capillary gel loader. A hole in the polymethyl-methacrylate (PMMA) lid admits a single capillary through the silicone seal and into a gel-filled Eppendorf tube. The solenoid valve can be used to remotely trigger loading. This device has been tested up to 7 MPa (1,000 psi).

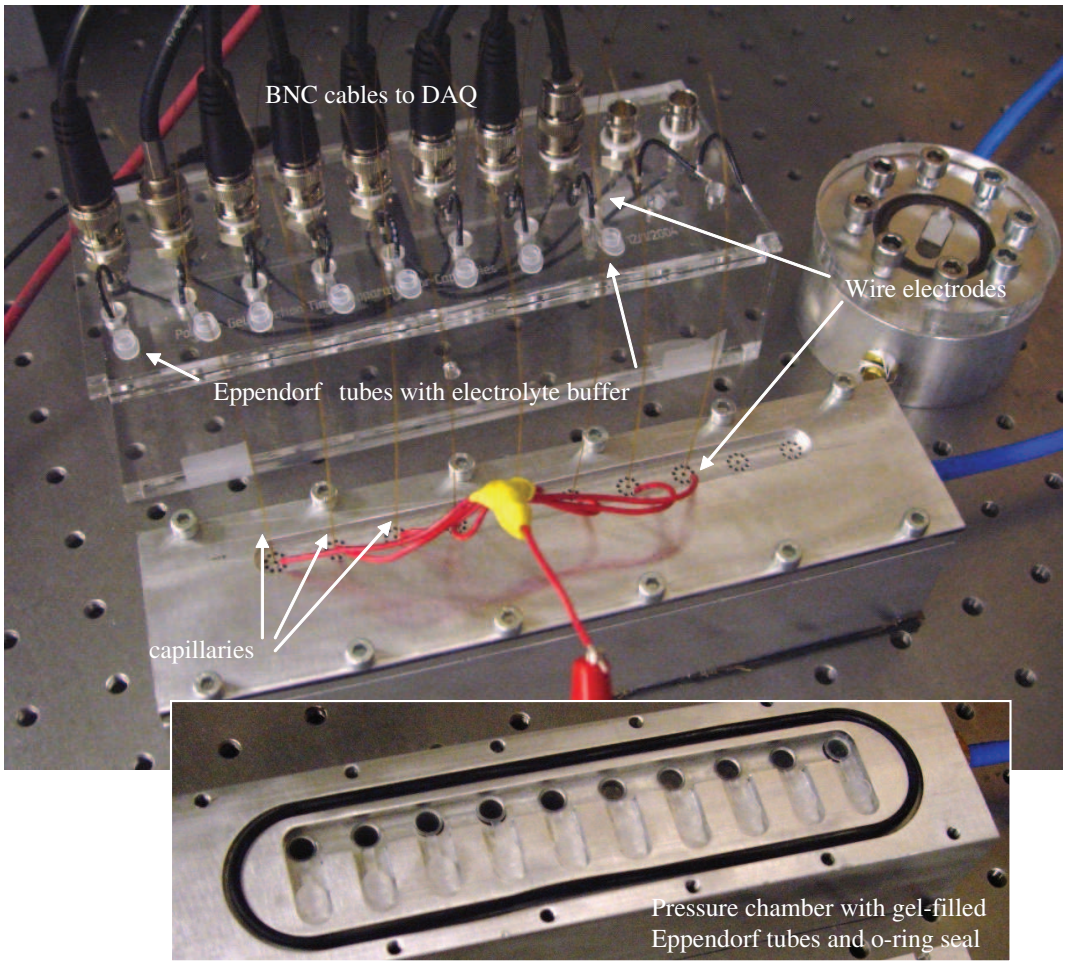

Figure 3-6: Eighty capillary gel loader. Ten independent Eppendorf tubes containing gel are loaded into the chamber (*bottom*). Through holes in the lid, eight capillaries per tube can be inserted. Ten independent electrical connections can be monitored via a data acquisition board (DAQ) to measure gel filling times. This device has been tested up to 350 kPa (50 psi).

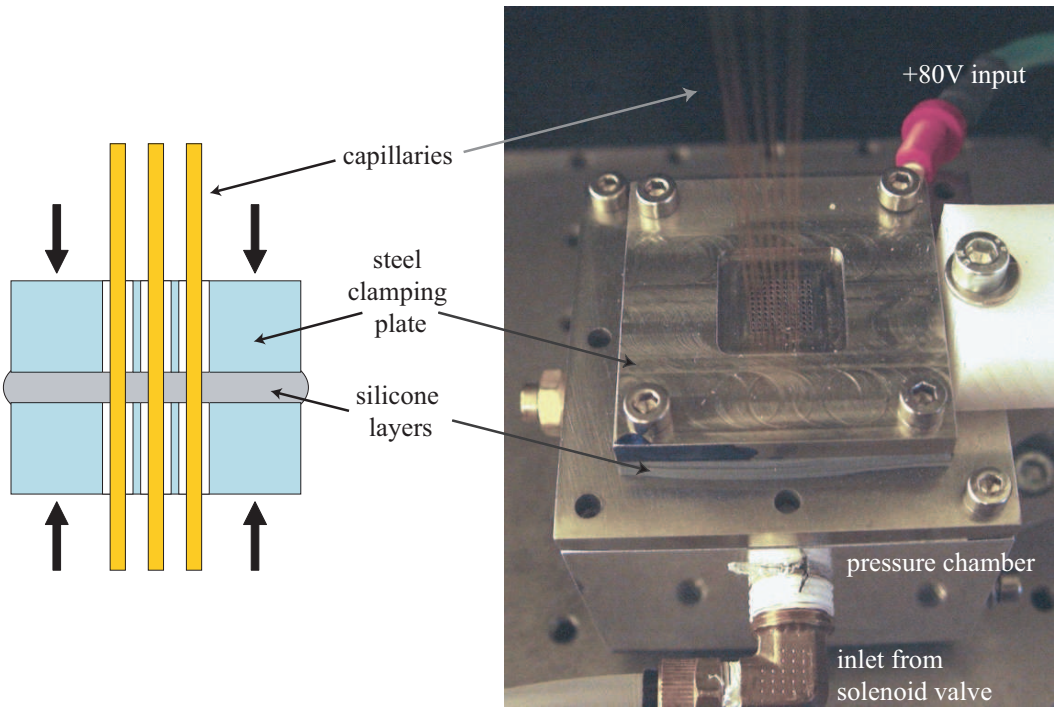

Figure 3-7: One-hundred capillary gel loader mated to capillary constraint device. The constraint device sandwich of steel-silicone-steel (*left*) is affixed to the 100 capillary chamber (*right*). The chamber contains a tray of gel. An electrode can be attached to measure gel filling times. This device has been tested up to 350 kPa (50 psi) to load 100 capillary arrays.

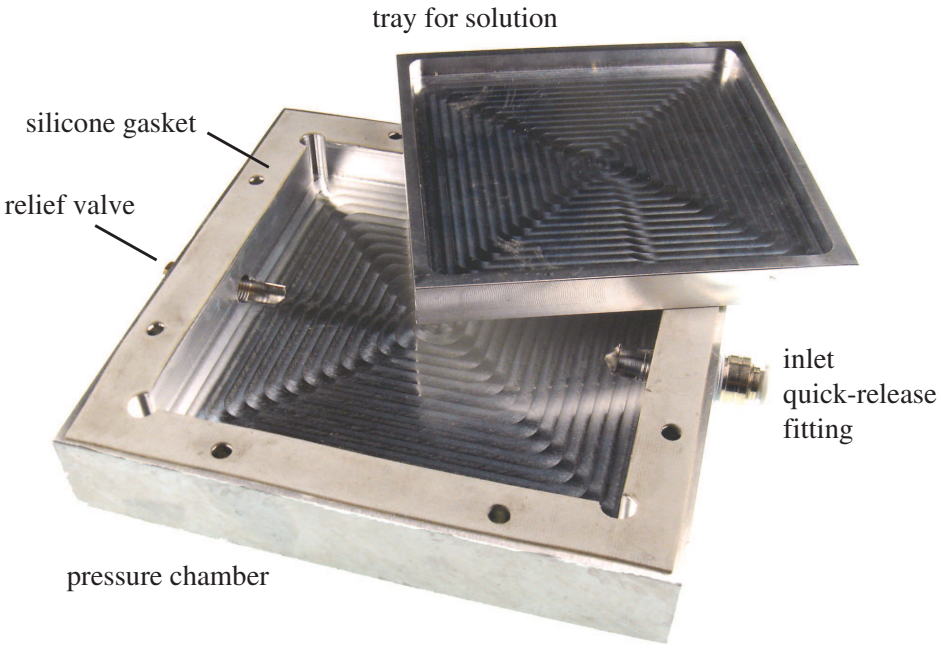

Figure 3-8: Ten-thousand capillary gel loader. The device is designed to interface to the 10,000 capillary constraint device. As the pressure chamber is affixed to the constraint device, all capillary tips enter the gel solution in the tray. This device has been tested up to 350 kPa (50 psi), albeit without capillaries inserted.

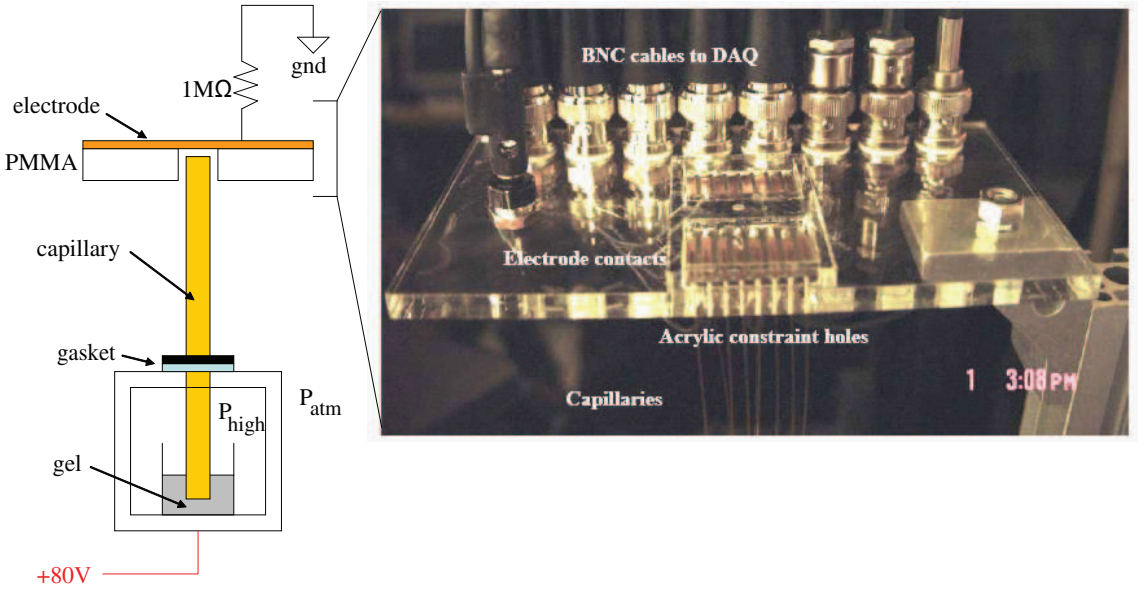

Figure 3-9: An instrument was designed (*left*) and built (*right*) to enable measurement of the times required to fill capillaries with gel using the gel loaders. The device measures 8 channels independently, by triggering on electrical connectivity between the gel loading device and electrodes at the opposite end of the capillary array.

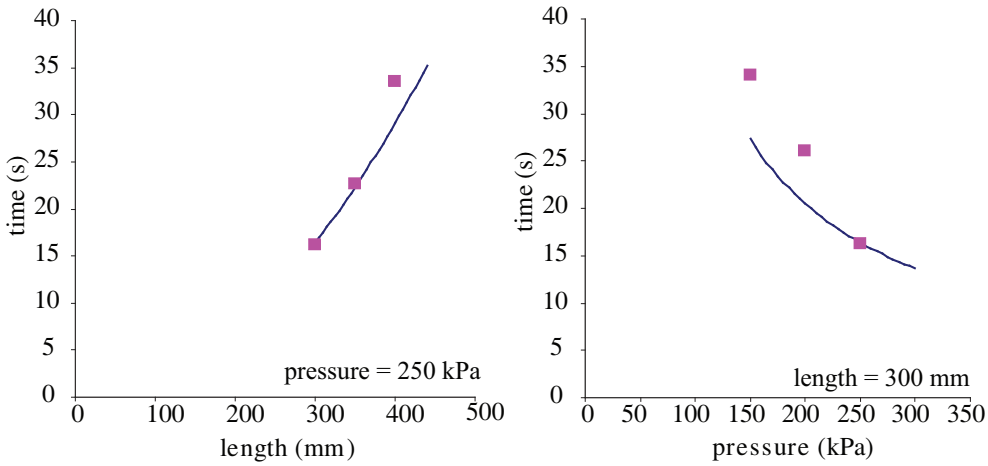

Figure 3-10: Experimental (points) and theoretical (line) gel filling times for 72 capillaries under 5 different parameter sets (one set is repeated between the two plots). d_i=75 μm, 300<l (mm)<500, 150<Δp (kPa)<250, μ=8000 Pa·s.

3.3 Results and discussion

3.3.1 Modeling and experimental data comparison

Using the 100 capillary gel loader with only eight capillaries inserted into a row, we measured the time required to fill each of them 45 times, for a total of 360 capillaries filling measurements. These measurements were organized as nine repeated runs, under five different conditions, each with eight capillaries. The 9×8=72 runs were averaged for each condition, and the results are shown along with the viscous flow model in Figure 3-10.

All 360 capillaries were successfully loaded during these experiments. The viscous flow model compares favorably with the experimental data, and can be used to reliably predict the filling time for parameters not tested. Filling time is proportional to viscosity and length squared, and inversely proportional to pressure and diameter squared, as indicated by Equation 3.2.

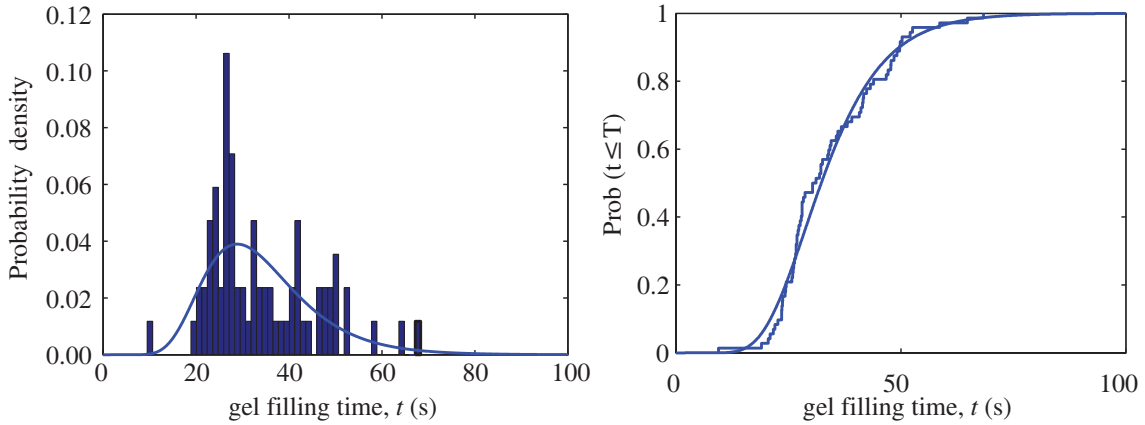

Figure 3-11: Probability density function (pdf) (*left*) and cumulative distribution function (cdf) (*right*) of gel filling times for capillaries with d_i=75 μm, l=300 mm, Δp=150 kPa, μ=8000 Pa·s, 72 trials, along with log-normal distribution model.

3.3.2 Statistical distribution

These 72 experiments per condition also provide an opportunity to asses the statistical distribution of capillary filling times, where the filling time is a random variable. For such a distributed random variable, one can define the probability density function (pdf). Like a smoothed out histogram, the pdf gives the probability that the value of the random variable will fall between given bounds. The cumulative distribution function (cdf) can also be defined for our distribution, as the probability that the random variable takes on a value less than some threshold.

For one parameter set (l=300 mm, Δp=150 kPa), the distribution of 72 filling times, or pdf, is shown along with the cdf in Figure 3-11. In this Figure we include modeled fits which shall be explained shortly. The pdf appears to be right-skewed (non-normal). That is, the median (30.8 s) is not equal to the mean (34.1 s) or mode (27 s). Fitting a model to this histogram for the pressure and length conditions can enable prediction of filling times for other parameters (e.g., untested pressure, length, viscosities, or numbers of capillaries).

Data of this distribution (right-skewed) can be caused by several factors. There is a lower bound (i.e., zero seconds) for the capillary filling times. Skewness can also result from start-up effects. Since no failures were observed in any of 360 fillings, our process could be characterized as having a long start-up period where failures are rare

resulting in right-skewed data, as in characterizing reliability data) [116].

There are numerous model families to fit a right-skewed distribution, such as the Weibull, gamma, chi-square, log-normal, and power log-normal. All of the preceding models were fit the data and the log-normal distribution was determined to be the best fitting.

The log-normal distribution is the probability distribution of a random variable whose logarithm is normally distributed. This differs only slightly from the normal distribution (by taking the logarithm of the exponent), but adds the flexibility of fitting skewed data. In general, a variable might be modeled as log-normal if it can be thought of as the multiplicative product of many small independent factors [116]. The probability density function for a log-normal distribution is defined as

$$f(x;\mu,\sigma) = \frac{e^{-(\ln x-\mu)^2/2\sigma^2}}{x\sigma\sqrt{2\pi}}, \tag{3.6}$$

where μ and σ are the mean and standard deviation of the variable's logarithm. The expected value and variance are, respectively,

$$\mathrm{E}(X) = e^{\mu+\sigma^2/2}, \tag{3.7}$$

and

$$\mathrm{Var}(X) = \left(e^{\sigma^2}-1\right)e^{2\mu+\sigma^2}. \tag{3.8}$$

For each parameter set, we found the best fit log-normal distribution (pdf) by minimizing the cross-correlation function between the data and the model. The results all five combinations of pressure and length tested are shown in Table 3.1 along with the parameters μ and σ from the log-normal fit of Equation 3.6. Using these parameters, the expected value and variance of the distribution are shown, as computed from Equation 3.7 and Equation 3.8. Aside from the trends in expected value already discussed, the variance appears to decrease as the expected value decreases, that is, as pressure increases or length decreases. The tight 95% confidence intervals indicate that the log-normal fit is good. The fit is less accurate at predicting the variance for

pressure (kPa)	150		200		250	
length (mm)	300		300		300	
	$\mathrm{E}(X)$	$\mathrm{Var}(X)$	$\mathrm{E}(X)$	$\mathrm{Var}(X)$	$\mathrm{E}(X)$	$\mathrm{Var}(X)$
exp. filling times (s)	34.1	130.8	26.1	175.5	16.2	70.8
log-normal fit	34.2	139.0	25.8	84.3	16.0	26.3
fitting parameters	μ	σ	μ	σ	μ	σ
with 95% CI	$3.47^{3.55}_{3.40}$	$0.34^{0.40}_{0.29}$	$3.19^{3.27}_{3.11}$	$0.35^{0.41}_{0.30}$	$2.72^{2.80}_{2.65}$	$0.31^{0.37}_{0.27}$

pressure (kPa)	250		250	
length (mm)	350		400	
	$\mathrm{E}(X)$	$\mathrm{Var}(X)$	$\mathrm{E}(X)$	$\mathrm{Var}(X)$
exp. filling times (s)	22.6	11.6	33.5	52.5
log-normal fit	22.6	9.0	33.5	48.6
fitting parameters	μ	σ	μ	σ
with 95% CI	$3.11^{3.14}_{3.08}$	$0.13^{0.16}_{0.11}$	$3.49^{3.54}_{3.44}$	$0.21^{0.25}_{0.18}$

Table 3.1: Gel filling time experimental data and best-fit log-normal models. For each set of pressures and lengths, the times required to fill 72 capillaries were measured. The expected values, $\mathrm{E}(X)$ and variances, $\mathrm{Var}(X)$ of this data set are shown. The best-fit log-normal distribution models were determined; their mean and variances are shown along with their fitting parameters with 95% confidence intervals.

the conditions of l=300 mm at Δp=200 kPa and 250 kPa.

Using these fits, the filling time for the 10,000 capillary array can be estimated from the cdf (as shown in Figure 3-11 (*right*)). For example, if d_i=75 μm, l=300 mm, Δp=150 kPa, μ=8000 Pa·s, 10,000 capillaries will require 107 s to fill. To summarize the application of this model to the UTMS for conditions not tested: (1) use the viscous flow model to predict the mean filling time, (2) estimate the variance from the measured trends in Table 3.1, compute the parameters μ and σ from Equation 3.7 and Equation 3.8, respectively, (3) compute the distributions, pdf and cdf (4) use these distributions, namely the cdf, to predict of the number of capillaries filled as a function of time (as a percentage).

Additionally, some distribution in filling times was systematic with capillary position in the row of eight capillaries. This is most likely due to pressure and fluid flow gradients inside of the gel loader, which have not been modeled in this work. Our experimental results indicate that capillaries at the perimeter of the array are more likely to require longer filling times; these were included in the statistical distribution analysis above.

3.3.3 Practicality

Through the many hundreds of trials, not only with the eight capillary gel timing apparatus, but with the single, 100, and 10,000 capillary gel loaders as well, some experience in practicality was gleaned. For gel loading apparatuses with less than ten capillaries, the silicone gasket proved to be a highly reliable and robust seal against the back pressure. In particular, the single capillary gel loader (see Figure 3-5), was used for gel loading hundreds of times, with pressures ranging from 150 kPa (22 psi) to 4200 kPa (600 psi), for a range of gel types. During this entire duration, the same silicone seal repeatedly sealed around the capillary after insertion. The pressure serves to deform it around the capillary's outer wall. For gel loaders with more than 10 capillaries, the apparatuses were interfaced directly to the capillary array constraint devices (see Chapter 6). These devices use a thicker, highly compressed silicone seal around the capillary perimeters instead. Again, they performed admirably through

Gel type	Gel loader (number of capillaries)			
	1	80	100	10,000
Spectrumedix PCR-QC	60 s	60-120 s	0-120 s	-
(∼3% LPA)	275 kPa	275 kPa	275 kPa	
Spectrumedix HR	120-180 s	-	-	-
(∼5%LPA)	3.4 MPa kPa			
PVP gel	60 s	-	-	-
	275 kPa			

Table 3.2: Typical filling times for capillaries in practice, for various gel matrices and gel loaders (number of capillaries). LPA–Linear poly(acrylamide), PVP-polyvinylpyrrolidone.

tens of experiments.

Regarding the seal between the gel loader lid and body, the o-ring works reliably up to 4200 kPa (600 psi) if sufficiently compressed, while the silicone gasket will be displaced by pressures above 700 kPa (100 psi). In addition, at the higher pressures, the gel loader will eject the capillaries if they are not restrained in place, because of the axial pressure on the fused silica capillary annulus. Through the various trials, a general set of filling times was obtained for the standard 300 mm long, 75 μm d_i capillary. These are summarized in Table 3.2.

The time required to flush the capillaries with an aqueous surfactant solution as also measured (not shown). As a rule of thumb, the capillaries should be flushed for the same amount of time as they were loaded with gel. An extra 10 s will account beyond this will account for any spatial variation in loading times.

3.4 Conclusions

Gel loaders that can accommodate 1-10,000 capillaries were designed, built, and tested. The time required to fill the capillaries with gel compared favorably with theoretical predictions based on viscous fluid pipe flow. While alternative ideas to lower the filling time were considered, such as shear rate oscillations or heating the gel, the static pressure pneumatic gel loader proved functional for hundreds of trials with arrays ranging from 1 to 100 capillaries. The statistical distribution of capillary

loading times and spatial systematic effects were considered. The loading times are well characterized by a log-normal distribution, since the data is right-skewed (non-normal) with a expected value of tens of seconds at a few hundred kPa. Log-normal fits to the hundreds of capillary filling time measurements enable the prediction of the performance for arrays of thousands of capillaries. In addition, several rules of thumb for utilization of this family of gel loaders were introduced, such as recommended pressures and durations.

Chapter 4

Micro-well array interface

Ultrahigh-throughput denaturing electrophoresis (>5000 channels) can enable rapid, low-cost mutation scans of human genomes. One design challenge for such instruments is successfully interfacing to the capillary channel array ends for briefly loading and collecting DNA while enabling electrophoresis, permitting DNA detection, and preventing evaporation. Micro-well arrays were developed for this application, both polymer and silicon-based, using laser drilling, electrical discharge machining, and micro-milling processes. The devices were used to load and subsequently collect 10^8 fluorescently-labeled DNA primers. Modeling and experimental results were used to determine the duration of pH stability of 250 nL solution, about 60 s for 8 μA electrophoretic current and 20 mM concentration. The devices are easily scalable to 10,000 micro-wells as required for pangenomic mutational spectrometry.

4.1 Introduction

Capillary arrays to date have been limited to hundreds of channels, interfaced to 96-384 well-plates. These plates contribute to channel number limitations, and their large volumes increase consumable costs. Collection of interesting separation fractions from capillaries has been shown for 12-16 channel rows, such as by Minarik *et al.* [117], who devised an agarose gel bed through which the capillary tips were dragged to spread their effluent spatially. Other work [118, 119] transports separation zones to collection

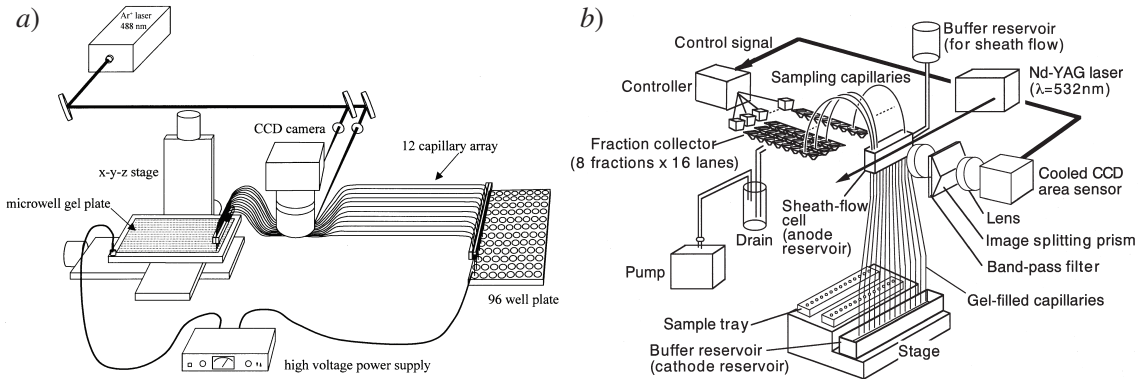

Figure 4-1: Examples of fraction collectors in the literature, capable of collecting the effluent from up to 16 capillaries by either dragging their tips through an agarose gel (reproduced from [117]) or using sheath flow (reproduced from [118]).

vials or capillaries using sheath flow. These examples from the literature are shown in Figure 4-1. These techniques would be very complex and costly to implement for 2-D array of 10,000 capillaries.

4.2 Design and manufacture

Our design involves loading and unloading biomolecules from a micro-well array containing up to 10,000 independent wells. A schematic of the device in relation to the UTMS is shown in Figure 4-2. In the multi-layer assembly, initial registration between the capillary array and micro-wells is made by a pair of guide pins. As the device is translated along these guide pins, the capillaries enter an array of conical thru-holes in polymethyl-methacrylate (PMMA) to self-align the tips. These overlapping cones will interface with capillaries that have maximum allowable radial error of 320 μm, that is, just touching each other As the device is further translated along the guide pins, the capillaries puncture a poly-dimethylsiloxane (PDMS) seal and enter tapered wells which further passively align the capillary tips to the detection system. Detection is permitted by a transparent wafer backing. The well layer can be either highly conductive silicon that can serve as an electrode for electrophoresis, or a thin metallic layer can be deposited on a non-conductive (e.g., PMMA) well layer for electrophoresis. The adhesive backed poly-dimethylsiloxane (PDMS) seal mitigates

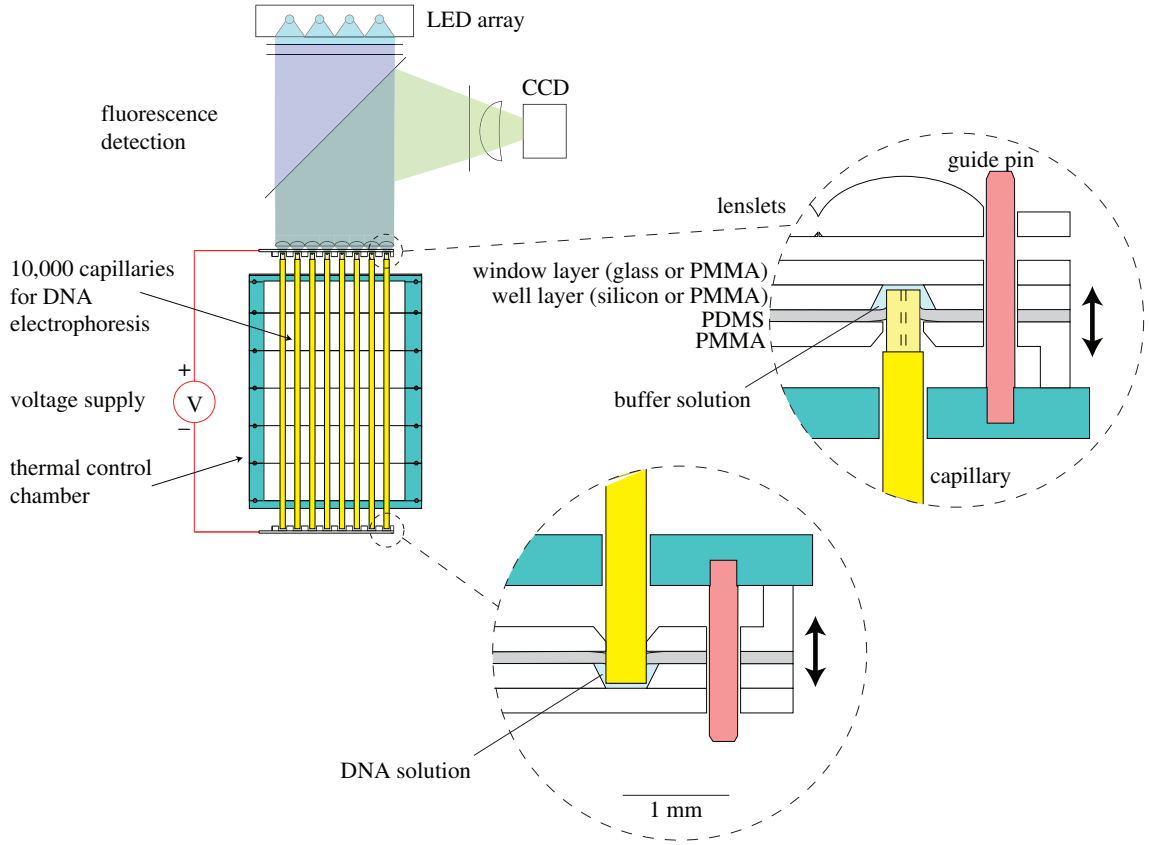

Figure 4-2: The micro-well array can interface to either end of the capillary array. Two designs were manufactured and tested: silicon/glass and PMMA throughout. At the upper end, the capillary tips are aligned for detection with a microlens array.

evaporation of a well's 250 nL solution for >48 hours and repeatedly seals around the shaft of inserted capillaries. The assembly is electrochemically inert, so collected DNA is stable for extended durations (several days in solution).

Both a PMMA and a silicon-based version of the micro-fluidic chip have been designed, fabricated, and tested. In the PMMA design, 1.6 mm thick wafers are laser machined with a high numerical aperture (NA) focusing lens to achieve tapered sidewalls. Laser machining is accomplished with an 80 W CO_2 laser with a 100 μm focal spot size directed using an x-y translation stage (Trotec, Speedy 100). Devices with 10,000 wells have been manufactured (see Figure 4-3). The wafer is then metallized with 100 nm of gold by evaporation and bonded to a blank PMMA wafer using dichloromethane. The PDMS seal is then applied.

For the silicon-based chip, hole array fabrication is accomplished using a faster,

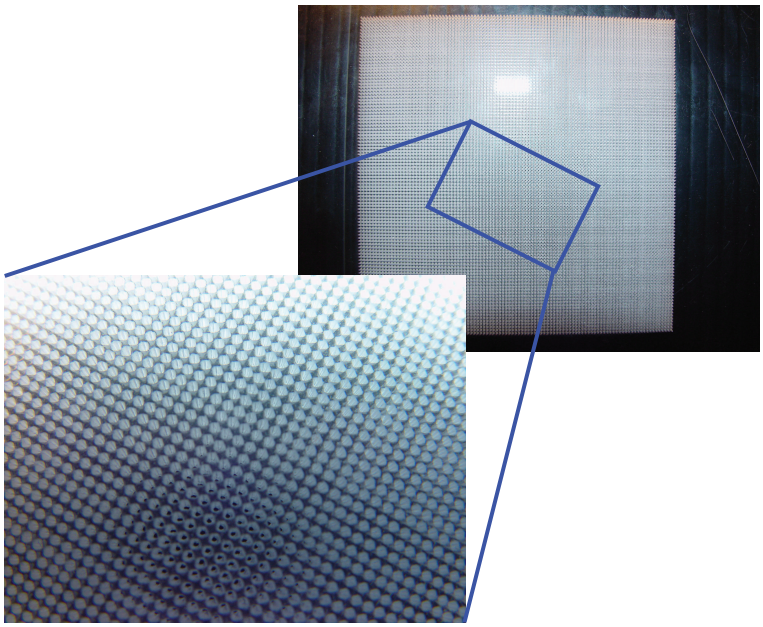

Figure 4-3: Ten-thousand well arrays have been manufactured from PMMA using an 80 W CO_2 laser with a 100 μm focal spot size directed using an x-y translation stage (Trotec, Speedy 100).

cheaper set of processes than conventional MEMS etching, as first described for a different application by Kanigan *et al.*, in our laboratory [120]. The array is created by electrical discharge machining (EDM). A plunger is fabricated first using wire EDM (Charmilles, Robofil 1020si), and then the hole array is die-sunk through a 500 μm thick, highly doped with antimony (0.0008-0.002 Ω-mm) silicon wafer in <1 min using a sink EDM (Charmilles, Roboform 30). The holes are spaced 1 mm apart, 700 μm and 400 μm wide at top and base, respectively, with process variation of approximately 10 μm. The plunger can be reused dozens of times with minimal degradation. This process and a comparison to anisotropic silicon etching are presented in Figure 4-4. After plunging, the wafer is cleaned, chemically-mechanically polished (CMP), anodically-bonded to a borosilicate wafer (400 °C at 800 V, limited to 5 mA for 5 hr), and PDMS sealed. The assembly is cut on a die-saw to a 50 mm $\times$ 50 mm square and mounted in cartridge. A PMMA layer with guide pin holes and conical holes matching the wells is used to seal the cartridge. The conical holes are drilled with a 381 μm diameter spot drill (size 0000') at 21,000 rpm in a CNC machining center (Haas, VF-OE). This entire assembly is shown for the silicon device in Figure 4-5.

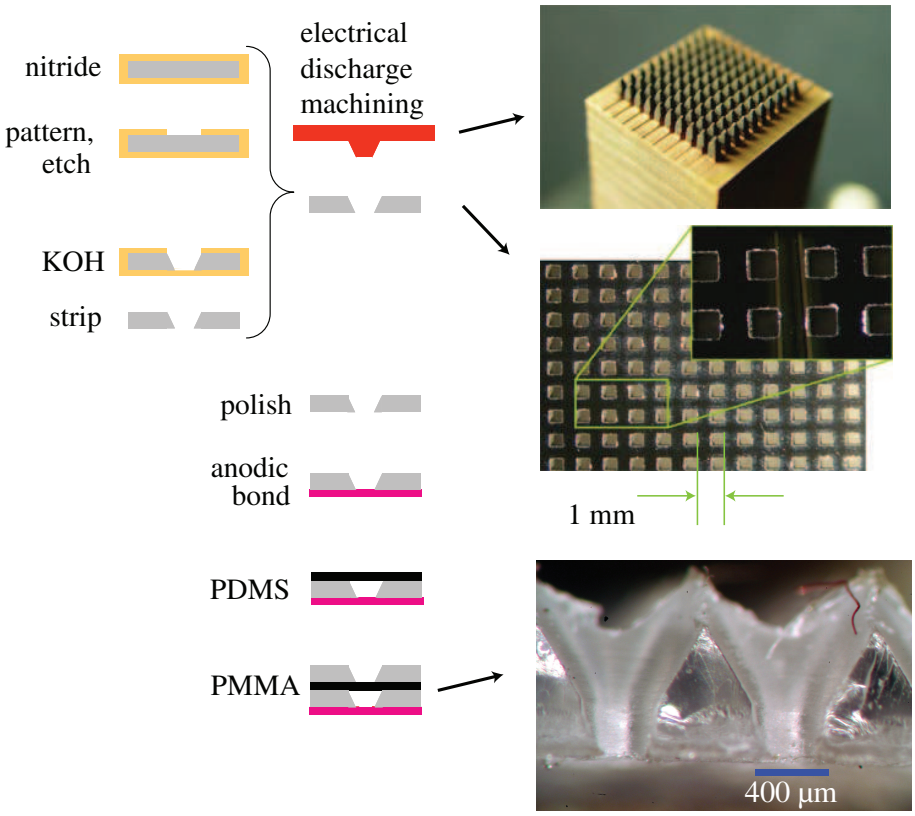

Figure 4-4: Manufacturing process for silicon-based micro-well array device, compared with anisotropic silicon etching. After plunging, the wafer is cleaned, chemically-mechanically polished (CMP), anodically-bonded to a borosilicate wafer, and PDMS sealed. A PMMA layer with an array of conical holes is attached last.

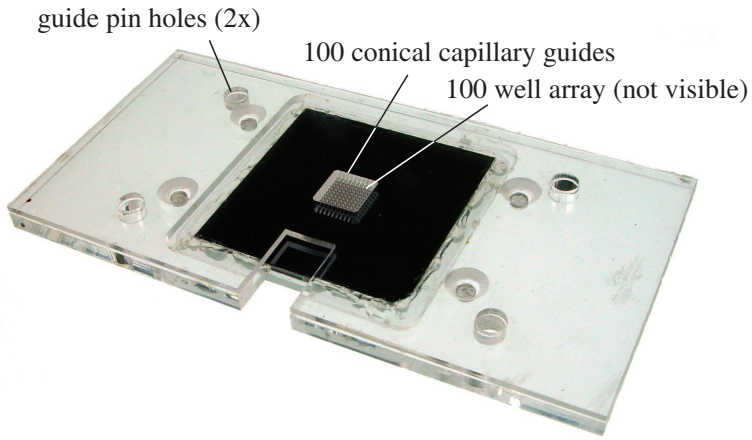

Figure 4-5: The silicon-based micro-well array, containing 100 250 nL wells, is cut square and glued into a PMMA cartridge which contains guide pin holes and conical holes to align the capillaries with the micro-wells.

In this work, the EDM process application to highly-doped silicon was extended to compatibility with traditional MEMS fabrication techniques. CMP can enable planarization of reflowed silicon for subsequent processing such as evaporation beam metal deposition and anodic bonding, both of which have been performed. The reflowed silicon was of amplitude 60 μm, as measured with a Dektak profilometer. The EDM manufacturing process for features in silicon offers advantages over MEMS anisotropic etching such as lower cost (no lithography required), speed (1 min to etch through wafer versus overnight for potassium hydroxide (KOH), and more freedom over the sidewall angles whereas KOH etching is limited to the crystal lattice orientation.

4.3 Experimental methods

4.3.1 DNA loading

To test DNA loading capability, a 20-base primer fluorescently-labeled with AlexaFluor 488 (Synthegen) at 16.6 nM concentration [121] and water control were electrokinetically injected from both an Eppendorf tube and the silicon-based micro-well array into a 125 mm long, 75 μm d_i, 360 μm d_o capillary (Polymicro Technologies, TSP075375) pre-filled with a non-crosslinked polyacrylamide matrix (Spectrumedix, PCR-QC). Detection was performed using side-column laser-induced fluorescence with a 100 mW argon-ion laser and photo-multiplier tube (PMT) while the DNA was undergoing electrophoresis at 10 V/mm. To inject from the micro-well array, the capillary end was plugged through the PDMS layer into a pre-filled well. A 10 s injection at 2 μA loaded 10^8 molecules [121]. This experimental setup is depicted in Figure 4-6. Reagents and sample preparation techniques are further described in Appendices A and B.

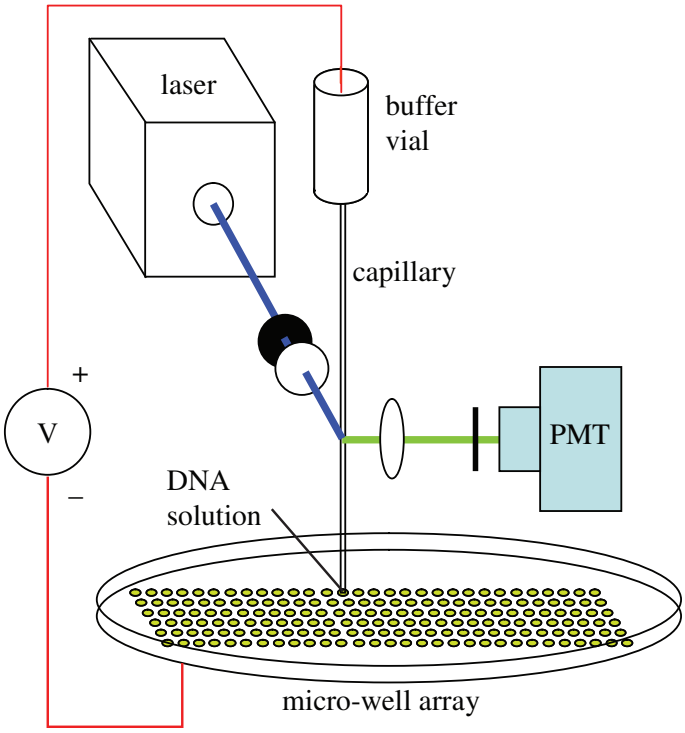

Figure 4-6: Experimental setup for detection of DNA loaded from the 250 nL silicon-based micro-well array. Side-column laser-induced fluorescence with a photo-multiplier tube comprises the detection system.

4.3.2 DNA collection and end-column detection

To test the micro-well's DNA collection and simultaneous end-column detection capability, experiments were performed on both the PMMA and silicon-based devices. In preparation for detection, the 20 μm polyimide coating was removed from the tip of the capillary (as shown in Figure 4-2), and the capillary was plugged into the micro-well array through the PDMS seal. This coating removal procedure is described in Appendix A. DNA was again injected electrokinetically, and then the end-column detection system described in Chapter 7 and shown schematically in Figure 4-2 was used to interrogate the contents of each buffer solution-filled micro-well simultaneously and continuously during electrophoresis.

4.4 Results and discussion

4.4.1 DNA loading, collection, and detection

For the DNA loading experiments, results in Figure 4-7*a* indicate that DNA can be successfully loaded from the 250 nL micro-wells as compared to a traditional Eppendorf tube containing 5 μL. Figure 4-7*b* and Figure 4-7*c* show the intensity of the CCD pixels covering the lenslet area versus time during electrophoresis for the PMMA and silicon-based micro-well arrays, respectively. In both cases, a peak with signal-to-noise ratio (SNR)>7 is observed, illustrating successful collection and detection of 10^8 primers using the micro-well.

The silicon-based device was created to avoid PMMA processing challenges as fabrication variation of approximately 100 μm combined with poor adhesion in the presence of the buffer solution led to unacceptable leakage between wells. These results demonstrate our ability to detect DNA through the window of the micro-well array, but the SNR is relatively poor. Improved SNR results with a common buffer reservoir, as opposed to the isolated micro-wells, as well as modified end-column detection system, are presented in Chapter 7. Another potential cause of the low SNR in the end-column detection data is pH changes in the buffer solution within the micro-

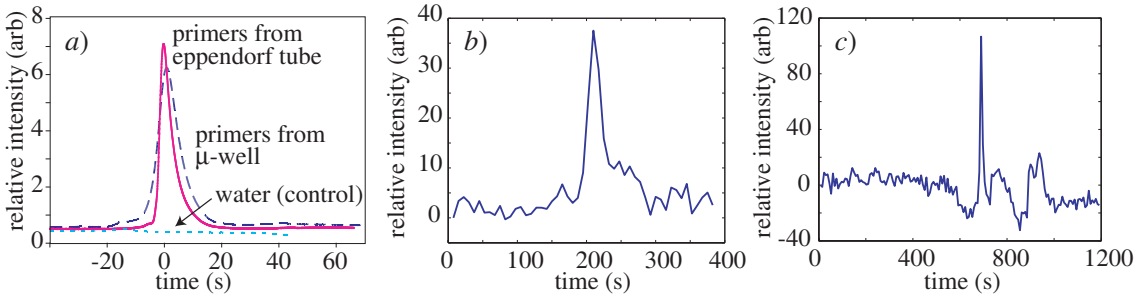

Figure 4-7: (*a*) Side-column laser-induced fluorescence detection of 10^8 primers loaded from the buffer micro-well array into a capillary compared to conventional loading from an Eppendorf tube. End-column detection of 10^8 primers collected into buffer well array using (*b*) PMMA and (*c*) silicon-based devices.

wells during electrophoresis. To investigate this possibility, theoretical calculations about the pH stability of a 250 nL buffer solution were undertaken.

4.4.2 Buffer pH stability

During electrophoresis, electrolysis is taking place. The electric current leads to the accumulation of anions at the anode, or equivalently, depletion of cations. This accumulations of ions at the anode by the electric current is an electrolytic modification to the buffer, which causes the pH to decrease, becoming more acidic. These effects have been modeled or experimentally measured by several researchers [122, 123, 124, 125, 126]. In this reaction, H^+ is generated to maintain neutrality according to

$$2H_2O \rightarrow O_2 + 4H^+ + 4e^-. \tag{4.1}$$

Correspondingly, the electrical conductivity and ionic strength of the buffer change, affecting:

1. Reproducibility of consecutive separations with the same buffer.

2. Migration time of analytes (DNA) through variations in electro-osmotic flow.

3. Electrophoretic mobilities of analytes sensitive to the pH (i.e., protonation of the DNA). pH change from 3-8 can change the electroosmotic velocity by a factor of 500% in fused silica capillaries [127].

The time that the buffer can maintain the pH can be estimated from the Faraday equation [128], given by

$$t = \frac{[\text{buffer}]\, VF}{i}, \tag{4.2}$$

where [buffer] is the buffer concentration, V is the volume of buffer solution, i is the current, and F is the Faraday constant equal to 96485.4 Coulomb to cause one equivalent of reaction. For a concentration and volume of 20 mM/L and 250 nL, respectively, and current of 8 μA in our micro-wells, the buffer will be unable to neutralize the generation of H^+ after 60 s. Assuming that the molecular weight, M_W, and density, ρ, of the buffer is equal to that of water (18 g/M, 1,000 g/L, respectively), H^+ will be generated at a rate of $i/F = 8.3 \times 10^{-11}$ M/s, and thus the $\rho V/M_W$=1.4×10^{-5} M of buffer will be completely gaseous in 46 hr. While the conversion from liquid to gas is not a concern, the buffer depletion in 60 s certainly is. This calculation does not account for an increase in OH^- or H^+ generation with time, nor the effects of electroosmotic flow, which will increase with time. These factors will lead to a net decrease in the neutralization time by about 10% [128]. Other assumptions are that the solution is electrically neutral, water hydrolysis is the only electrode process occurring, ion mobilities and the dissociation constant of the buffer are equal to those at infinite dilution, the influence of pH and conductivity changes on the electric current is negligible, and that molecular diffusion is negligible.

To further explore this concern, we measured the voltage required to maintain current, i, as function of time in the micro-wells. The results are shown in Figure 4-8. The voltage for the 250 nL micro-well increases as the pH changes, while that for the 3 mL vial stays constant. Using Equation 4.2 above, the 3 mL vial can maintain the pH for 200 hr. Thus, the micro-wells should only interface with the capillary array for 60 s for electrokinetic loading or collection at 8 μA to maintain pH neutrality. The AlexaFluor dye is reportedly insensitive for 4<pH<10, but fluorescence is reduced below that [129]. This effect could contribute to the low SNR. The preceding factors were considered in the design of the buffer reservoirs (see Chapter 5) into which the capillary array ends are terminated for the majority of the electrophoretic run.

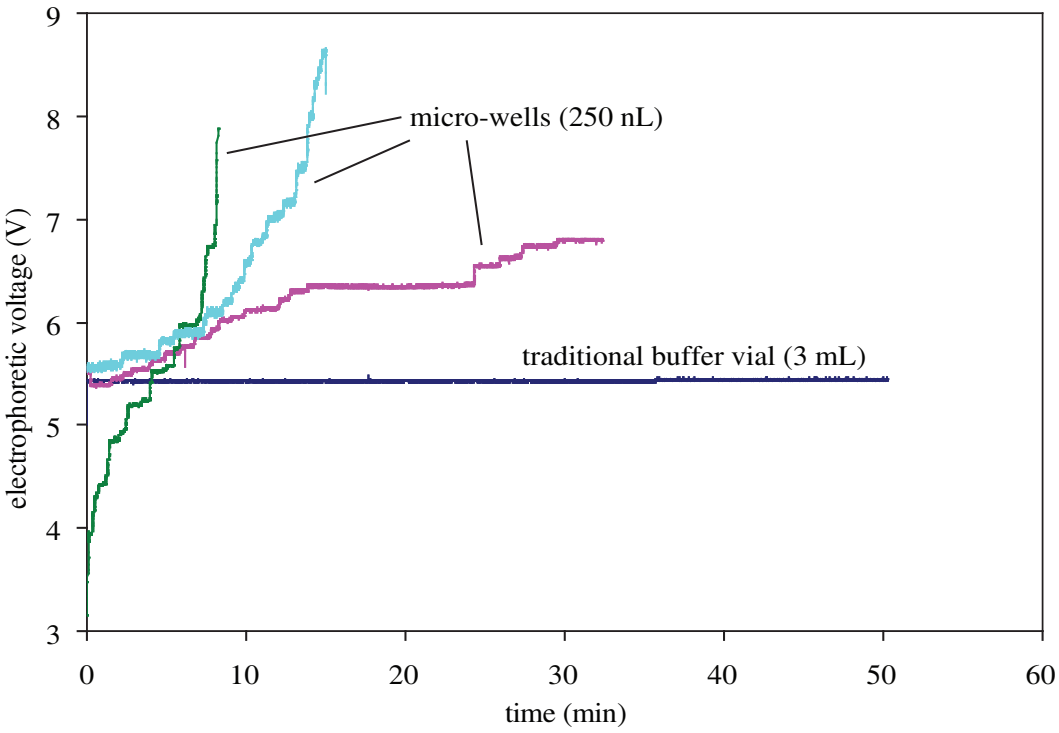

Figure 4-8: Electrophoretic voltage required to maintain a constant current of 8 μA in a 300 mm capillary filled with linear polyacrylamide (Spectrumedix, PCR-QC).

Other researchers (e.g., [119, 130]) have overcome this challenge by separating the buffer chamber from the analysis chamber with a semipermeable membrane. This configuration permits the flow of small buffer molecules for pH stability while retaining the larger DNA molecules for analysis. A cellulose dialysis membrane was considered for our application. The buffering ions are <0.1 kD in size, while the DNA would be considerably larger. Each nucleotide has a molecular weight of 330 g/M, so a 20 base primer is 6.6 kD while a 100 base double-stranded gene fragment is 660 kD. The recommended membrane molecular weight cutoff (M_WCO) is determined by satisfying

$$M_W\text{CO} < \frac{1}{2} M_W\text{retained solute}, \tag{4.3}$$

and

$$\frac{M_W\text{retained}}{M_W\text{passed}} \geq 25. \tag{4.4}$$

A membrane with 0.2 kD<M_WCO<330 kD would satisfy Equation 4.3. Equation 4.4 results in 660 kD/0.1 kD = 6600. So 6-8 kD cellulose membranes, 0.07 mm thick and flexible, were purchased (Spectrum, spectra/Por RC). Testing with the micro-well arrays to combine the goals of fraction collection and pH stability for extended durations (>60 s) remains for future work.

4.4.3 Independent capillary control

The prospect of independently controlling capillary electrophoresis in each of the capillaries within a 10,000 capillary array is enticing. Were it possible, one could collect the DNA from a single capillary while keeping the analytes in the others stationary. Or one could vary the voltage on the array spatially to account for inherent differences to achieve migration velocity uniformity. Or one could imagine modulating the voltage to enhance the separation efficiency.

We considered this possibility with a few calculations about dielectric breakdown, the one concern with this idea. If a capillary is being driven at 5-10 kV while its neighbor, 1 mm away is stopped (grounded), sparking between them is a real possibility and unallowable. Independent wiring to each capillary micro-well would require

on the order of 6 μm between wires, assuming 6 μm wide wires, and 400 μm wide holes spaced 1,000 μm apart. In air, merely 18 V is required to bridge this 6 μm gap, with its inherent dielectric strength of 3×10^6 V/m. Few materials would be up to the insulation job in fact. One candidate material would be sputtered SiO_2, used in integrated circuit manufacturing. With a dielectric strength of 1×10^9 V/m, it could withstand a potential difference of 6 kV between wires spaced 6 μm apart before breakdown. We decided to use a common potential for all capillaries in the array for this phase of the research because of the complexity and cost of this approach, but it remains an interesting avenue for future work.

4.5 Conclusions

To manipulate DNA, we have utilized both a PMMA and silicon-based bioMEMS design platform for interfacing to capillaries, thus simultaneously enabling electrokinetic biomolecule loading, detection, and fraction collection of 10^8 DNA primer molecules in independent wells. The silicon-based micro-well array is simple and cheap to fabricate, while meeting the design functional requirements for brief (e.g., less than 60 s) array interface durations. If longer durations are essential, a device with integrated semipermeable membrane appears offer promise. Additionally, insulation with sputtered SiO_2 may enable independent capillary electrophoresis control without dielectric breakdown.

The application of this micro-well array for loading up to 100 DNA samples simultaneously is described in Chapter 8. This device is an integral subsystem of the ultra-high throughput mutational spectrometer which will ultimately enable population-wide pangenomic analyses to uncover the genetic causes of common diseases.

Chapter 5

Buffer reservoirs

For electrophoresis, a buffer solution is required to transfer the electric potential from the cathode or anode to the respective ends of the separation channel, capillary, or slab. Buffer reservoirs were designed and manufactured to interface with each end of the capillary array for this purpose. Within the buffer reservoirs, a variety of buffer solutions were used, as described in Appendix A. In this Chapter we present the design of these reservoirs, focusing particularly on a passive connector which allows the capillary tips to be cleanly inserted and aligned for optical detection, DNA loading or collection from/into the micro-well array, etc. We address issues with the buffer reservoir design such as pH stability, electrical and fluidic connectivity, insertion accuracy and repeatability, and mitigation of cross-talk between channels. For most of the Chapter, the discussion centers on a 100-port reservoir, but devices with 5-1,000 ports have been designed, fabricated, and tested. These devices serve as an essential subsystem of the UTMS for high throughput denaturing capillary electrophoresis.

5.1 Introduction

Traditionally, slab gel or capillary electrophoresis employs relatively large 3-300 mL reservoir of 1× buffer solution at both ends of the electrophoretic separation. The design becomes somewhat more complex for multi-capillary instrumentation, as the

capillaries must all protrude into the solution, considering gravity and leakage. We seek to design a buffer reservoir device into which 1-10,000 capillaries can be readily inserted without leakage. Additionally, the buffering solution must be of sufficient volume to maintain pH during electrophoresis. Furthermore, the design should not interfere with the requirements of end-column detection: namely that the capillary tips are optically accessible and aligned radially and axially.

In this work we devised a variety of buffer reservoirs, each designed to interface to the capillary array and maintain pH for a proscribed duration during electrophoresis. Previously, in Chapter 4, we discussed the requirements for pH stability, and the buffer well designs described in this Chapter adhere to these volume requirements. After a brief discussion of the required volumes for electrophoresis with 10,000 capillaries, we introduce the heart of the buffer reservoir design–the micro-fluidic connector, which permits repeated interfacing to the capillary array. Lastly, we present improvements and future work for the design to semi-isolate the capillaries from each other to prevent optical contamination, or "cross-talk" between neighboring capillaries.

5.2 Buffer reservoir volume for pH stability

In Chapter 4, Section 4.4.2, we discussed the chemical reaction of electrolysis during electrophoresis imposes a finite duration on the pH stability of a buffering solution. Using Equation 4.2, we can calculate that 150 mL of buffer at a concentration of 20 mM/L is required to maintain 8 μA per capillary in 10,000 capillaries for 1 hr. This sets a practical lower limit on the volume capacity of the buffer reservoir for the UTMS.

5.3 Passively aligned 100-port micro-fluidic connector

To interface the buffer reservoir with the capillary array, we developed a repeatable, accurate, micro-fluidic connector, and tested the device at 100 capillary capacity. In

the design, the capillaries are passively aligned for connection with a combination of precision guide pins and conical tapered holes, and sealed with a PDMS layer after insertion. Using the device, an array of 100 loosely-assembled capillaries with 1 mm spacing can be aligned to 20 μm radial accuracy and high repeatability, with only 5% failure after 10,000 insertions. We have confirmed electrical and fluidic connectivity as well. This device can be simply manufactured and adapted for numerous micro-fluidic interface applications. The relevance of this work goes beyond our capillary array electrophoresis instrumentation, as widespread application of micro-fluidic systems has been hindered by the challenge of repeatable, accurate interfacing between integrated channels or wells and external loading, detecting, or stimulating devices.

Despite the compelling advantages of microfluidics (e.g., lower reagent volumes or template numbers, higher surface area to volume ratio, and more rapid, portable analysis), there are numerous complications for micro-fluidic device design and manufacturing. A key challenge is fluidic connectivity, which is often required to deliver or withdraw nanoliter-scale volumes from the device [131, 132]. These fluidic connections are often an essential link between an integrated channel or well and external loading, detecting, or stimulating devices. Previous work has been limited to only a few micro-fluidic ports [133, 134], whereas hundreds or thousands of ports can be required in applications which fully exploit the potential of miniaturization, such as the UTMS.

As a proof-of-concept for the 10,000 capillary UTMS, we developed a 100-port micro-fluidic connector to service either end of the capillary array. The application of this connector as a part of the buffer reservoirs is shown in Figure 5-1. We require that the connection be made rapidly (<10 s) with inherent sealing to mitigate evaporation and prevent contamination. We desire to have the individual capillaries simultaneously and passively aligned to less than 50 μm radially to facilitate mechanical access and optical detection. Furthermore, the micro-fluidic connector must tolerate a high-voltage (10 kV), chemically-active environment.

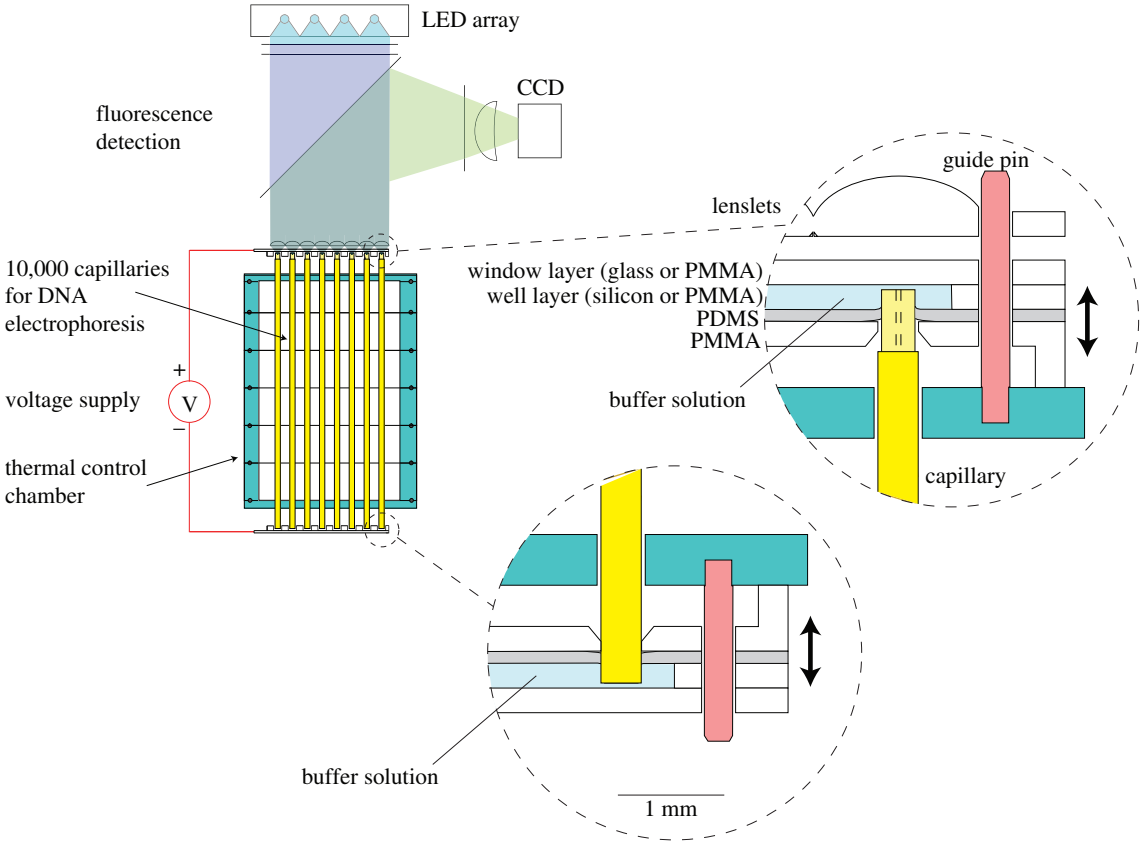

Figure 5-1: The buffer reservoir, much like the micro-well array, can interface to either end of the capillary array. At the upper end, the capillary tips are aligned for detection with a microlens array. The buffer reservoirs rely on the micro-fluidic connector to interface to the capillary array.

	Random errors (mm)	Root-sum-squared random errors (RSS) (mm)	Average of random &RSS errors (mm)
Δx,Δy	0.38	0.21	0.30

Table 5.1: Error budget for capillary tip alignment structural loop.

5.3.1 Connector alignment theory

In order to predict repeatability and accuracy of insertion for the capillaries into a connection interface, we created a system error budget [135]. With the capillary tip as the "tool," and the fluid connection port as the "workpiece," we proceeded to trace random errors through the structural loop to predict the maximum mechanical alignment error. Random errors were tallied from the capillary tip to a guide pin interface to conical entry ports through eight coordinate systems. These errors were dominated by machining tolerances and capillary radial misalignment. Random error results are displayed in Table 5.1; systematic errors were negligible, as were axial (z) errors. The structural loop is shown in the design in Figure 5-2, to be further discussed.

Thus the predicted maximum alignment error between the capillary tips and the conical fluidic ports is 0.30 mm. We therefore designed the entry port for each capillary to have a radius large enough to tolerate this misalignment; 0.5 mm radius entry ports can tolerate a misalignment of 0.32 mm.

5.3.2 Connector materials and methods

The main challenge in mechanical design of the fluid connector is to permit the maximum predicted error while enabling reliable, repeatable, and accurate interfacing. This is initially accomplished with the use of two 6 mm steel guide pins to align the device prior to capillary insertion, as shown in Figure 5-2.

The connector is translated along the guide pins in use, causing each capillary tip to be directed by a pair of conical tapers (primary and secondary) into a 400 μm hole. These conical ports independently and passively bend the capillaries into

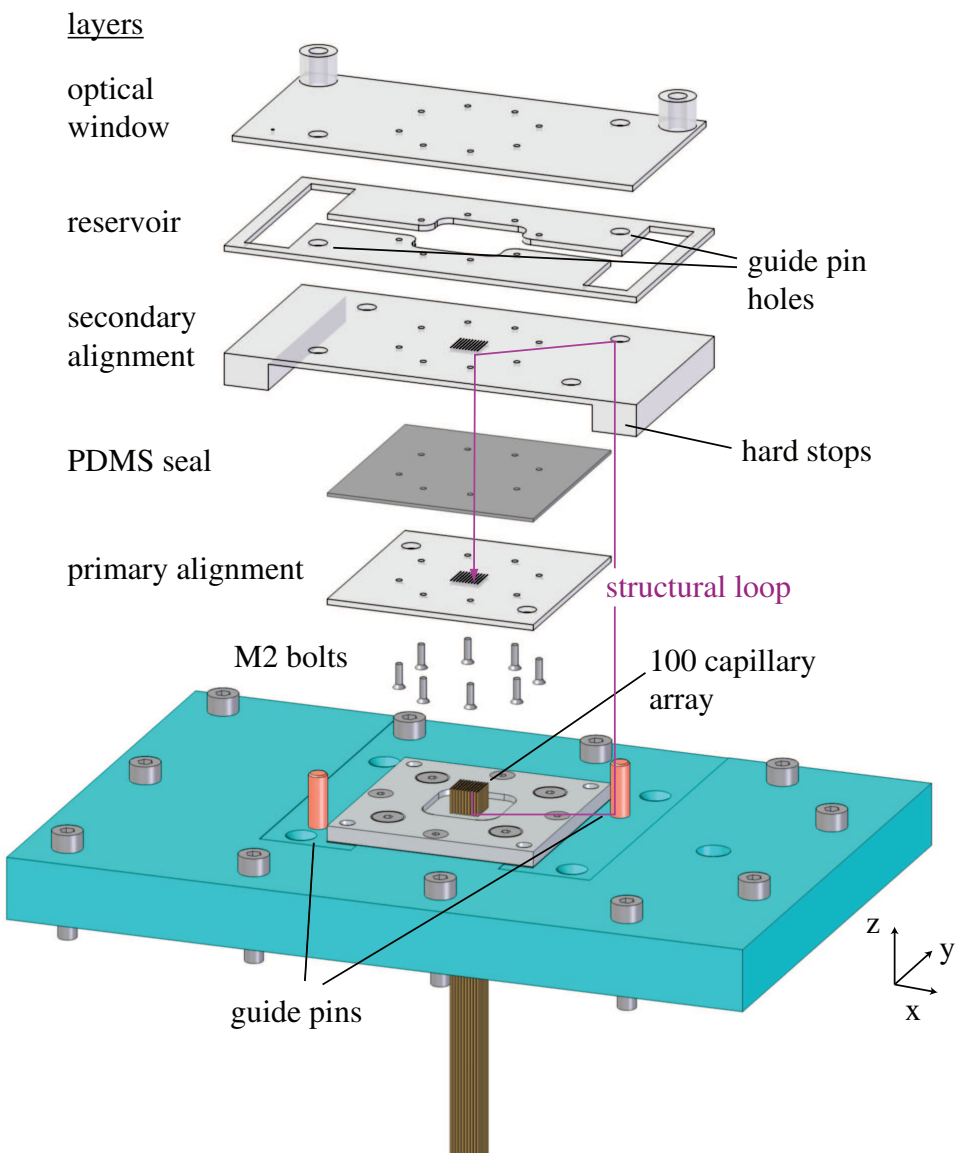

Figure 5-2: Exploded model of the 140 mm × 70 mm micro-fluidic connector above the capillary array and guide pins. Entry ports on the primary alignment layer are registered to those on the secondary alignment layer. This secondary alignment layer has reamed guide pin holes which are machined in the same setup as the conical ports for accurate registration. After the guide pins provide coarse capillary alignment, the primary and secondary alignment layers engage the capillaries.

an aligned grid with 20 μm radial accuracy, spaced 1 mm apart. The conical ports are drilled with a 381 μm diameter spot drill (size 0000') at 21,000 rpm in a CNC machining center (Haas, VF-OE). The aligned capillary array within the connector is shown below in Figure 5-3.

The entire device (see Figure 5-2), is constructed using four separate layers of 1.6 mm thick polymethyl-methacrylate (PMMA), fluidically sealed with a layer of poly-dimethylsiloxane (PDMS) and bonded with PMMA solvent methylene dichloride. Each PMMA layer is laser cut (Trotec, Speedy 100) and selectively milled. As the capillaries are inserted into the device, they pass through the PDMS layer, which and seals around their perimeters. When the capillaries are removed, the silicone reseals, preventing escape of fluid from the reservoir. Translation along the guide pins is limited by hard stops on the secondary alignment layer, preventing capillary damage. During electrophoretic separation of DNA within the capillaries, a 5 mL buffer solution in the reservoir layer permits application of high voltage (up to 10 kV tested) to the capillaries via a platinum wire electrode, as well as collection of the eluted DNA. Detection of DNA is permitted by a PMMA optical window layer. (Figure 5-3 photograph taken through this window.)

In addition to the 100 capillary capacity buffer reservoirs described, we designed and manufactured a variety of buffer reservoirs with capacities of $\sim$5-1,000 capillaries. The $\sim$1,000 capillary buffer reservoir, with capability to continually recirculate buffer through the shallow channel where the capillary tips reside for end-column detection, is shown in Figure 5-4. The recirculation features assures that a continuous supply of buffer is available for pH stability and that the fluorescently-labeled DNA effluent is whisked away to prevent cross-talk. This devices was used in some preliminary experiments with superb pH stability. Subsequent designs avoided recirculation due to its complexity.

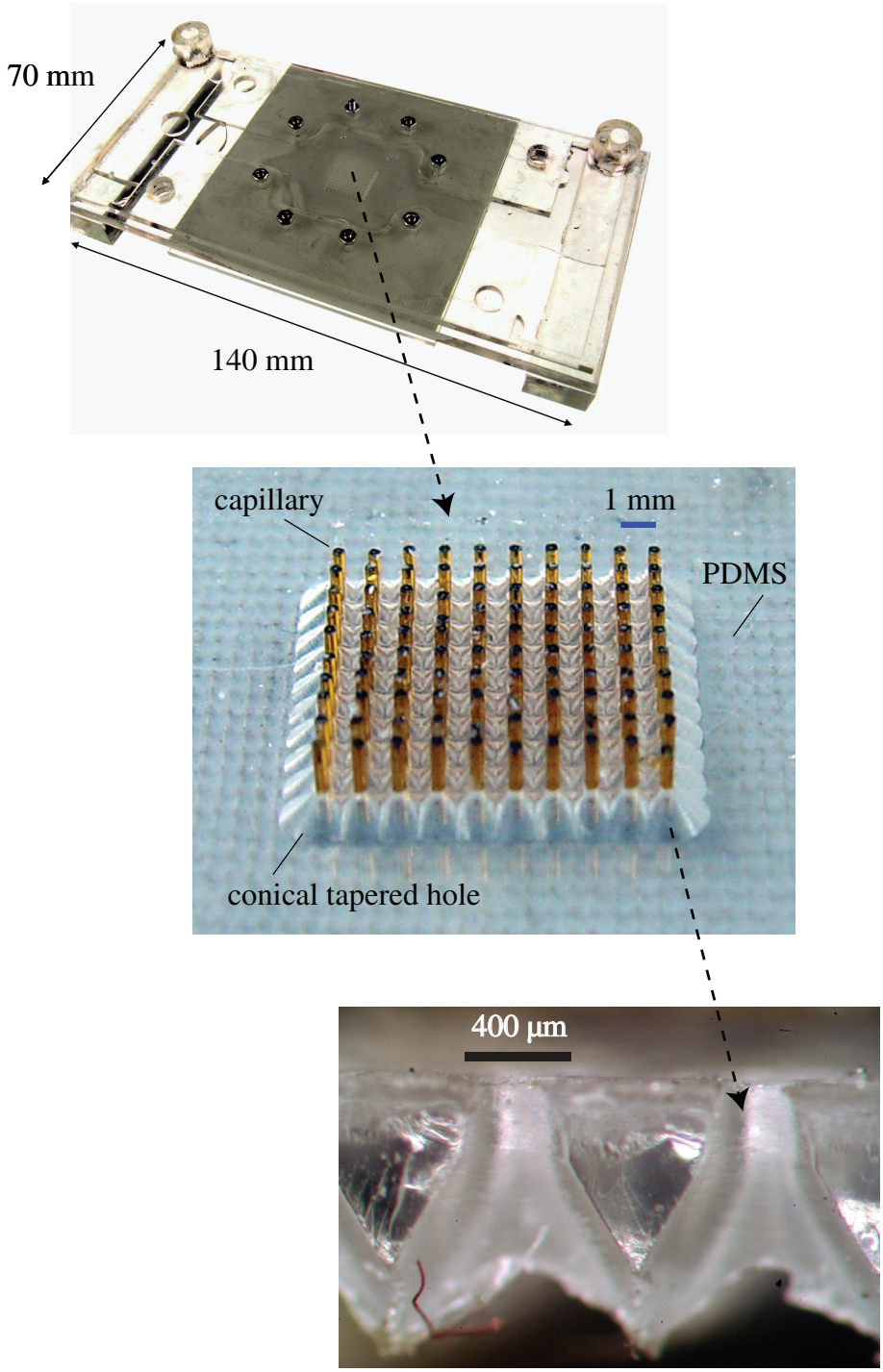

Figure 5-3: Photograph of 100 capillaries (diameter 360 μm) inserted and sealed into the micro-fluidic connector. The conical ports, spaced 1 mm apart, align the capillaries accurately to 20 μm radially. The PDMS layer seals and inhibits evaporation. Insertion of the 100 ports requires approximately five seconds by hand.

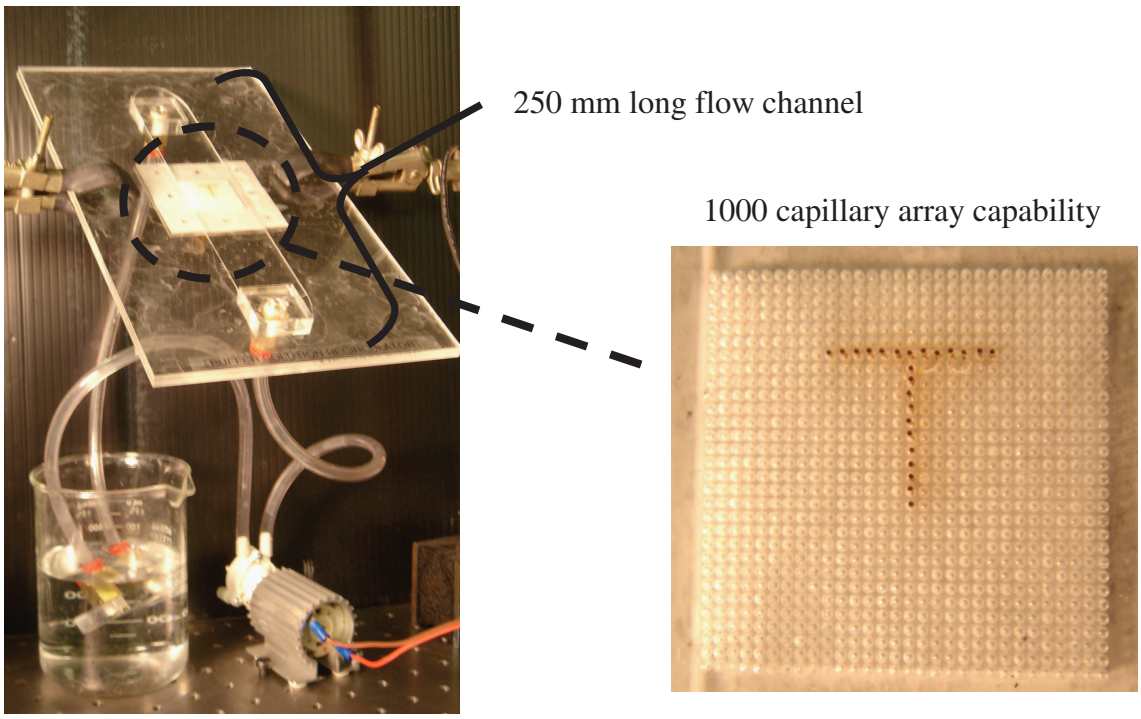

Figure 5-4: The ~1,000 capillary buffer reservoir, with buffer recirculating pump. A capillary array in the shape of a "T" is inserted for testing. Recirculation from a large tank prevents pH changes and removes DNA from the capillary tips hydrodynamically, but the design is somewhat complex.

5.4 Results and discussion

5.4.1 Capillary radial alignment accuracy

For each of the capillaries shown in Figure 5-3, we measured the capillary tip alignment before and after insertion into the 100 port connector. For the 100 capillaries tips protruding 8 mm from an elastomer constraint device described in Chapter 6, the 3σ radial deviation from a regular grid with 1 mm pitch is 187 ± 7 μm. Since the holes in the fluidic ports are only 400 μm in diameter, the 360 μm diameter capillaries must be aligned to, at worst, 20 μm radially after insertion.

5.4.2 Repeatability

The repeatability of insertion was measured by a 100 trial experiment. The microfluidic connector was inserted to the 100 capillary array 100 times, resulting in a total of 10,000 individual capillary insertions. During the course of these trials, 6 capillaries were rendered unusable. In total, 9517 out of 10,000 capillaries were successfully inserted, 95.2%.

5.4.3 Electrical and fluidic connectivity

To verify electrical and fluidic connectivity, at test of the resistance of the capillary array was conducted. From the cathode, the current flows through the buffer solution in the reservoir, through the gel within n capillaries in the array, and though another buffer solution reservoir. The number of capillaries that have been filled with gel can be determine by measuring the total resistance in the system. As shown in Figure 5-5, the total resistance as a function of n will decrease monotonically. An Ohmmeter was used to measure the resistance of the system through a pair of platinum electrodes in either buffer reservoir. By measuring the resistance of the system as $\sim$19 MΩ, we can conclude that all 100 capillaries are filled with gel, and that the buffer reservoirs, engaged with the 100 port microfluidic connectors, enable electrical and fluidic connectivity to the capillary array.

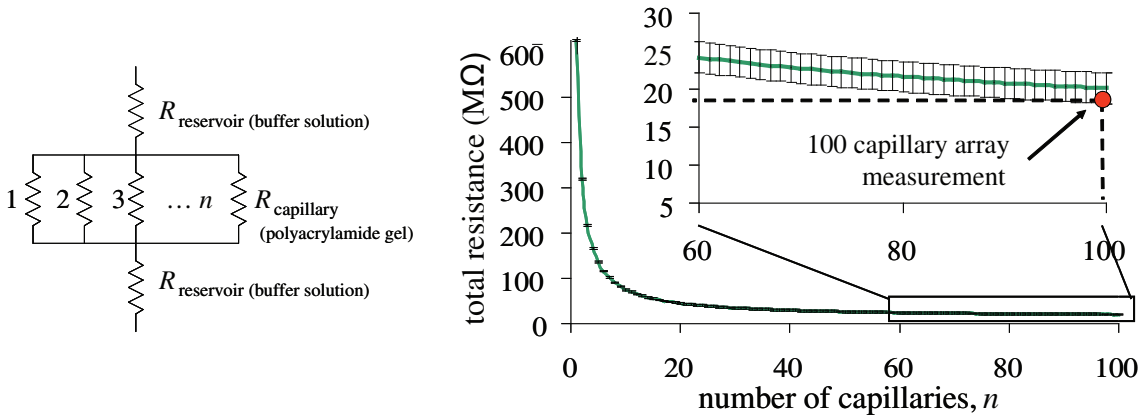

Figure 5-5: Electrical resistance measurement (circle) for the 100 capillary array allows the determination of the number of capillaries filled with sieving matrix gel n, as well as verifies electrical and fluidic connectivity to these n capillaries. The total system resistance was measured as $\sim$19 MΩ.

5.5 Semi-isolation of capillaries

Upon fluorescently-labeled DNA elution into from the capillary array into a common buffer reservoir, "cross-talk" between neighboring capillaries can occur. Semi-isolation of these capillaries is required to prevent the signal from one capillary from being detected at another. In Chapter 4, Section 4.4.2, several approaches to mitigating this were presented, such as isolating each capillary's elution well from a larger common buffer reservoir by means of a semipermeable, dialysis membrane, or briefly collecting eluted DNA in micro-wells that can sustain electrophoresis for a few minutes without a pH change.

We also designed implemented semi-isolation well as features in the common buffer reservoir optical window layer (see Figure 5-2). These pocket features assured that the DNA migrating towards the electrode during electrophoresis was on a plane below the detection plane. Initial results were encouraging, but this remains an active area of future work.

5.6 Conclusions

This proof-of-concept describes the design of a 100 port micro-fluidic connector enabled buffer reservoir for application to high throughput capillary electrophoresis instrumentation as well as the design of a 1,000 port buffer reservoir. A volume of >150 mL within these reservoirs is stable for an hour at typical electrophoretic currents in a 10,000 capillary array, and in addition we have discussed techniques for overcoming cross-talk between neighboring capillaries. We have completed devices for both ends of the capillary array. This interface technique allows for a repeatable, accurate method of introducing a fluid reservoir, electrical connectivity, and 20 μm radial alignment to 100 capillaries, and potentially more, by hand.

This interface is a low cost, easily constructed method of alignment for multiple ports, which in future work will be scaled up to a 10,000 port device. Applications to other micro-fluidic fields, such as cell manipulations and DNA applications, as well as microchip interfacing and chemical mixing are simple, require only minor modifications to the device for chemical compatibility and fluid requirements.

Chapter 6

Assembly and constraint technology for large arrays of capillaries

Discovering the genetic causes of common diseases may require pan-genomic mutation scanning of all genes in a million people. An increase in throughput of genetic analysis instrumentation by several orders of magnitude is essential to undertake such an ambitious task. This Chapter describes the design, manufacture, and testing of assembly and constraint technologies for arrays containing as many as 10,000 capillaries packed into 1-D rows or 2-D arrays with 1 mm spacing. Capillaries are sealed around their perimeters against pneumatic and hydrodynamic pressures useful for automated capillary array electrophoresis. We show that both ends of the arrayed capillaries are aligned axially to 11±5 μm accuracy and repeatability. Radially, the capillary tips are aligned well enough for insertion into arrays of sample wells and subsequent end-of-capillary fluorescence detection. Using our semi-automated assembly machine, we demonstrate 100 capillary array assembly in 12 minutes. This array assembly and constraint technology could be incorporated into commercial capillary instruments, and may enable a new generation of ultrahigh throughput instruments with 2-D arrays of 10,000 or more capillaries.

6.1 Introduction

Dramatically increased throughput (e.g., 100 channels→10,000 channels) in biological instrumentation can enable systematic studies across thousands of samples simultaneously. For DNA analysis applications such as sequencing and mutation detection, this increased throughput is typically achieved by patterning more channels on a planar substrate using semiconductor manufacturing techniques or manufacturing arrays of multiple capillary tubes in which analyses can be performed.

Semiconductor manufacturing techniques have enabled devices with several hundred channels on a glass substrate [90, 91]. These devices have proven useful where rapid, portable genetic analysis is required; examples include crime scenes [136] and extraterrestrial rovers [137]. Applying these device designs to ultrahigh throughput tasks requiring 10,000 separation channels becomes increasingly complex when considering interfacing dozens of thermally-controlled substrates with macro-scale equipment for loading, separation, detection, and collection.

Capillary array electrophoresis instrumentation containing several hundred capillaries has been commercialized (e.g., Applied Biosystems, Foster City, CA; Amersham Biosciences Corp., Piscataway, NJ; SpectruMedix, LLC, State College, PA). Capillaries are typically 300-1000 mm long, with 360 μm outer diameter (d_o) and 75 μm inner diameter (d_i) (e.g., Polymicro TSP075375). The capillary rows in such instruments are typically purchased as pre-packaged assemblies, with 16 to 96 capillaries manually bonded in a linear array onto support brackets [Personal communication with Joe Macomber, Polymicro Technologies, Inc., 2006]. Improvements to the optical properties of these linear arrays have enabled rows containing as many as 550 capillaries [138].

Two-dimensional arrays containing up to 32 capillaries have been demonstrated by Zhang *et al.* [139], complete with DNA sequencing data from all capillaries. In their impressive work that could practically be scaled to 1,000 capillaries, the capillaries were permanently bonded with epoxy into drilled holes in alignment plates. Capillary alignment was limited to hole clearance, sufficient for the detection method

used. Analytes were detected fluorescently in a sheath flow from the capillary tips. Fluorescence was excited with an elliptically shaped laser beam that illuminated all sample streams simultaneously from the side and the emission was captured from the capillary ends by a CCD following a wavelength dispersing prism. From the teachings of Zhang *et al.*, we learned that this end-of-column detection technique could work well for arrays containing as many as 1,000 capillaries; arrays with 10,000 capillaries would likely require a 10 W laser, sheath flow volume of approximately 50 L, and bulky and expensive optical elements.

The assembly and constraint of tightly-packed 2-D arrays of many thousands of capillaries, as required for the pangenomic mutation discovery study proposed, is impractical using these approaches. Capillary arrays for ultrahigh throughput applications must be densely packed (maximum 1,000 μm intercapillary spacing) into a 2-D array with the capillary ends organized regularly to enable access. Specifically, both of the capillaries' ends should be aligned in a plane with less than 100 μm axial variation for interfacing with planar, microfabricated loading and detecting devices. In contrast with previous work, permanently bonding both aligned ends of a 10,000 capillary array would be extremely difficult. Assembly errors during this process would be permanent and costly. A replaceable constraint method would allow capillary tip alignment axially and radially and also permit capillary replacement for periodic array refurbishment. Radially, the capillary tips should not contact their neighbors and permit isolated immersion in sub-μL wells, channels, and the like, which are particularly useful for collection of eluted fractions of DNA. Isolated wells spaced 1,000 μm apart, 900 μm wide would allow 360 μm capillary tip radial variation of 270 μm. The capillaries should be held rigidly in all translational degrees-of-freedom, as forces on the order of 1 N are expected to occur during access to and manipulation of the array. To control the temperature of the capillaries during analyte separation in the range of 50-80 °C, as is commonly required (see Chapter 2), we require free standing, parallel capillaries rather than a closely packed bundle.

In this paper, we describe the design, manufacture, and testing of a constraint device for arrays of 100 to 10,000 capillaries and the metrics used to evaluate its

performance. We also present the design and evaluation of a semi-automated assembly machine for loading capillaries into the constraint device. Capillary array functionality is demonstrated with detection of electrophoresed DNA molecules in concentrations useful for DNA sequencing and mutation detection applications in Chapter 8.

6.2 Materials and methods

6.2.1 Capillary constraint device

Conceptual design

We set out to design a device which could constrain very large arrays of densely packed capillaries. Shown in Figure 6-1, the device consists of a sandwich of steel, silicone, and steel that contains an array of holes. Capillaries are inserted through the holes, spaced 1 mm apart, and then the sandwich is clamped. This clamping compresses and deforms the silicone to seal around the capillaries and align them radially. Errors in the initial capillary position are averaged by elastic deformation of the silicone. The design has the additional advantage of high load capacity and stiffness in the x,y plane (see Figure 6-1) [135].

This design has a number of advantages over previous methods of assembling capillary arrays as well as semiconductor manufacturing of etched channels in a planar substrate. The device can accommodate a single 1-D row of capillaries, as well as 2-D arrays containing thousands of capillary tubes, while permitting replacement. The capillaries can be aligned axially and clamped *en masse*. The constraint device seals against pneumatic and hydraulic pressure. Pneumatic pressure application at one end of the array is required for injection of DNA sieving media and/or hydrodynamic loading of analytes. Hydraulic pressure sealing between constraint devices located at either end of the array (see Figure 6-1*b*) can permit flow of a fluid for thermal control. The free-standing capillaries are amenable thermal control in a cross-flow water heat exchanger, due to inherently low thermal capacity and surrounding space

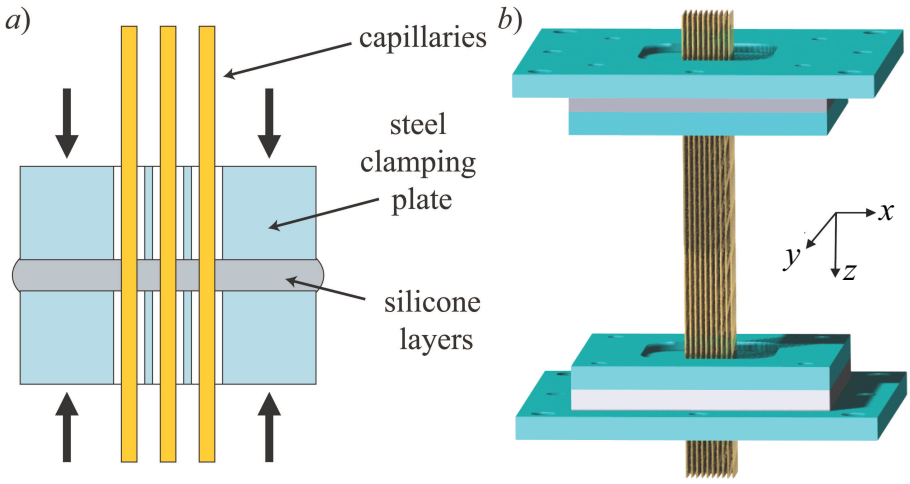

Figure 6-1: Schematic of the constraint device (a) and its application to constraint of a 100 capillary array (b). After capillaries are loosely inserted through holes in the device, clamping deforms the silicone layers to locate and seal around the capillaries.

for flow. In addition, the constraint devices accurately and repeatedly locates an array of capillaries locally and globally, essential for interfacing to loading, detection, and collection devices.

To design the parts for the constraint device, we performed a mechanics analysis using closed-form equations and Finite Element Analysis (FEA). When the sandwich is compressed by a bolt pattern, a resulting pressure compresses the silicone layers and causes the steel plates to bulge outward. We desire that the steel plates do not deform elastically by more than 15% of the silicone deformation, so that the plates remain relatively planar for capillary tip alignment as well as apply sufficient clamping pressure over their entire surfaces. The silicone deformation is calculated from

$$\delta_{\text{silicone}} = \frac{LT}{AEnKd}, \tag{6.1}$$

where the silicone properties are thickness L , cross-sectional area A , elastic modulus E , and the compressive force is defined by bolt torque T, number of bolts n, bolt diameter d, with torque coefficient $K \approx 0.2$ [140]. For a 100 capillary constraint device, four 5 mm diameter bolts with 1 N·m of torque compressing 4.8 mm thick silicone with 0.0025 m^2 area and modulus of 5×10^6 Pa gives $\delta_{\text{silicone}} \simeq 1.5$ mm. A stainless

steel type 304 plate of 3 mm thickness only deflects 22 μm under this pressure of 1.6 MPa, so $\delta_{\text{steel}}/\delta_{\text{silicone}}$=1.4%. For a 10,000 capillary constraint device, the same clamping pressure requires a stainless steel plate type 440C of thickness 3.5 mm and nine bolts to reduce the deformation to 0.14 mm, or 9.3%, as determined using FEA. Thus we settled on these bolt, steel plate, torque, and silicone specifications for the design of our constraint devices.

Fabrication

Manufacturing is achieved using a combination of microelectrode discharge machining (EDM) and laser micro-machining processes. These parts are shown in Figure 6-2. To fabricate the steel plates, a tellurium copper pin array is first cut on a wire EDM (Charmilles, Robofil 1020si) after selective milling to permit clamping bolts. The pins are sized 350 μm $\times$ 350 μm, spaced 1 mm apart in a square array. For arrays containing 10,000 holes, the wire EDM job is performed with 0.254 mm wire in 36 hours with one part rotation of 90°. The pin array is then used as the tool in a die sink EDM (Charmilles, Roboform 30) to etch 10,000 square holes simultaneously in hardened 3.5 mm thick 440C stainless steel. An overburn setting of 90 μm removing an average of 4 mm^3/min of steel required 40 hours to etch through the part. The hole array in the silicone layer is pre-drilled with a 75 W CO_2 laser (Trotec Speedy 100). This process yields 200 μm diameter holes located with 25 μm accuracy.

Performance evaluation

The accuracy and repeatability of capillary location in the array are critical performance criteria. By accuracy, we desire to know how close the capillaries are to their desired locations after bias error compensation. For repeatability, we ask how close the capillaries are to their previous locations after the constraint device is reclamped (e.g., after capillary replacement). These questions are relevant locally and globally, axially and radially, as functions of the length that the capillary protrudes from the constraint device.

We established an experimental methodology to address these questions. The

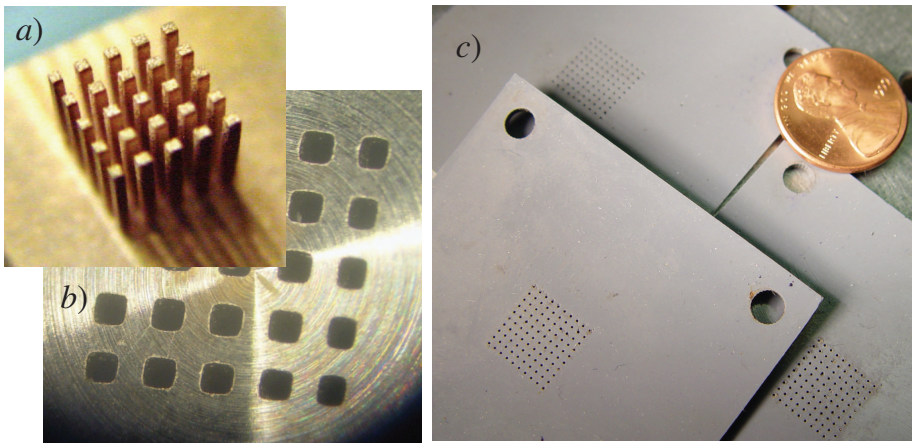

Figure 6-2: Parts for a capillary array constraint device. A tellurium copper pin array (a) is used to etch square holes in the steel plates (b), made using a combination of wire and sink electrical discharge machining, while the silicone parts (c) are laser cut. These parts, with 25 to 100 holes, are shown for clarity; parts with 2500 to 10,000 pins/holes have been manufactured.

notation is indicated in Figure 6-3. First we measured the positions of the capillary tips radially and axially within the array using a microscope (Zeiss, Stemi SV 11). Our axial measurements were physically limited to single rows of capillaries. Next we performed a least squares fit to position a best-fit grid (radial measurements) or line (axial measurements) with 1 mm period. As shown in Figure 6-3, we then calculated the lateral errors, Δx_{ij}, Δy_{ij}, of the actual capillary positions, x_{ij}, y_{ij}, from the best-fit grid, as well as the axial errors, Δz_p, from the best-fit line. This was performed for $I \times J$ positions in the x,y plane and K positions in the z direction. This process was repeated for T trials, where each trial t represents an reclamping of the capillary constraint device.

The axial accuracy of the capillary tips is then given by

$$\rho_a = \frac{1}{T}\sum_{t=1}^{T}\sqrt{\frac{1}{K}\sum_{k=1}^{K}\left(\Delta z_{kt} - \overline{\Delta z_t}\right)^2}, \tag{6.2}$$

where we have averaged the standard deviations of the tip errors for all capillary positions across all trials. The axial repeatability, σ_a, requires a slight switch in the order of operations to instead average the standard deviations of the tip errors for all

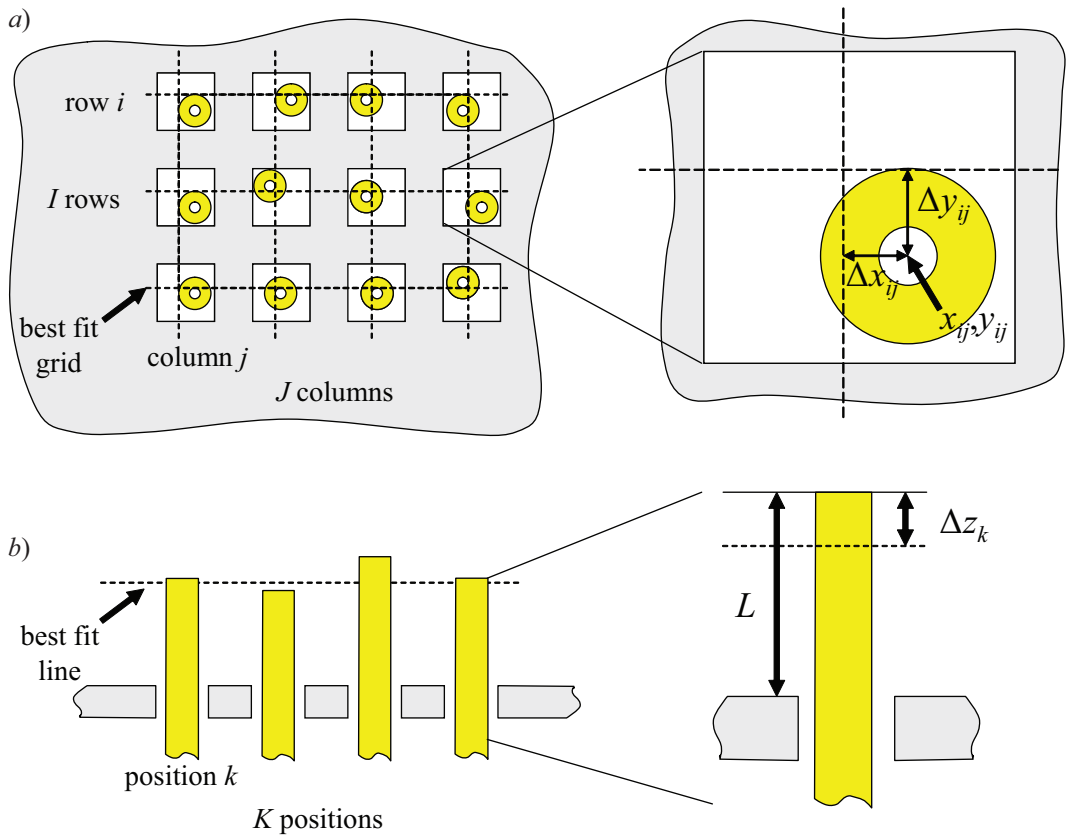

Figure 6-3: Notation for measurement of the accuracy and repeatability of the capillary tip alignment in the constraint device, both radially (a) and axially (b).

trials across all positions, as in

$$\sigma_a = \frac{1}{K}\sum_{k=1}^{K}\sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(\Delta z_{kt} - \overline{\Delta z_k}\right)^2}. \tag{6.3}$$

The radial accuracy and repeatability require a transformation between Cartesian and polar coordinates. To determine the radial accuracy, ρ_r, we apply the transformation $\sqrt{\rho_{\Delta x}^2 + \rho_{\Delta y}^2}$, as given more explicitly by

$$\rho_r = \sqrt{\left(\frac{1}{T}\sum_{t=1}^{T}\sqrt{\frac{1}{IJ}\sum_{i=1}^{I}\sum_{j=1}^{J}\left(\Delta x_{ijt} - \overline{\Delta x_t}\right)^2}\right)^2 + \left(\frac{1}{T}\sum_{t=1}^{T}\sqrt{\frac{1}{IJ}\sum_{i=1}^{I}\sum_{j=1}^{J}\left(\Delta y_{ijt} - \overline{\Delta y_t}\right)^2}\right)^2}. \tag{6.4}$$

Thus, the lateral accuracies in the Cartesian space are computed by averaging the standard deviation of the capillary tip errors for all positions across all trials. This preserves the spread of Δx, Δy more correctly than converting directly to $\Delta r_{ij} = \sqrt{\Delta x_{ij}^2 + \Delta y_{ij}^2}$ and taking the standard deviation of Δr_{ij}. To determine the radial repeatability, we switch the order of operations so that we have the average of the standard deviations of the trials across all of the positions. This radial repeatability is

$$\sigma_r = \sqrt{\left(\frac{1}{IJ}\sum_{i=1}^{I}\sum_{j=1}^{J}\sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(\Delta x_{ijt} - \overline{\Delta x_{ij}}\right)^2}\right)^2 + \left(\frac{1}{IJ}\sum_{i=1}^{I}\sum_{j=1}^{J}\sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(\Delta y_{ijt} - \overline{\Delta y_{ij}}\right)^2}\right)^2}. \tag{6.5}$$

The radial accuracy, ρ_r, was determined locally for a set of 12 capillaries, as well as globally for a set of 100 capillaries. In all measurements, the tips of the capillaries were aligned in a plane, but the length of the capillary protruding from the constraint device was varied from L=0-12 mm. The radial repeatability, σ_r, was determined from T=3 trials, on a set of 12 capillaries protruding L=9 mm. The axial accuracy and repeatability were measured from K=5 capillary positions reclamped T=3 times with a fixed protrusion of L=9 mm.

Since we intend for the capillary constraint device to be useful for up to 10,000 capillaries, we require satisfactory performance at 3ρ accuracy and 3σ repeatability, which statistically accounts for 9,973/10,000 capillaries, assuming a normal distribu-

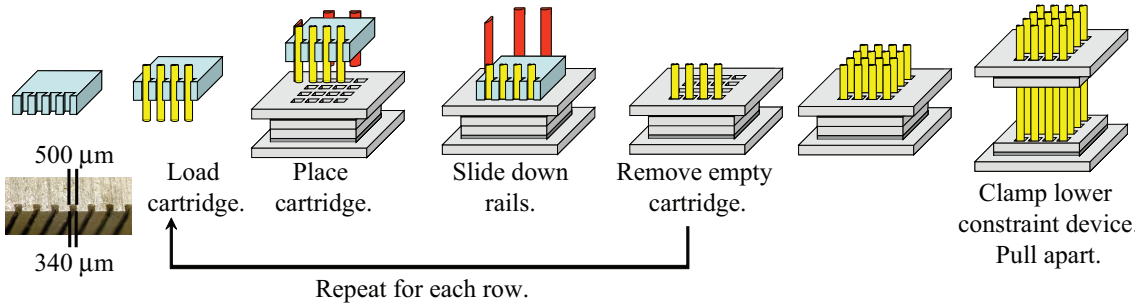

Figure 6-4: Capillary array assembly procedure using the semi-automated machine. Row-by-row assembly is performed using a cartridge to align and insert a row of capillaries into corresponding holes in the constraint device, as guided by rails. A photograph of the trapezoidal troughs in the cartridge is shown with critical dimensions.

tion of errors. Additionally, uncertainty in all of the measurements arises from the resolution of the images captured through the microscope at 2.5× to 10× magnification. Uncertainty for axial accuracy, local radial accuracy, radial and axial repeatability, is 5 μm. For global accuracy, images captured at the lower magnification of 2.5× have a resolution uncertainty of 7 μm.

6.2.2 Capillary assembly

To assemble the capillary arrays efficiently, we designed and constructed a semi-automated assembly machine for row-by-row assembly. The procedure for assembling arrays using this machine is depicted in Figure 6-4. The key feature of this design is a stainless steel cartridge that contains a row of evenly spaced trapezoidal troughs accurately manufactured by EDM to <5 μm process variation. Each trough pinches a capillary upon application of a preload force, thus constraining it for insertion into the array. This is similar to a technique utilized in the optical fiber industry [141]. Our cartridge is easily filled with capillaries by rolling a bundle of them across the troughs.

The machine is designed for semi-automated assembly of arrays of 100 capillaries, as shown in Figure 6-5. In operation, a linear stage is automatically indexed to successive row positions in the array. At each position, the loaded cartridge is repeatedly attached to the stage using kinematic couplings and translated to insert the capillary

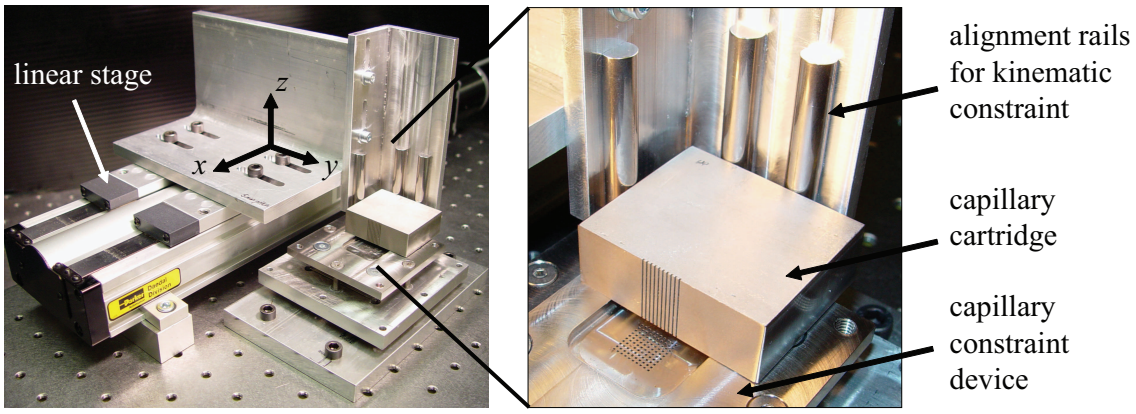

Figure 6-5: Capillary array assembly machine. An automated linear translation stage indexes to each row. The alignment rails position the loaded cartridge for insertion of a row of capillaries (*not shown*) into the constraint device.

row.

We desired to characterize the performance of the assembly machine by comparing manual to semi-automated assembly times. In addition we measured the cartridge's ability to align the capillaries for insertion into the constraint device, $\rho_{\Delta x}$ and $\rho_{\Delta y}$, from Equation 6.4. The Cartesian coordinate system was retained as non-random differences in $\rho_{\Delta x}$ and $\rho_{\Delta y}$ were noted. This accuracy was determined for various lengths of protrusion, from 2.9 to 15 mm.

We also designed, manufactured, and tested another assembly technique, which was subsequently abandoned. The design consists of a staggered, tapered hole array on an x-y translation stage through which capillaries can be quickly aligned and inserted into a corresponding array hole in the constraint device. The x-y stage is amenable to indexing to successive array holes. This design was abandoned due to the serial nature of assembly as compared to the more parallel row-by-row assembly.

6.3 Results

6.3.1 Constraint device

We have designed, manufactured, and tested capillary array constraint devices for arrays of 100 to 10,000 capillaries, and subsequently assembled arrays in them. The

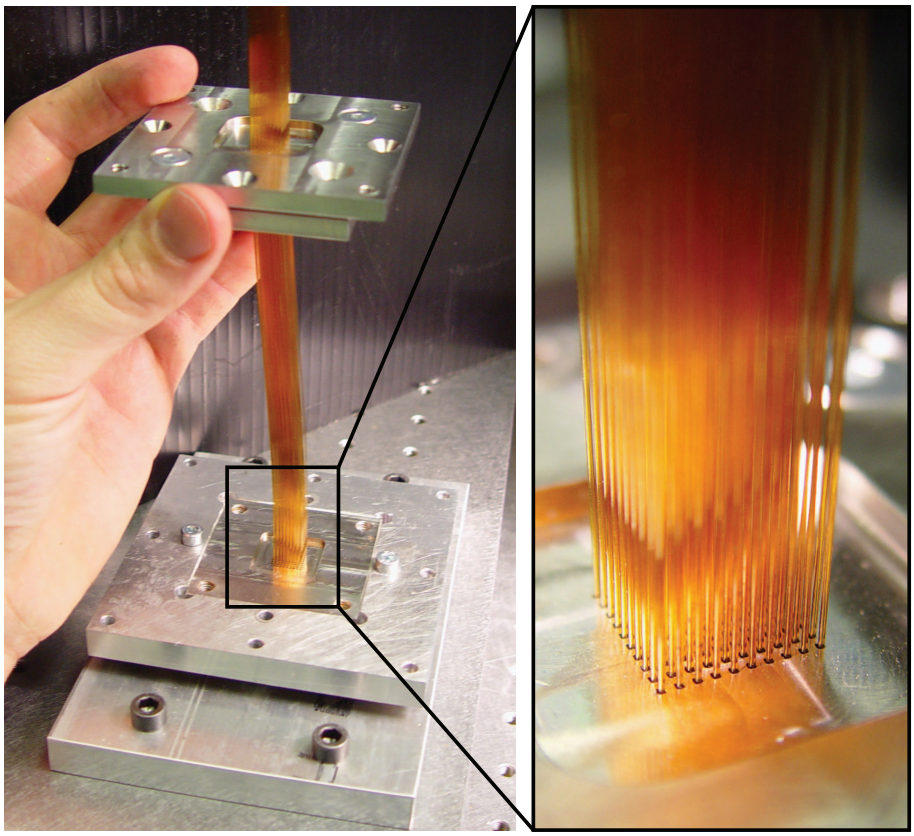

Figure 6-6: Assembled 100 capillary array. Capillaries 300 mm long, 360 μm outer diameter are assembled into a pair of constraint devices, one at each end. The capillary tips are aligned to enable access for analyte loading and detection. The array can be assembled in about 12 minutes.

capillaries are 360 μm outer diameter, 75 μm inner diameter, 300 mm long are arrayed with 1 mm spacing. Figure 6-6 and Figure 6-7 show arrays of 100 and 10,000 capillaries, respectively.

After assembling the arrays, we proceeded to measure the capillary alignment accuracy and repeatability. The local radial accuracy for a neighborhood of 12 capillaries is shown as a function of protrusion length in Figure 6-8. At protrusions larger than 10 mm, there is a statistical likelihood that two capillaries will touch each other.

We also measured the radial errors of all of the capillaries in a 100 capillary array at a protrusion of $L = 7$ mm, as given by

$$\Delta r_{ij} = \sqrt{\Delta x_{ij}^2 + \Delta y_{ij}^2}. \tag{6.6}$$

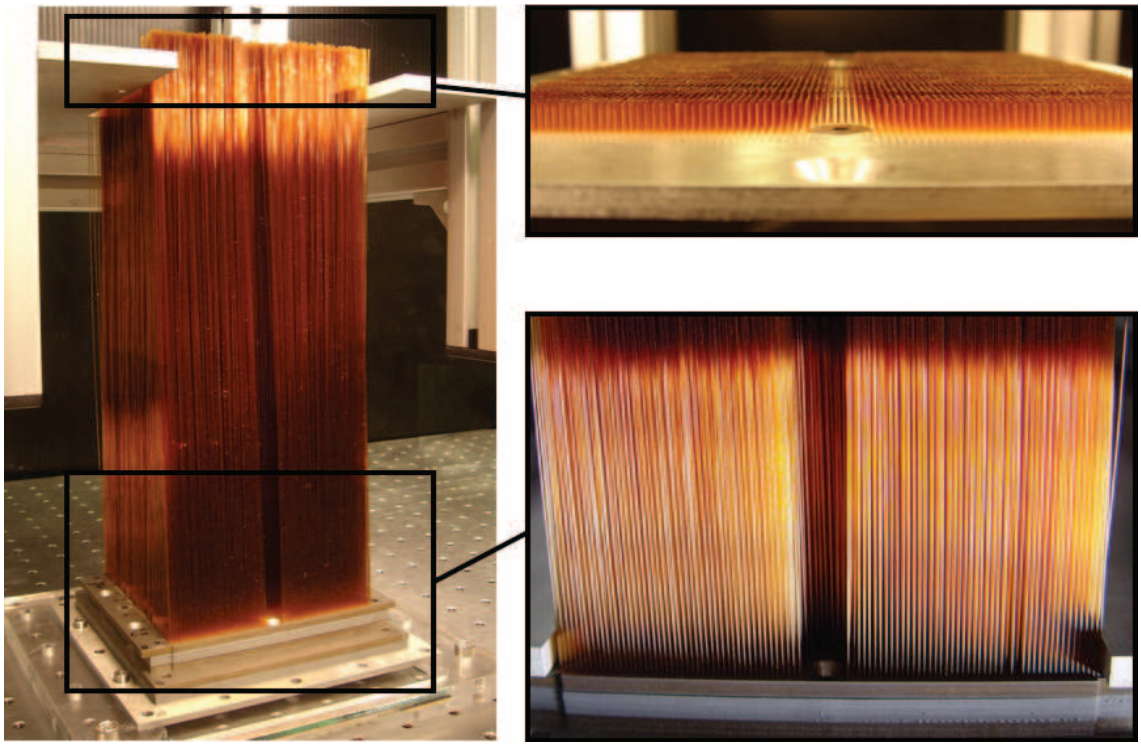

Figure 6-7: Assembled 10,000 capillary array. As a proof-of-concept, we assembled the 10,000 capillary array in about 100 hours, using a pair of constraint devices. Manufacturing of the constraint devices required about 500 hours. In this work, actually 9,867 capillaries were successfully inserted into an array with 10,060 holes.

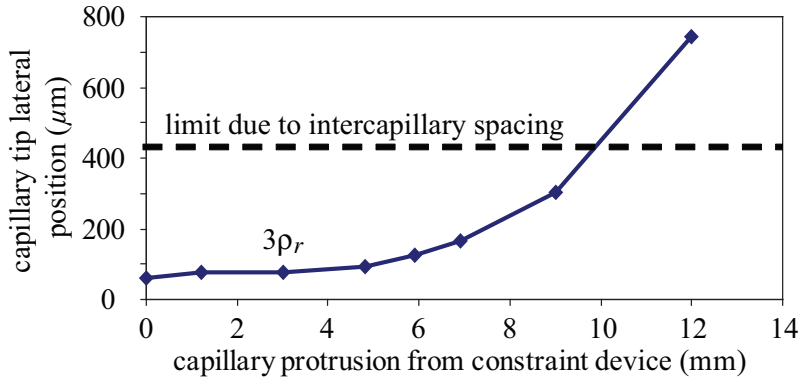

Figure 6-8: Local radial accuracy of 12 capillary tips as a function of length of protrusion from the capillary constraint device. For protrusions over 10 mm, capillary neighbors may touch each other radially.

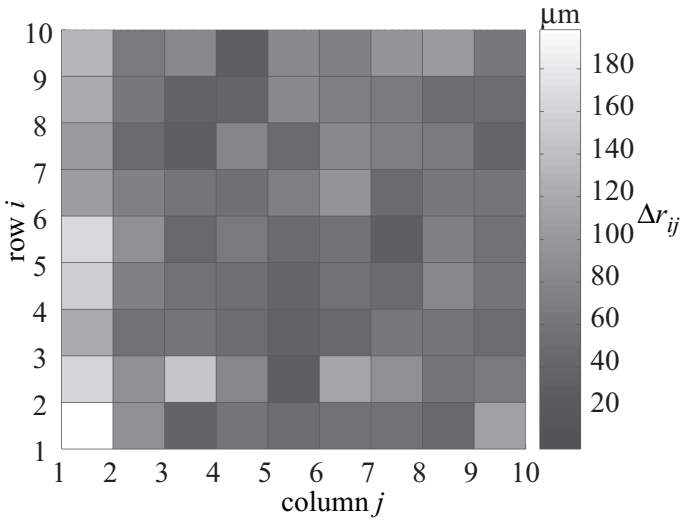

Figure 6-9: Global radial accuracy error map of 100 capillary tips held in the capillary constraint device, $L = 7$.

These resulting global error map is shown in Figure 6-9. The distribution of Δx_{ij} and Δy_{ij} errors is Gaussian (*not shown*), thus they can be assumed random. The $3\rho_r$ (radial accuracy) over the entire array is 265±7 μm. The largest radial errors in the global map are at the perimeter. Removing these perimeter capillaries reduces $3\rho_r$ to 223±7 μm. Radial repeatability for the array was also measured to be $3\sigma_r$=94±5 μm (T=3, L=9). Axial accuracy and repeatability were independent of protrusion length. They were measured and found to be $3\rho_a$=$3\sigma_a$=11±5 μm (T=3, L=9).

When the constraint device is clamped, the silicone deformation of 1.5 mm retains the capillaries under and an axial force of 3±0.5 N per capillary. This is 3-15× greater than the force required to puncture a thin silicone septum useful for sealing a buffer reservoir. When clamped, the array can readily seal against 700 kPa N_2, or 250 mm water head. By unclamping the constraint device, we have demonstrated capillary replacement with the 100 capillary array. Any capillary in the array can be identified by position, removed, and a new one rethreaded.

6.3.2 10,000 capillary array testing

Detection of optical emission from the 10,000 capillary array has been performed to verify its fidelity and capillary alignment. Figure 6-10 shows the 10,000 capillary

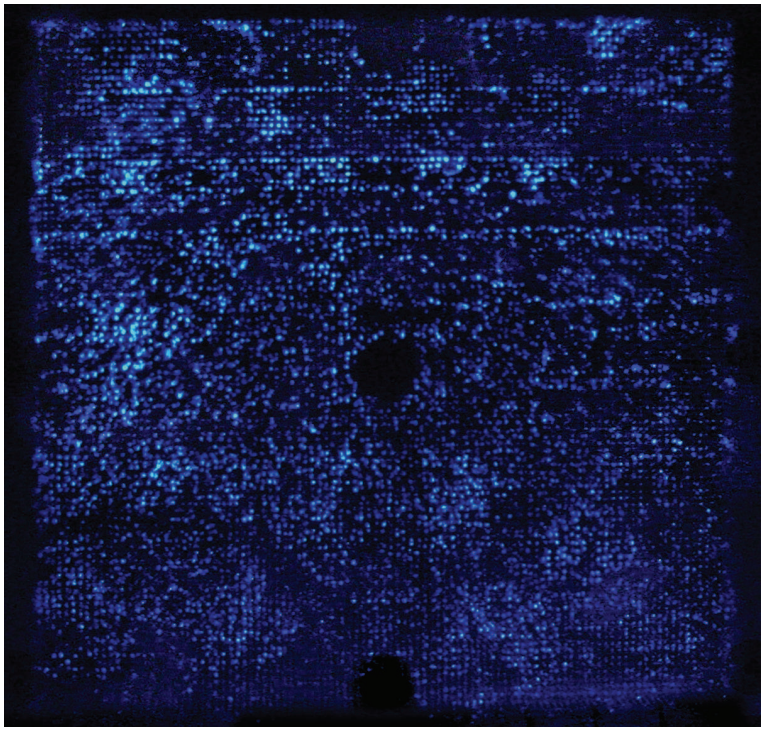

Figure 6-10: Points of light emanating from the 10,000 capillary array demonstrating the capillaries fidelity and alignment. Diffused light was used to illuminate one end of the capillary array and image was captured at the other.

array that we have manufactured under transillumination, imitating the conditions of fluorescence emission.

6.3.3 Assembly

We have used the assembly machine for arrays containing up to 100 capillaries. Assembly using the semi-automated row-by-row strategy requires approximately 0.2 hr, as compared to 0.5 hr for manual assembly, a 56% reduction. Following Taylorism [142], we conducted a time and motion study (see Table 6.1) and recognized that substantial time is saved in the capillary placement and tip insertion, as it is done *en masse* for an entire row. With assembly of longer rows of capillaries, challenges were encountered with retaining the capillaries in the troughs during insertion and aligning the capillaries with the holes in the constraint devices. Thus, though the semi-automated assembly machine was quite useful for assembling 100 capillary arrays, we opted for manual assembly of larger arrays. We expect that additional design

tasks		manual assembly (per capillary)	10^2 capillary array semi-automated assembly (per row)	10^3 capillary array semi-automated assembly (per row)	10^4 capillary array semi-automated assembly (per row)
capillary acq.	(s)	7	15	20	27
placement	(s)	3	3	3	3
tip insertion	(s)	2	2	2	2
feed through	(s)	6	60	80	107
total	(s)	18	80	105	138
			(per array)	(per array)	(per array)
total, manual	(hr)		0.5	5	50
total, semi-auto.	(hr)		0.2	0.9*	3.8*
improvement	(%)		56%	81%	91%

*estimated

Table 6.1: Capillary array assembly times for manual and semi-automated assembly, for arrays of $I \times J$ capillaries.

improvements in future work will enable a reduction in assembly time of 81% for a 1,000 capillary array (from 5 hr to 0.9 hr) and 92% for a 10,000 capillary array (from 50 hr to 3.8 hr). These results are summarized in Table 6.1.

The alignment of the capillaries in the cartridge for assembly has been measured as well. Shown in Figure 6-11, the $3\rho_{\Delta x}$ and $3\rho_{\Delta y}$ accuracies are measured as a function of capillary protrusion length from the cartridge. Even at protrusions of 15 mm, the capillary tips can glide into the holes in the constraint device without hindrance. The accuracy in the x direction is somewhat independent of protrusion length due to the asymmetrical constraint offered by the trapezoidal cross-section. In that direction the capillary rests against the steel face of the cartridge.

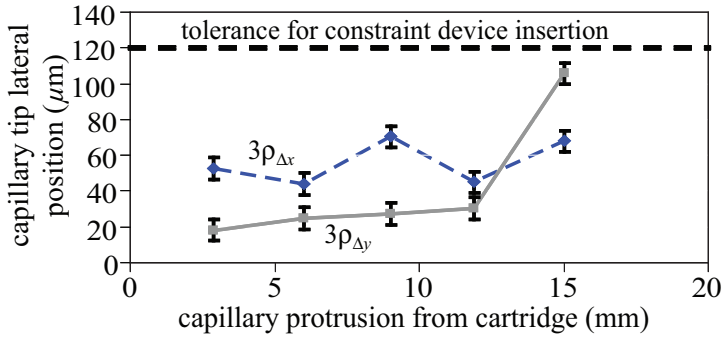

Figure 6-11: The cartridge performance as measured by the accuracy of the capillary tip position vs. length of capillary protruding from the cartridge. The threshold for insertion into the capillary array constraint device is also shown.

6.4 Discussion

We have demonstrated a facile means for constraint of hundreds to thousands of capillaries. In a neighborhood of a dozen capillaries protruding less than 6 mm from the constraint device, 99.73% of their tips do not deviate by more than 100 μm radially from their desired positions on a grid with 1 mm period. If protrusion greater than 6 mm is desired, as in fluorescence detection from the 75 μm central lumen of the capillary, a secondary guide with conical holes can be utilized to redirect the capillary tips to within tolerance. In Chapter 5 have report on such guide technology, termed a passively aligned microfluidic connector. This device has the ability to align the tips to 25 μm radial accuracy over the entire array for loading and detection of analytes and can be used in tandem with the constraint device. For an entire 100 capillary array, the capillary tips have radial accuracy errors of 265$\pm$7 μm when protruding 7 mm, which is sufficient for entrance into 900 μm wells spaced 1,000 μm apart. This error can be reduced by shortening the protrusion or, again, utilizing a secondary guide if lengthy protrusions are required for access. When the constraint device is reclamped, as in capillary replacement, 99.73% of the capillaries return to their original radial positions within 94$\pm$5 μm. Thus, array refurbishment does not hinder performance. Axially, the capillaries are aligned in a plane and stay that way after reclamping, to within 11$\pm$5 μm. For loading from a planar array of sample wells (described in Chapter 4) or detecting fluorescence through a transparent window enclosing an

electrolyte solution (described in Chapters 7 and 8), this axial variation is virtually negligible.

In previous capillary array constraint technologies, replacing an inoperable, relatively inexpensive (∼$5) capillary after merely 300 runs or bad fortune typically requires replacement of the entire bundle of 16-96 capillaries (∼$2K to $5K), or a substrate with potentially hundreds of channels. In our device, the capillaries are replaceable, which reduces the cost of refurbishing an array by 100× to 1,000× as compared to replacing the array. For the 10,000 capillary array, such refurbishment would prudently be performed after a significant portion of the array is non-functional. This device could contribute to consumable cost and downtime reduction for current capillary array electrophoresis instrumentation, as well as enabling a new generation of ultrahigh throughput instruments with 2-D arrays of many thousands of capillaries.

Capillary array assembly time and cost is dramatically reduced by using the semi-automated assembly machine, as compared to manual assembly, since it eliminates time-consuming handling. The simple, accurate, and repeatable procedure for row-by-row assembly makes feasible the rapid manufacture of arrays of hundreds of capillaries. In this assembly machine, the cartridge positions the capillaries more accurately than the 120 μm tolerance required with allowable protrusions from 0 to 15 mm. Applications for thousands of capillaries for ultrahigh throughput electrophoresis instrumentation are probable, but have not been demonstrated in this work. Detection of fluorescently-labeled DNA by end-column detection interfaced to the capillary array is described in Chapter 8. These result will show that the capillary arrays are functional.

Chapter 7

End-column fluorescence detection system

Ultrahigh-throughput denaturing electrophoresis (>5000 channels) can enable pangenomic scans of the human genome. Instruments with such high throughput present signal detection challenges. We report on a two-dimensional capillary array fluorescence detection system featuring a light-emitting-diode (LED) array, lenslet array, and charged-coupled-device (CCD) detector. The LED array provides scalable excitation illumination. A typical lenslet array from the device increases the field intensity by >70,000X and subsequently collects emission with a 0.16 numerical aperture (NA). This detection technology is scalable to arrays with greater than 5000 channels and is expected to detect fluorescently-labeled electrophoretic "peaks" containing 10^7 or more molecules. Applications are further described in Chapter 8.

7.1 Introduction

Detection of mutant DNA segments within each separation channel is the foremost challenge in the design of the UTMS. We seek to measure fluorescent peaks containing 10^7–10^{10} molecules across 10,000 independent channels. Numerous capillary array detectors have been developed, but none satisfy these functional requirements.

Initially Zagurski [143] and later Mathies [144, 112] described techniques for scan-

ning across a planar capillary array to detect fluorescence. These designs utilize a laser and moving confocal microscope to serially interrogate the capillaries. This technology has proven effective for etched channel, micro-electro-mechanical system (MEMS) based sequencers with up to 384 channels [90] and has also been utilized in 768 lane sequencers created from etched glass plates [91].

Yeung [145, 146] reported a detection system in which a laser beam focused to a line continuously illuminates the side of a planar capillary array and a CCD camera detected fluorescence from all of the capillaries simultaneously. Quesada [147] reported a similar system which utilizes fiber optics to illuminate the capillaries. These technologies have been extended to non-planar, two-dimensional capillary arrays [148] and other manifestations in which the laser excitation passes through multiple capillaries [149, 150]. These techniques require high laser powers either to distribute the beam across the full array and/or to overcome beam weakening as it passes through each capillary by reflections and refractions.

Kambara [151, 152] and Dovichi [153, 154, 139] have pursued a side column detection strategy which is similar to Yeung and Quesada's, except that the detection is performed after the fluorescent molecules have exited the capillaries into a sheath flow cuvette. The number of capillaries observable was limited by the distance over which the laser beam was collimated ($\sim$4 mm for a 488 nm laser). In a two-dimensional array configuration, limitations include a costly optical system to collect the emitted fluorescence efficiently; large volumes of uniform, degassed, particle-free buffer solution flow; and laser powers of greater than 1 W required for greater than 1,000 capillaries. Other post-column detection techniques, such as Bottani [155], utilize electric field focusing. Unfortunately, none of these published approaches appear to offer the promise of scanning 10,000 capillaries with the detection criteria we believe to be required for a serious pangenomic scan for mutations in large human populations. Thus, we set out to design and build a detector that would continuously detect fluorescent molecules just prior to the time they exited the capillary gel into an aqueous solution of electrolyte.

Our device appears to overcome the limitations of previous designs and enables

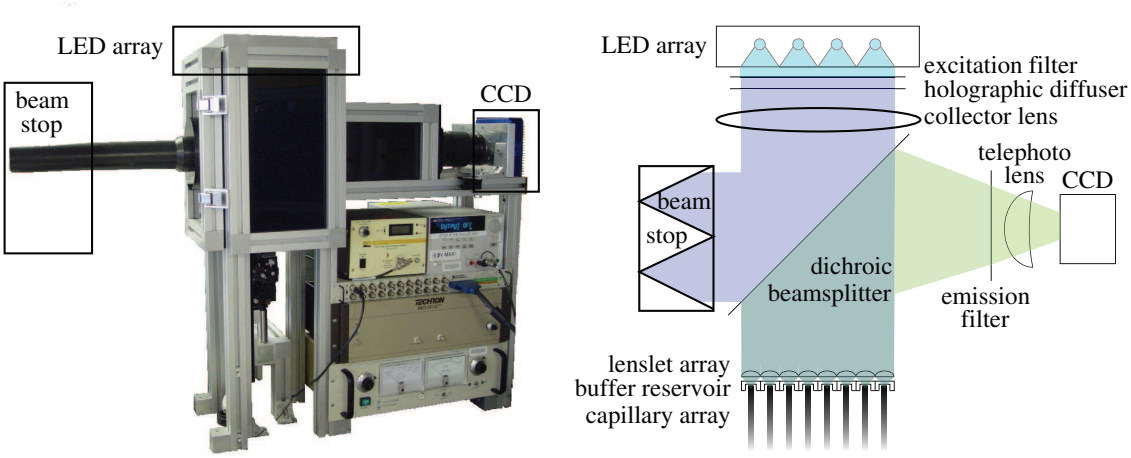

Figure 7-1: Photograph (*left*) and schematic (*right*) illustrate the key features of the optical detection system.

detection from arrays containing >5000 capillaries. Fluorescence is excited and detected from the ends of the capillaries. We utilize an LED array that can scale to illuminate any size area uniformly with sufficient fluence to excite biological fluorescence. A lenslet array registered to the capillary tips is used to focus excitation light as well as collect and project the emitted light to a CCD detector. This lenslet array can also scale to accommodate many thousands of capillaries, thus eliminating the need in previous work for costly or complex scanning optics. This system is furthermore capable of integrating with real-time mutant fraction collection. The ability to sequester such fractions of DNA eluted from the capillary column can be advantageous for subsequent processing steps, such as sequencing.

7.2 Apparatus overview

The two-dimensional end-column detection system is shown in Figure 7-1. In operation, light propagates from the LED array through a dichroic beamsplitter to a lenslet array. Each lenslet focuses a portion of the incident beam onto a capillary's tip. If a fluorescent molecule is present within the capillary's tip, the wavelength shifted emission is captured by the lenslet and propagated to the CCD for imaging.

A 42 LED array provides the illumination. After spectrally filtering, the intensity

output of 13.2 W/m^2 peaks at 482 nm with a corresponding bandwidth, W_h, equal to 20 nm. The LED's (Lumileds, model Luxeon III) are water-cooled to 7 °C using a serpentine heat exchanger to keep the intensity and wavelength output constant. Individual collimators are mounted to the LEDs to direct the their illumination.

The set of spectral filters (Omega, Models 1073, 3084, 3105, 512DRSP) include excitation and emission filters and well as the dichroic beam splitter. These were selected to optimize signal-to-noise ratio (SNR) while maintaining a bright emission, as further described in the next Section. Unconventionally, the dichroic beam splitter transmits the excitation and reflects emission due to optimal LED and CCD packaging constraints. An emission filter is placed in front of a telephoto lens to minimize off-axis illumination (see Figure 7-1).

A holographic diffuser (Edmund Industrial Optics) was implemented to spread the illumination across the field-of-view more uniformly. This element is comprised of a polycarbonate sheet etched to introduce complex and irregular phase shifts to the incident wavefront [156]. This element has >50% transmission and efficiently distributes the intensity with <10° divergence.

A collector lens (50 mm clear aperture, 100 mm focal length) concentrates the excitation intensity by an areal factor of 25. This component improves the signal-to-noise ratio of fluorescence detection by increasing the power density available for fluorophore excitation.

The intensity of background light that passes through the emission filters was reduced by 50% by designing and manufacturing a beam stop, as shown in Figure 7-1. This element consists of a set of opposed cones manufactured by 3-D printing with photo-polymerized resin, integrated to the beam path with poly-vinyl-chloride pipe.

The camera assembly includes a telephoto lens and CCD detector. The lens (Zeiss, model Apo-Makro-Planar T 4/120) features eight elements to minimize aberrations and is used to image the fluorescent molecules with a 0.25× magnification. Thus the system is a telescope. The CCD detector (Apogee, U10) is thermo-electrically cooled to −10 °C and has 2048×2048 pixels. This CCD enables a 1/3 Hz sampling rate.

The 16-bit well depth provides ample dynamic range to meet the 10^3 requirement, even with significant background subtraction; 12-bit images were typically gathered to expedite readout time.

A light tight enclosure surrounds the optical detection system to reduce background illumination. The enclosure reduces the CCD detection noise to 8 photons/pixel/s, including thermal noise, readout noise, and stray light.

7.3 Spectral transmission modeling and measurement

The propagation of electromagnetic radiation through the optical system can be modeled in numerous ways to yield insights about the design. In this Section, we consider the power as a function of wavelength to aid in the selection of spectral filters for maximizing the SNR of the detected fluorescence emission.

From the design shown in Figure 7-1, there are many components that have spectrally dependent properties. These include the LED array, excitation and emission filters, dichroic beamsplitter, fluorophore excitation and emission properties, and the CCD quantum efficiency. In this analysis, we will ignore the wavelength dependence of the fluorophore quantum yield.

We seek to choose an optimal set of these components to maximize our ability to detect fluorescence. For the fluorescent probe, we choose fluorescein dye (6-FAM) for its ready availability and low cost. In some experiments, dyes with nearly identical absorption/emission were used, such as AlexaFluor 488. More discussion of the choice of dyes and their application in available in Appendix B. For the LEDs we chose the highest power output available with the best excitation match to the fluorophores.

Given these two components, LEDs and dyes, we computed the power as a function of wavelength measured by the CCD for a variety of spectral filter combinations. In this algorithm, we consecutively multiply the LED power output, excitation filter transmission, dichroic transmission, fluorophore absorption. This energy is then

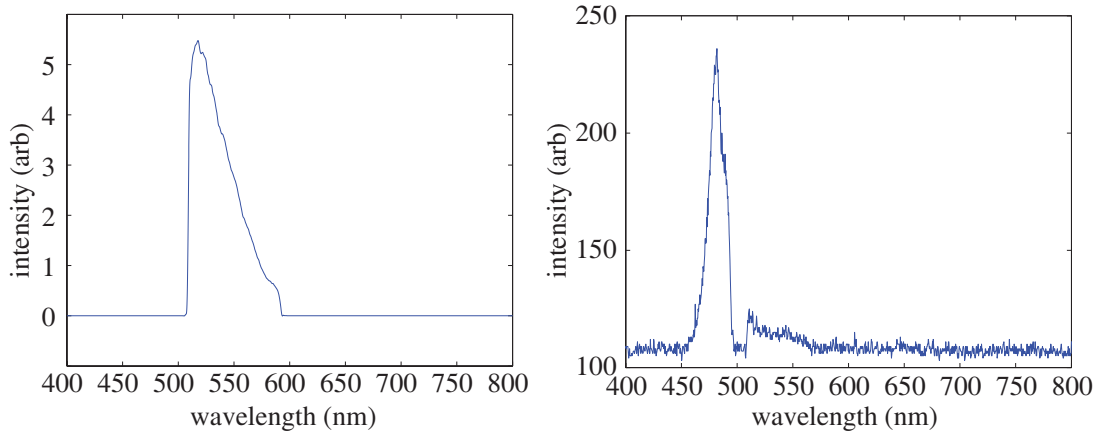

Figure 7-2: Simulated (*left*) and experimental (*right*) power spectrums measured by the fluorescence detection system. A set of excitation, emission, and dichroic filters were chosen from the simulation results and implemented for the experiment. The power measured in the 400-500 nm band is due to filter leakage and can be eliminated by subtracting a background spectral measurement without emission peak present.

wavelength shifted to the fluorophore emission spectrum, after uniform (not wavelength dependent) attenuation by a combination of fluor concentration, extinction coefficient, volume, and quantum yield. This emission spectrum is then consecutively multiplied by the dichroic reflection, emission filter transmission, and CCD quantum efficiency.

The simulation result for a filter combination is shown alongside of experimental measurement of the power spectrum using the modeled filters in Figure 7-2. Experimentally, more power was measured in the 400-500 nm band than expected. This can be attributed to leakage in the filters, most likely due to off-axis transmission and reflection. This problem was overcome by subtracting a background measurement, in which the emission signal is not present.

The excitation spectrum was further measured, as shown in Figure 7-3, to understand the total power available for exciting fluorescence. From these results, we see that 132 mW (13.2 W/m^2) of spectrally filtered, diffuse, incoherent illumination is available for fluorophore excitation across the 10,000 capillary array area (100 mm $\times$ 100 mm).

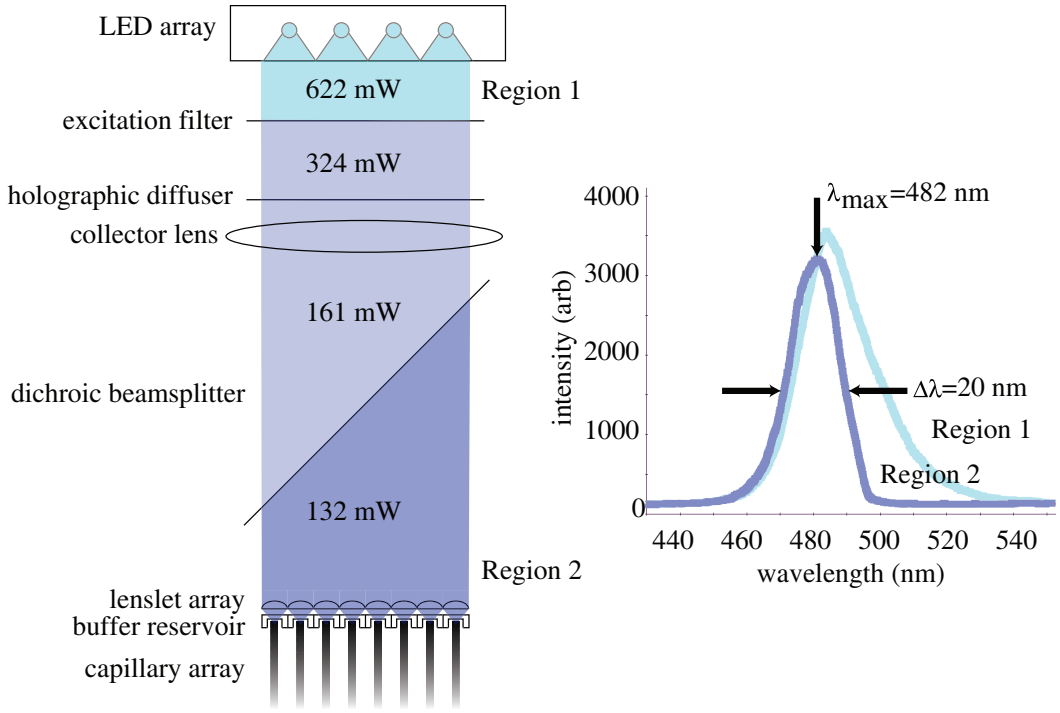

Figure 7-3: Experimental measurement of the spectrum and power of illumination available for fluorophore excitation across the 10,000 capillary array area. The attenuating effect of the spectral filters can be observed as well.

7.4 Lenslet array modeling

The lenslet array enables an efficient increase in excitation intensity, emission collection, and spatial magnification. A 25 lenslet prototype is shown in Figure 7-4. The benefits of the implementation of these light refracting elements are best conveyed using some optical calculations. The square of the numerical aperture (NA) of a lens is a measure of its light gathering power [157]. NA is defined as

$$\mathrm{NA} = n \sin\theta, \tag{7.1}$$

where n is the refractive index of the propagation medium (n_{air}=1) and θ is the half angle of the maximum cone of light focused up by the lens. The focal length, f, of a plano-convex lenses is defined as

$$f = \frac{1}{P} = \frac{R_c}{n_{\mathrm{lens}} - n_{\mathrm{air}}}, \tag{7.2}$$

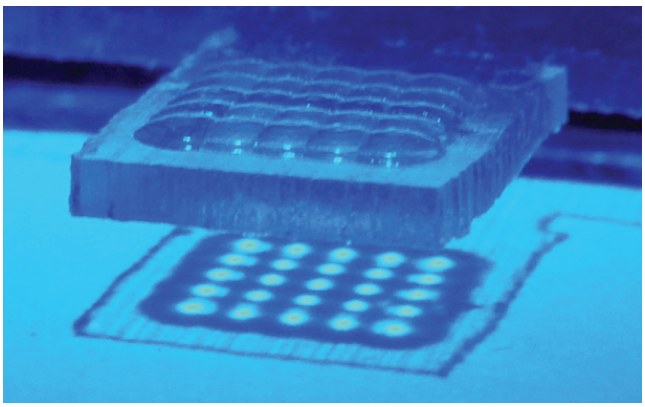

Figure 7-4: An injection-molded, polymethyl-methacrylate (PMMA) lenslet array photograph is shown. The lenslets each have a 1 mm aperture and 3 mm focal length. The corresponding focal spots from LED illumination are shown below the lenslet array.

where P is the lens power, n_{lens} is refractive index of the lens material (n_{PMMA}=1.49), and R_c is the radius of curvature of the lens. Combining these equations using geometry we have

$$\text{NA} = \frac{r}{\sqrt{r^2 + f^2}} = \frac{r}{\sqrt{r^2 + \left(\frac{R_c}{n_{\text{lens}} - n_{\text{air}}}\right)^2}}, \tag{7.3}$$

where r is the lens radius. We have manufactured lenses with r=500 μm and 500<R_c (μm)<3000, thus having 0.05<NA<0.44. The benefits of these lenses can now be computed. The first benefit conferred is the increase in the excitation intensity as the light is focused by the lenslet onto the capillary. The intensity will increase by the ratio of the areas of the lenslet to that of its focal spot. For a lens with collimated input, the focal spot width [158], w, is

$$w = 1.22\frac{\lambda}{\text{NA}}, \tag{7.4}$$

Excitation intensity increase	$\frac{A_{\text{lenslet}}}{A_{\text{focus}}}$	>70,000
Solid angle of emission collection	$\frac{\Omega_{\text{lenslet}}}{\Omega_{\text{telephoto lens}}}$	9.3
Spatial magnification	$\frac{A_{\text{lenslet}}}{A_{\text{capillary}}}$	178

Table 7.1: The use of lenslets greatly improves the end-column detection performance. The benefits conferred are increased excitation intensity, increased solid angle of emission collection, and spatial magnification of the emission. For PMMA lenslets with r=500 μm, NA=0.16.

so therefore the area ratio is defined as

$$\frac{A_{\text{lenslet}}}{A_{\text{focus}}} = \frac{\pi r^2}{\pi \left(0.61 \frac{\lambda}{\text{NA}}\right)^2}, \tag{7.5}$$

where λ is the wavelength of the excitation light. For lenslets with NA=0.16, having a preferable focal length to NA=0.44, the excitation intensity amplification is shown in Table 7.1.

The solid angle over which the emission is detected is also improved with the lenslets. The solid angle, Ω, is geometrically related to the numerical aperture as

$$\Omega = 2\pi \left(1 - \cos\left(\sin^{-1} \frac{\text{NA}}{n_{\text{air}}}\right)\right). \tag{7.6}$$

The telephoto lens has NA=0.05 while our lenslets have NA=0.16. The improvement in the emission collection is shown in Table 7.1.

The lenslets also provide a spatial magnification of the fluorescent molecules inside of the capillary. The emission light fills the entire lenslet area, and the lenslet is then imaged by the telephoto lens. Thus the improvement in spatial magnification can be expressed as shown in Table 7.1. The capillary inner diameter, d_i=75 μm.

7.4.1 Focal volume

As the lenslets excite and collect fluorescence within their focal volume, it is useful to compute theoretically what that volume is. Then, given a concentration of

fluorophores, one can estimate how many molecules are being detected.

For this estimation of the focal volume, we assume a cylindrical shape. The focal volume of a lens under these assumptions is actually somewhat elliptical [159], but a cylinder can serve for first-order estimation of the volume. In this cylinder, the cross-sectional area, a, is computed from the diameter of the central lobe of the Airy pattern (diffraction-limited focal spot intensity distribution though a lens from a collimated source), as given by Equation 7.4, according to

$$a = \pi \left(\frac{w}{2}\right)^2. \tag{7.7}$$

The length of the cylinder approximation is given by the depth of focus, as given by

$$\Delta z = \frac{\lambda}{4\left(1 - \sqrt{1 - \mathrm{NA}^2}\right)}. \tag{7.8}$$

For typical lenses with r=0.5 mm and R_c=1.5 mm, NA=0.16 and the focal volume has a=1.7 μm^2, and Δz=9.3 μm, resulting in a total volume, V_f, of 100 μm^3. Relaxing R_c to 2.5 mm decreases NA to 0.1 and increases the focal volume to 750 μm^3.

7.5 Lens array fabrication

As our application requires arrays of thousands of tiny high numerical aperture (NA) lenses, tightly packed with 100% fill factor (the ratio of the active refracting area to the total contiguous area occupied by the lens array, f_f), with apertures of 1 mm and f/#<5, we developed a novel manufacturing technique for producing them. The technique, which utilizes a combination of ball-end milling and "micro-forging," enables injection molding into low cost, high accuracy, and low roughness molds.

The technique involves first performing a rough cut with a ball-end mill. The final shape and sag height are then achieved by pressing a sphere of equal diameter into the milled divot. Using this process, we have fabricated molds for rectangular arrays of 1-10,000 spherical lenses with apertures ranging from 0.25-1.6 mm, sag height ranging

from 3-130 μm, inter-lens spacing of 0.25-2 mm, and fill factors from 0-100%. Mold profiles have average roughness (Ra) and figure error, defined as the average absolute deviation from the desired geometry, of 68 nm and 354 nm, respectively for the most challenging conditions of 100% fill factor with 1 mm aperture lenses. The required forging force was modeled as a modified open-die forging process and experimentally verified to increase nearly linearly with surface area. This novel micro-forging finishing process coupled with milling is easily adapted to lenticular arrays, scalable for thousands of individual lenses, and enables fill factors up to 100%. Expected limitations include milling machine range and accuracy (limited to >0.25 mm aperture lenses in our work), and molding equipment tonnage rating.

7.5.1 Lens array fabrication introduction

Current micro-lens array mold master manufacturing techniques include photoresist reflow [160, 161], laser beam shaping of photoresist [162] or glass [163], photothermal expansion [164], ion exchange [165], diamond turning, and micro-droplet printing [166]. Save diamond turning, these processes can often deliver exquisite replication of 15-500 μm aperture lenses that are often limited to square close-packed (f_f=π/4=78%) or approaching hexagonally close packed (f_f=91%). Efforts to increase f_f have not resulted accurate lenses over the full aperture [167]. High numerical apertures can also be challenging, as typical lens' sags of this size limited to 1-20 μm. These processes can also be relatively expensive and time consuming to implement, and are limited in area that can be patterned to about several hundred millimeters.

Diamond turning, for costly machine (e.g., Precitech) and tooling (e.g., Chardon Tool, K&Y Diamond), can produce microlens arrays with 250 nm figure error and 9 μm roughness over 1 mm apertures, as well as a variety of other sizes. This process can be used for aspheres, a tremendous advantage over other processes, and costs can be distributed if the master is replicated, such as by injection molding. Limitations include the area that can be patterned, high start up costs, and fragile tooling. Sag height can be limiting as well, as the tools would need to match the part slope over large distances (e.g., several millimeters). In this work, we sought to develop a low

cost, flexible mold fabrication technique that could match the figure error of diamond turning, while compromising somewhat on roughness.

7.5.2 Fabrication methods

The combined milling and micro-forging process is shown schematically in Figure 7-5, accompanied by photographs of a 100 lens mold master and injection-molded lens array. To fabricate the lens array molds, we first performed a rough cut with a ball-end mill. The final shape and sag height are then achieved by pressing a sphere of equal diameter into the milled divot. The reasoning behind the process is that the mill determines the majority of the mold figure, and the subsequent forging with the with a ground, polished sphere imparts a near perfect figure and roughness to the mold without substantial deformation that could affect neighboring lenses in a tightly-packed array. After mold fabrication, injection molding is performed, typically with polymethyl methacrylate (PMMA), which is harder, more transparent, has lower ultra-violet absorption and less shrinkage than other commonly used optical polymers such as polycarbonate and polystyrene. In practice, mold masters were created using a titanium-nitride coated ball-end mills (McMaster) and a tungsten-carbide spheres (McMaster). Using this process, we have fabricated molds for rectangular arrays of 1-10,000 spherical lenses with apertures ranging from 0.25-1.6 mm, sag height ranging from 3-130 μm, inter-lens spacing of 0.25-2 mm, and f_f ranging from 0-100%.

A machining center (Haas, VF-OE) was used for lens mold fabrication. The system has 5.0 μm accuracy and 2.5 μm repeatability in all axes with work volume of 0.8 m $\times$ 0.4 m $\times$ 0.5 m. Experiments to determine required forging force were performed on a 90 kN hydraulic press (Devin, LP-500).

The quality of the lens molds produced with this technique was determined using two parameters: figure error and average roughness (Ra). The figure error is defined as the average absolute deviation from the desired shape and provides a measure of how well the process formed the desired mold geometry. Average roughness is a commonly used roughness parameter defined as the average absolute deviation from the mean. These two parameters were calculated from mold surface profiles measured using a

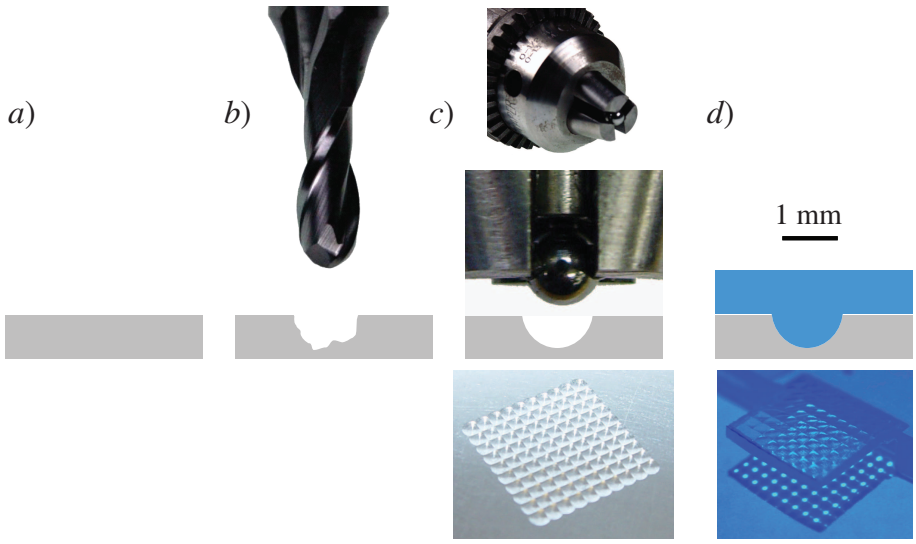

Figure 7-5: Schematic of lens array mold manufacturing process and photographs of 100 lens array mold and molded part. (*a*) A blank aluminum mold is faced and polished. (*b*) A rotating ball-end mill is used to cut an array of divots. (*c*) A tungsten-carbide sphere is lowered onto the surface to deform the divots to the final shape. (*d*) The mold is used to injection mold lens arrays, as we have done for arrays containing up to 10,000 lenses.

stylus profilometer (Mitutoyo, SV-3000) with a 2 μm radius tip. This instrument was used to obtain uni-directional scans with 1 μm resolution in the lateral direction (across the mold), x, and 1 nm resolution in the transverse direction, z. From this raw data, we implemented several algorithms to measure figure error and surface roughness.

The figure error was calculated using an algorithm that compares the measured surface profile to the desired shape. By shifting the desired shape in the x and z directions with respect to the measured profile, the algorithm finds the optimal fit (least error) of the desired shape to the measured data and computes their average absolute difference to determine the figure error.

To compute the average roughness, we first detrend the measured profile by separating the low frequency profile associated with the mold's round shape from the high frequency profile associated with surface roughness. This is accomplished by dividing the measured profile into an integer number of sections and successively fitting and subtracting a 2^{nd}, 1^{st}, and 0^{th} order polynomial from each section. For each section

the average roughness, r_i, is calculated as the average absolute deviation from the mean, and then these roughnesses, r_i, are averaged to yield the average roughness of the profile. The optimal number of sections was determined by performing this operation on simulated noise data with known roughness; the resulting sections have $\sim$32 μm width.

7.5.3 Forging theory

As we seek to create spherical lens impressions in an aluminum mold master for injection molding, theoretical modeling of the required force was performed. We modified the open-die forging model by Kalpakjian [168] to derive a new model for hemispherical impressions on a planar substrate, assuming that the tool is much more stiff than the workpiece, as

$$F = Y_f \pi A_s \left(1 + \frac{\mu d}{3z}\right), \tag{7.9}$$

where F is the forging force required, Y_f is the flow stress of the material-estimated by the true stress at 100% true strain, A_s is the surface area of the spherical impression, μ is the coefficient of friction typically assumed to be 0.2, d is the lens aperture, and z is the depth of the forge (sag height of the lens). Both A_s and z are functions of d and R, the lens radius of curvature, as given respectively by

$$z(d, R) = R - \sqrt{R^2 - \frac{d}{2}^2}, \tag{7.10}$$

and

$$A_s(d, R) = 2\pi R^2 \left(1 - \frac{\sqrt{R^2 - \frac{d}{2}^2}}{R}\right). \tag{7.11}$$

Thus, this model can be utilized to predict the required force, F, to create a spherical impression of depth z and diameter d for a range of mold master materials, lens sizes and sag heights.

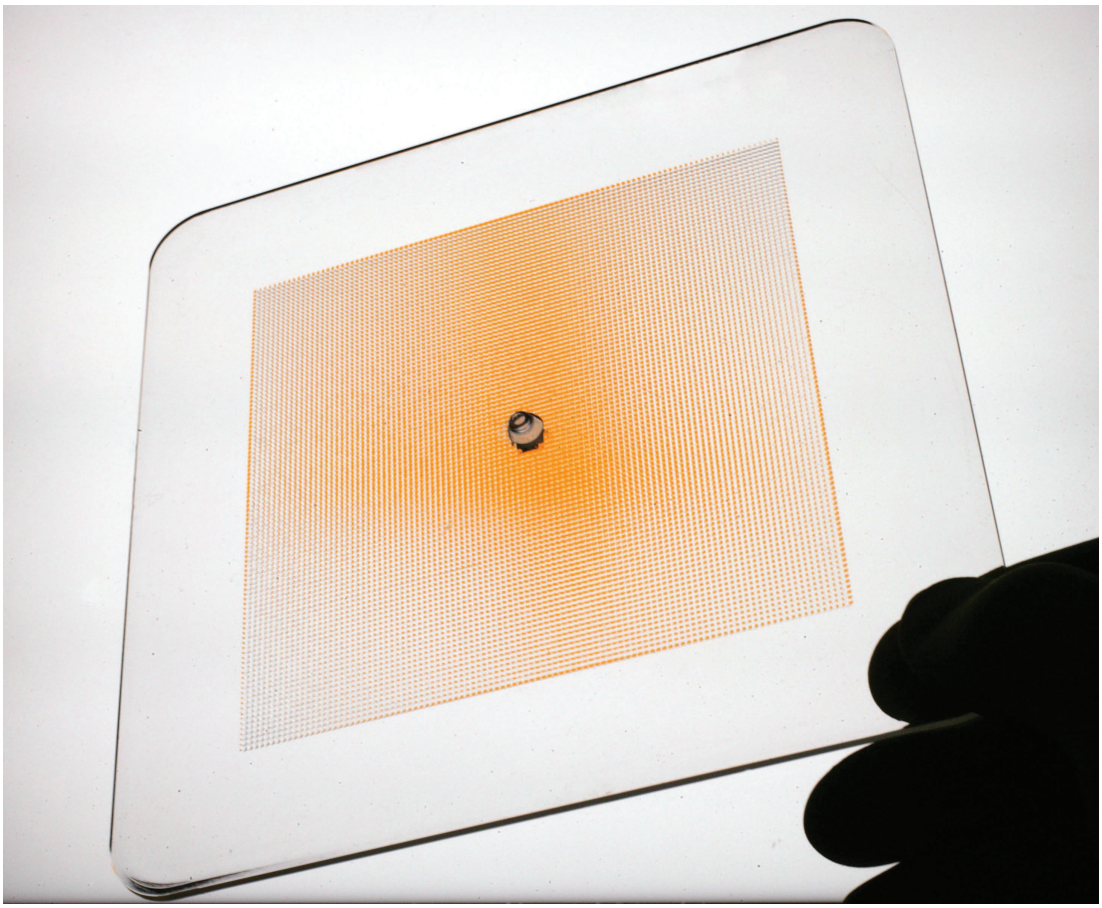

Figure 7-6: Ten-thousand PMMA lenses injection molded from a micro-forged mold. The convex lenses are each 1 mm aperture, 2.5 mm radius of curvature with 100 μm sag height.

7.5.4 Lens array fabrication results and discussion

We have used this technique to manufacture molds, and subsequently lenslet arrays, for a variety of lens curvatures, apertures, spacing, manufacturing technique, and number of lenslets. To give an indication of ability of this process to make molds containing thousands of tightly-packed (f_f=100%) lenses, Figure 7-6 shows a 10,000 lens injection molded part.

Using a hydraulic press to micro-forge impressions in a blank aluminum mold substrate without any milling, we were able to compare the open-die forging model to experimental forces as shown in Figure 7-7. These experiments were conducted with several sphere diameters (10-25.4 mm) in aluminum 6061 substrates. The results

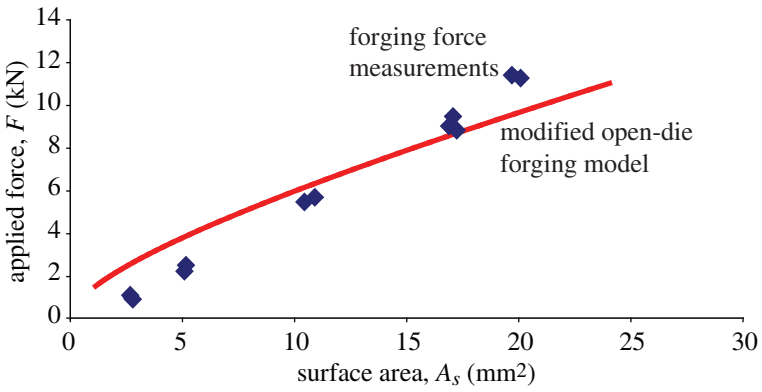

Figure 7-7: Forging force theory and experimental force measurements versus molded surface area. Data from a series of experiments with spheres of varying radii of curvature: 2.5 mm, 5 mm, 12.7 mm.

indicate that one can well predict the required forging force for typical lens sizes. For the smallest surfaces areas, the model deviates the experimental results by up to three times. Thus the model is conservative. Typical lenses with 1.41 mm aperture, 2.5 mm radius of curvature, have surface area of 1.6 mm^2 and require 1 kN of forging force. Our machining center (Haas VF-OE), used for this work, has 25 kN forging force capability.

Numerous lens array molds were fabricated, measured with the stylus profilometer, and analyzed as described. The results for the mold figure error and roughness are shown in Figure 7-8.

For lens arrays with $f_f \geq 80\%$ the figure error decreases monotonically with increasing mill depth (decreasing forge depth). Since the forging process does not remove material, but plastically deforms it, the neighboring lenses in tightly packed arrays are affected by the forging process. On the other hand, the roughness generally increases with increased mill depth. This can be attributed to the ball end-mill tool, which has roughness much larger than the polished tungsten-carbide sphere. The forging sphere was measured to have a 206 nm figure error and 19 nm roughness.

The advantage of the combined process (milling + forging) is clear when comparing the figure error and roughness of the lens mold to that created by either process alone. Purely milled divots typically have 1.7 μm figure error and 300 nm roughness,

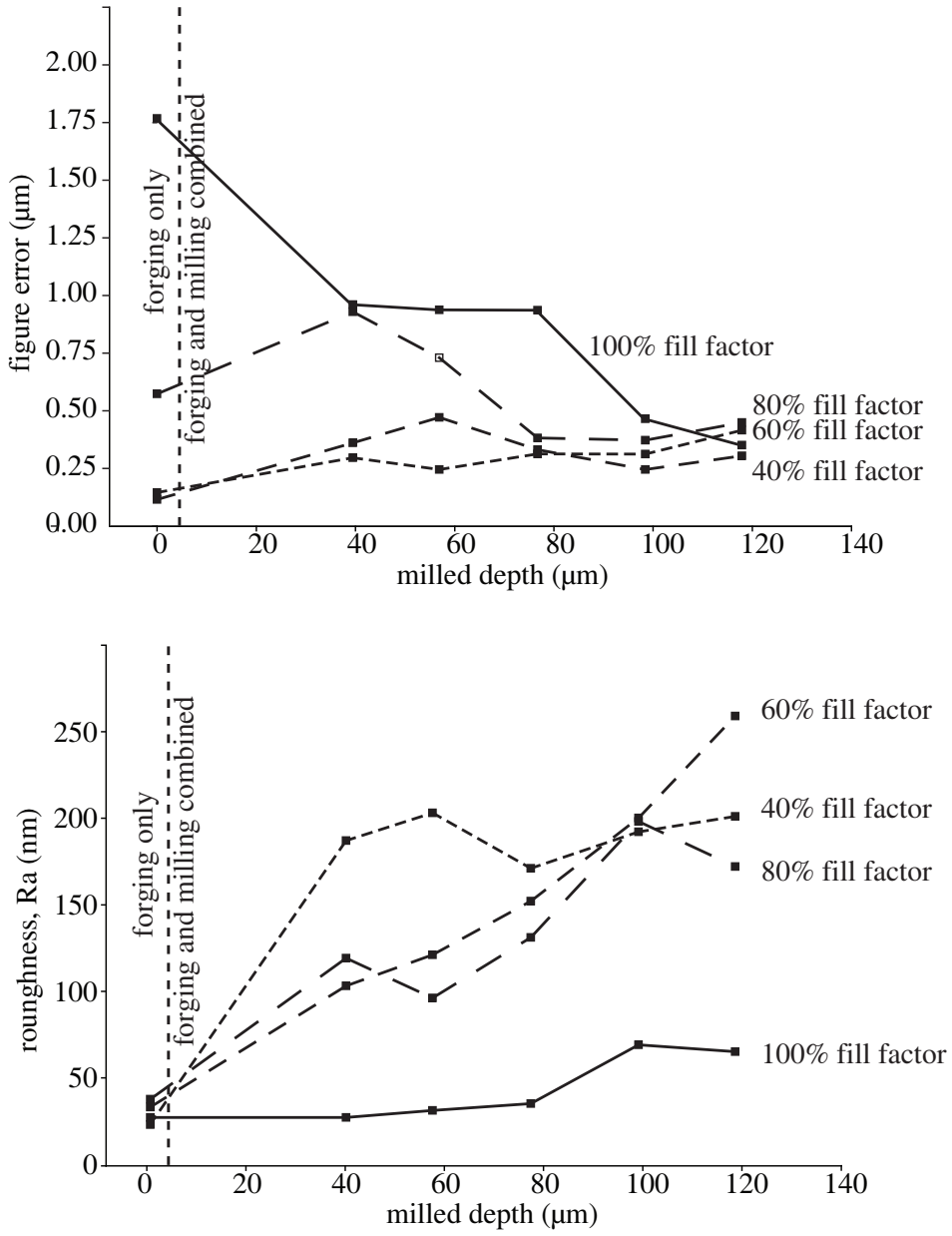

Figure 7-8: Figure error (*top*) and roughness (*bottom*) of lens array molds. All lenses have 120 μm total sag height, so increasing milled depth implies that a larger percentage of the final figure was determined by milling. As milled depth increases, the figure error generally improves (*top*). The effect is most pronounced for fill factors $\geq$80%. As milled depth increases, the roughness increases (degrades) somewhat to a plateau, but is still small (*bottom*). For fill factors less than 80% forging alone has a superior combination of figure error and roughness. For 100% fill factor lens arrays, milling to the full depth and subsequently forging to the full depth results in the best figure with reasonable roughness.

both independent of f_f. Purely forged divots have figure error that depends strongly on f_f (see Figure 7-8 *left*), at worse 1.8 μm for f_f=100%, and roughness of 31 nm independent of fill factor. In Figure 7-8, the left most data points corresponding to milled depth equals zero imply that the forge only was used. The optimal combination of milling and forging results in a figure error that is 4-5$\times$ better than either process alone, while roughness remains good at $\sim$68 nm.

For f_f=100% molds manufactured using the optimum combination of milling and forging (milling to full depth followed by forging to the same depth), measurements of the resulting injection molded PMMA lens array showed that the figure error increased from 354 nm to 939 nm while roughness decreased from 68 nm to 36 nm. This is reasonable given the 1-2% linear shrinkage of PMMA upon cooling to room temperature after molding, and corresponding attenuation of high frequency features.

We also fabricated f_f=100% lens array molds with a range of apertures, 0.25-1 mm, and constant radius of curvature equal to 2.5 mm. Roughness was independent of lens aperture, averaging 75 nm. Figure error improved as lens aperture was reduced: 1 mm diameter, 50.5 μm sag lenses had 0.35 μm figure error (see Figure 7-8 *left*) , 0.5 mm diameter, 12.5 μm sag lenses had from 0.15 μm figure error, and 0.25 mm diameter, 3.1 μm sag lenses had from 0.11 μm figure error. These figure errors were evaluated over 80% of the full lens aperture. Molds with $<$0.25 μm lenses had poor figure error, $>$1μm, largely due to the accuracy of the machining center and end mill.

Several drawbacks to the process should also be considered. Prior to mold fabrication, calibration of the depth of milling and forging can be challenging. In our experience, a microscope objective and length standard can be used to measure the lens aperture, or impression diameter, with 10 μm resolution. This uncertainty has corresponding effect on depth of about 1 μm. Other effects contribute to process error but are more difficult to quantify. Thermal expansion of the machining center, and other relatively low frequency transients (with durations on the order of tens of minutes) can affect lens depth, or sag. In addition, elastic "springback," upon unloading the forge can contribute to figure errors, primarily in the lens sag. All combined, we measured these factors to affect the sag height of the lenses by, at most, 20 μm. For

lenses with large surface areas (e.g., >57 mm^2 each), the required forging forces may exceed typical machining center capabilities.

7.5.5 Lens array fabrication conclusions

For lens arrays with $f_f \geq 80\%$ the combination of milling and micro-forging offers great potential for the fabrication of molds with low figure error and roughness. For such tightly-packed lens arrays varying in lens aperture from 0.25-1 mm, we have demonstrated that figure error will be $\leq$354 nm and roughness will be $\sim$75 nm. We have used this process to mold lens arrays containing more than 10,000 lens elements, over a 100 mm $\times$ 100 mm area. Should lens arrays with $f_f < 80\%$ be desired, forging alone can provided figure errors of <270 nm and roughness of 31 nm. Thus, this flexible, low-cost, scalable process could supplant many other lens array manufacturing processes that operate within this size and density range.

7.6 Alternative design considerations

7.6.1 LED array

In the 42 LED array used in the detection system, each LED is bonded to a hexagonal printed circuit board (PCB) backing. This substrate limits the packing density, and hence the power density, of the array, to 13.2 W/m^2 after spectral filtering as stated in Section 7.3, or 62 W/m^2 without filtering. These measurement are taken at a distance of 300 mm.

In contrast, a typical argon-ion laser (e.g., Ion Laser Technology, 5425ASL), outputs $\sim$2600 W/m^2. If the design of the fluorescence detection system, specifically the lenslet array or microscope objective in the case of a laser scanning system, is such that the fluorophores are not saturated or photobleached, then increasing the power density will increase the SNR. Thus we set out to explore alternative excitation sources and configurations.

A laser excitation for the 10,000 capillary system would be possible, but would

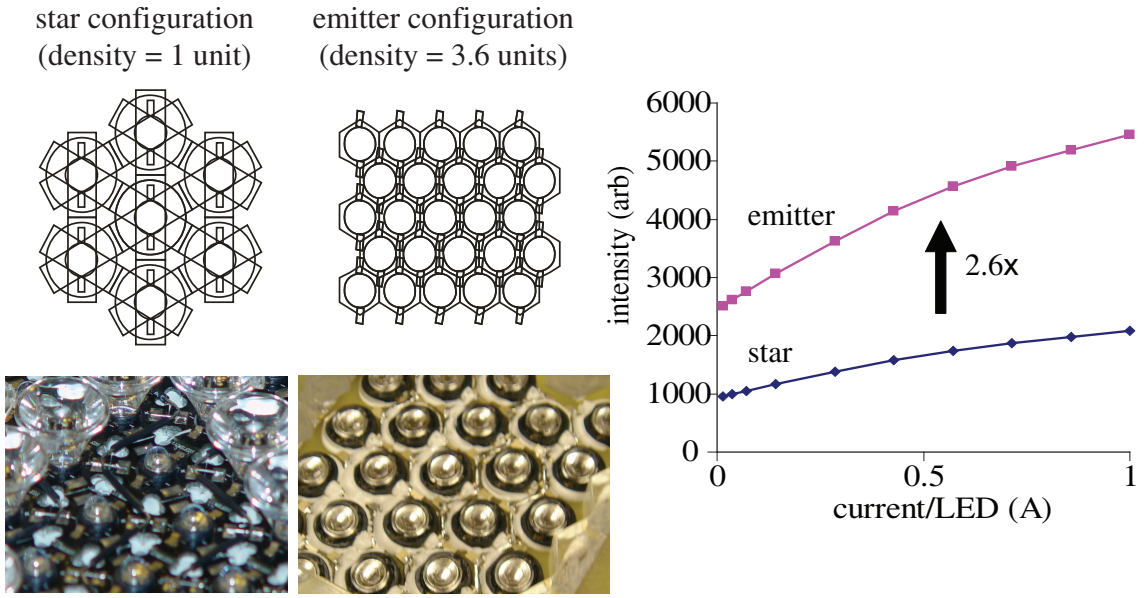

Figure 7-9: The LEDs can be hexagonally close packed on "star" PCB backings, or directly as emitters. The emitter configuration is 3.6× more dense than the former, and produces about 2.6× higher power density at all currents. The LEDs can be driven at 1 A each (at ∼3 V).

require a laser power of 2 W to maintain the energy density of the single laser-single capillary system. Other design configurations, as discussed in Section 7.1, also require laser powers of several watts. Lasers of this power and wavelength are expensive, large, and complex. Furthermore, large collimating lenses (i.e., 150 mm clear aperture) would be required.

The possibility of densely packing the LEDs into a new array configuration was considered. By utilizing the LED emitters without their hexagonal, "star," PCB backing, we were able to produce what is most likely the highest power density LED array to date (as of 2005). The assembly has a custom PCB backing (PCB Express), and a matching set of hexagonal collimators. With this array, we achieved 163 W/m^2 at 300 mm distance, or 2.6× the previous power density. These results are shown in Figure 7-9.

Though the emitter array was able to output a higher power density, the heat dissipation (∼65 W) over the 25 LED array proved problematic. After a meltdown, we reverted to the previous packing density. If heat conduction from the LED array can

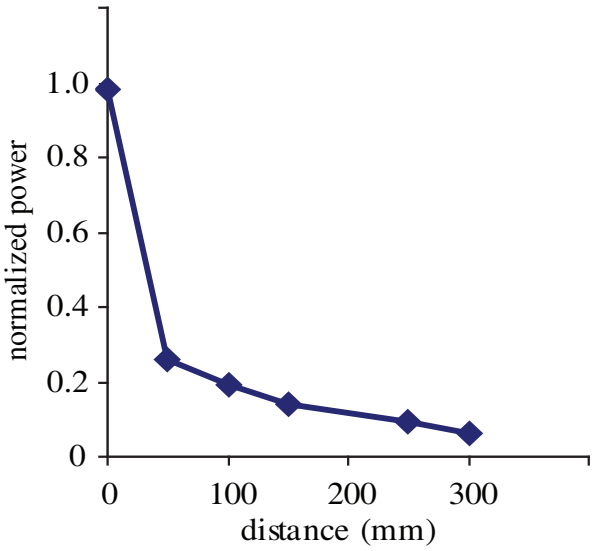

Figure 7-10: Optical power versus distance for the LED array. The LED is a diffuse light source, so the power decays rapidly as the distance is increased. At a distance of 300 mm, the power is only 6% of that at the source.

be improved in future work, the emitter array can potentially improve the detection SNR by 2.6×.

7.6.2 Collimator

In Section 7.3, and particularly in Figure 7-2, we discussed how undesired measurement of power in the 400-500 nm band can require background subtraction. This stray light is most likely the result of off-axis reflections and transmission in the optical system, with effectively shorter wavelengths, since the LEDs are inherently extended, incoherent illumination sources. To overcome this, and thus improve the raw signal-to-noise of the measure spectra, we explored the possibility of collimating the LED array.

A collimated source (e.g., a laser beam) will retain its intensity over a long distance. In contrast, we observed a large divergence, and corresponding loss of power, in our LED array, as a function of distance from the source, as shown in Figure 7-10. This data further supports the hypothesis that off-axis rays are sneaking through the filters, as such power decay implies a wide divergence angle.

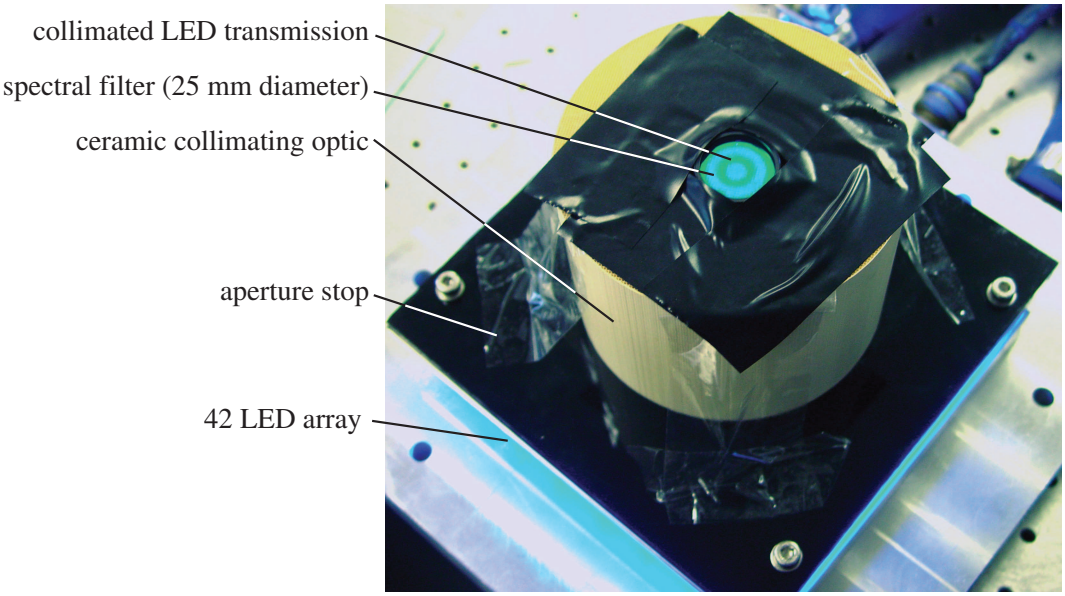

Figure 7-11: The collimating optic, a ceramic cylinder with square thru-holes, can be attached to the LED array to collimate by attenuating off-axis light rays.

We implemented a collimating optic to correct the problem. The collimator is an extruded ceramic catalytic converter substrate (NGK, Japan). The cylindrical part, shown attached to the LED array in Figure 7-11, works by attenuating off-axis rays. Thus the light emitted from the collimator has a much smaller divergence than without. This smaller divergence enables better spectral filtering. The power spectrum of the LED illumination through the excitation spectral filter is shown with and without the collimator in Figure 7-12. The off-axis illumination has an shorter wavelength as expected, and the collimated light can be more effectively spectrally filtered.

However, despite the encouraging collimation results, we chose to abandon this collimator because of its reduction in power. The cross sectional area of the ceramic, coupled with attenuation of off-axis light, dramatically reduced the power available to excite fluorescence.

7.6.3 Impact-ionizing CCD

Instead of a low-noise astronomy grade CCD camera, another type of CCD technology, called impact-ionization, was considered. This technology works by multiplying

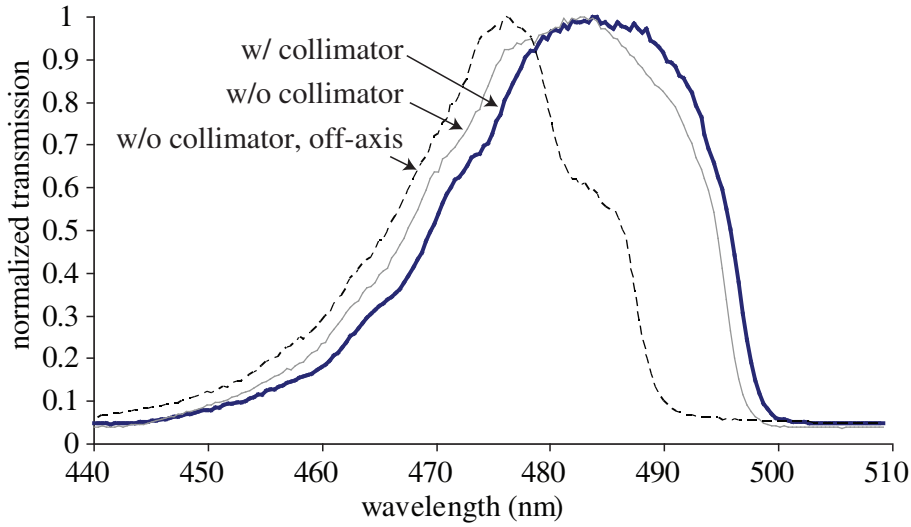

Figure 7-12: Power spectrum of the LEDs measured through the spectral filter, with and without the collimating optic. In addition, the off-axis measurement illustrates that shorter wavelengths can undesirably pass through the spectral filter off-axis. The collimator improves the ability to spectrally filter, but with substantial reduction in power.

photo-generated electron counts in a manner similar to a photo-multiplier tube, by a cascading electron effect. This technology is advantageous when the application is read-noise limited, as in low light situations where high frequency (e.g., 30-60 Hz) images are required.

Numerous experiments were performed with this CCD and it was determined that we are not limited by read-out noise, but rather dark (thermal noise), field effects, and shutter timing. Variation in the source illumination is also suspected to limit the noise threshold. Thus, we utilized the low noise astronomy-grade CCD instead. With Peltier cooling to -10 °C, the dark noise from the CCD is only 1 e^-/pixel/s.

7.7 Results and discussion

In the preceding sections, the hypothesized benefits of the lenslet array were presented, including increasing excitation intensity, increasing magnification, and increasing the numerical aperture (NA) over which fluorescence emission is collected. We sought to verify these benefits, and explore their limitations.

7.7.1 Side-column detection

The theoretical benefits that the lenslets offer to the detection system performance have been experimentally verified. Figure 7-13 shows the SNR increase as a function the lenslet presence as well as CCD exposure time. For this experiment, an empty capillary was placed randomly underneath the lenslet array. After recording a reference image (Figure 7-13*a*), the capillary was filled with 200 nM Alexa 488 conjugated to BSA protein. Sample preparation is described in Appendix B. Figure 7-13*b* shows the fluorescent probes within the capillary after background subtraction. The signal from the bare capillary (*solid line*) is compared to the signal where the capillary is underneath a lenslet's focal volume (*dashed line*) in the plot in Figure 7-13 as a function of exposure time. The SNR for the region of capillary underneath the lenslet is approximately 38× greater for all exposure durations. SNR is defined as the signal amplitude divided by the standard deviation of the noise. This is the result of the increase in both excitation intensity and emission collection solid angle.

The spatial magnification can also be computed from Figure 7-13*b* from the ratio of the number of bright pixels in the two regions. The spatial magnification measured is 100×. This is slightly lower than the 178× theoretical estimate (Table 7.1) due to capillary wall reflections.

7.7.2 End-column detection

Experiments with both the fluorophore solution and electrophoresis of labeled primers were conducted to demonstrate end-column detection. In this Chapter, we only present results from the fluorophore solution. All electrophoresis results are presented in Chapter 8. Figure 7-14 shows capillary end-column detection of the fluorophore solution. In this experiment the capillary was coaxially aligned with a lenslet, a reference image was captured, and the capillary was then filled with the fluorophore solution. After subtracting the reference image, the intensity of the single lenslet is clearly seen, indicating successful end-column detection with SNR=10.

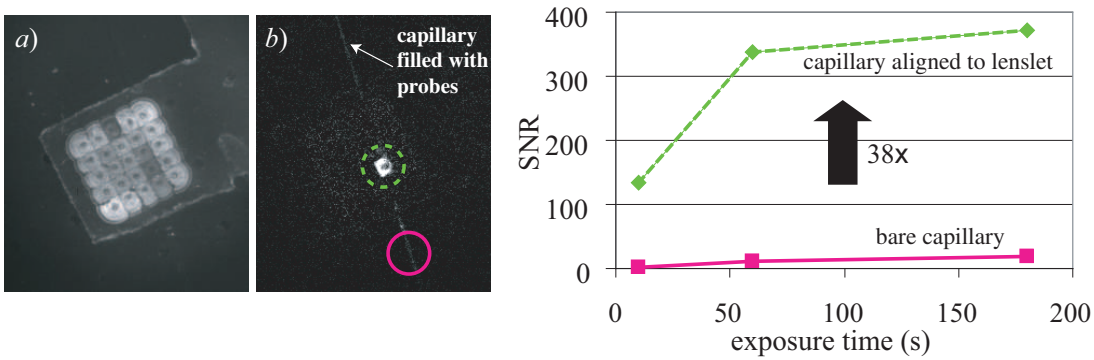

Figure 7-13: Measured benefits conferred by the lenslets towards detection sensitivity. Image *a* shows the lenslet array with an empty capillary positioned below it and spanning the length of the array. In image *b*, the capillary has been filled with the 200 nM Alexa 488 solution. The background has been subtracted to isolate the signal. The improvement to the fluorescent probe measurement SNR is shown in the plot. The 38× increase in SNR is attributed to both increase in excitation intensity and emission collection solid angle. The spatial magnification is ∼100×.

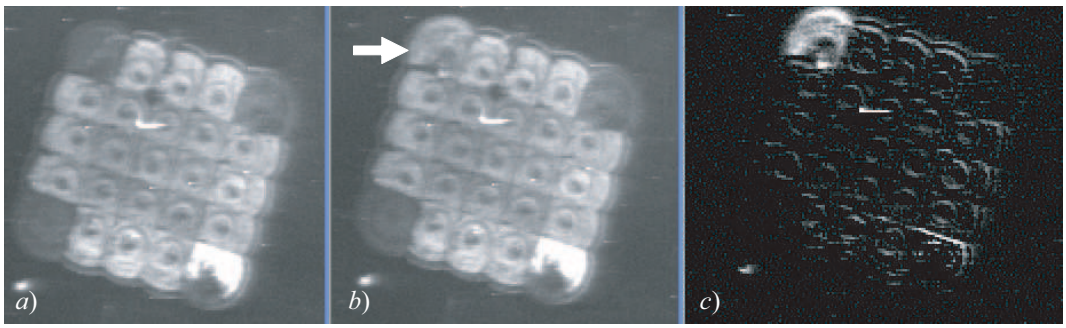

Figure 7-14: Demonstration of end-column detection using the lenslet array. Image *a* shows the lenslet array with an empty capillary coaxially positioned below it. The center image *b* shows the lenslet array after the capillary has been filled with filled with 200 nM Alexa 488 conjugated to BSA protein. In image *c*, the difference is shown (*b*-*a*). The SNR is 10.

7.7.3 Saturation

These results, combined with the theoretical discussion presented above, raise some interesting questions about the lenslet array design. If increasing the lenslet NA increases the magnification, intensity of the excitation light, and solid angle over which fluorescence is collected, then why not strive to make the NA as high as possible?

Were the response of the fluorophores linear with excitation, the answer would be to increase the NA as much as possible. However, because real fluorophores can saturate, there will be an optimal NA. Below this saturation threshold, increasing the input intensity will increase the SNR because the fluorophores will return an output roughly proportional to input. However, once saturation is reached, increasing input intensity further will not further increase output (or SNR) because all of the fluorophores in the focal volume are saturated. This behavior and plateau is depicted in Figure 7-15*a*. This leads to an optimum NA. Below this optimum increasing NA increases SNR because of the lenslet benefits of increasing magnification, excitation intensity, and solid angle of collection. Above this optimum, decreasing NA increases SNR because there will be a larger focal volume within which fluorophores will be saturated.*

Thus we set out to experimentally converge on the optimum NA for maximum SNR. The results are shown in Figure 7-15*b,c*. For these experiments, a solution of 200 nM AlexaFluor 488 conjugated BSA, as described in Appendix B, was used for end column detection from capillaries in the UTMS. For relatively high NA lenslets, the saturation plateau is quickly reached, resulting in a low SNR as a function of input intensity. For the medium NA, the saturation plateau is reached around I_{in}=1600 units. However, for the lowest NA (0.10) the SNR continues to increase throughout the range of input intensity. The maximum SNR that can be achieved with this fluorophore solution is shown in Figure 7-15*c*, for a range of NA. The NA is optimized when all fluorophores in the focal volume are saturated, and further increasing the focal volume would cease fluorophore saturation. Thus we settled on NA=0.1 for our end-column detection experiments.

*Increasing NA decreases focal volume as indicated by Equation 7.4 and Equation 7.8.

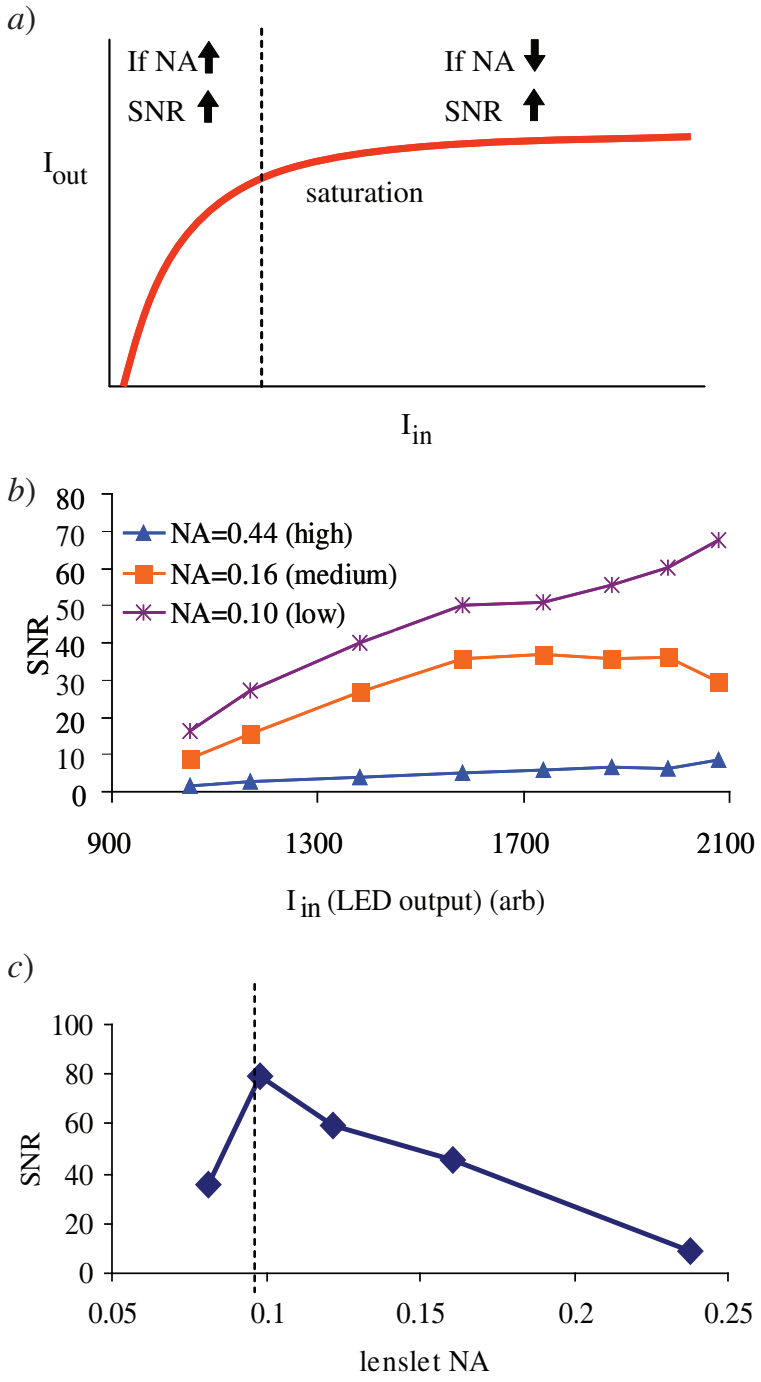

Figure 7-15: Fluorophore saturation phenomenon predicted (*a*) and experimentally verified (*b*). Beyond a saturation plateau, increasing the input intensity does not increase output intensity (or SNR). The optimum lenslet NA is at the knee of the saturation curve, where the focal volume over which the fluorophores are saturated is as large as possible. (*c*) This optimum NA has been experimentally verified for a 200 nM Alexa 488 solution, as 0.1.

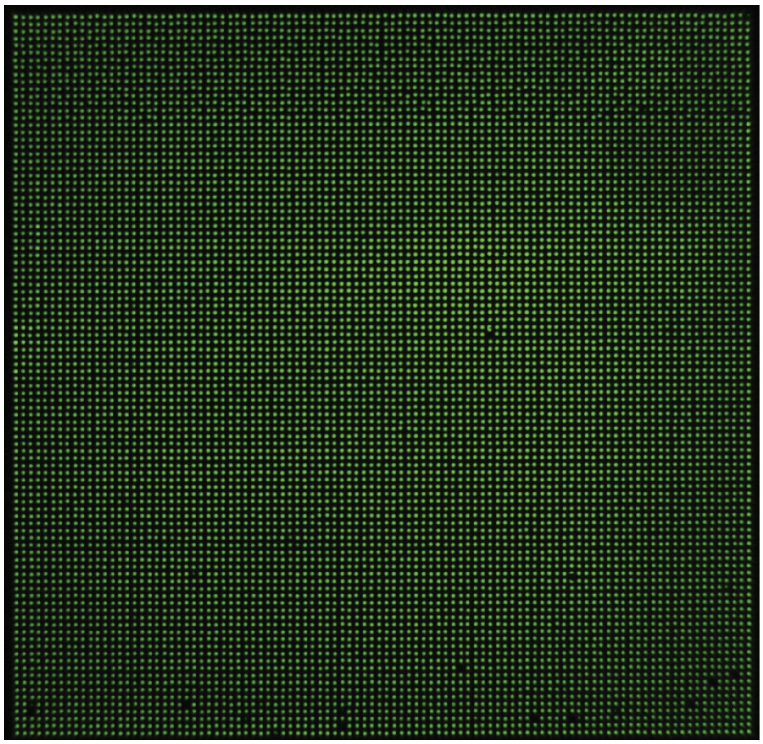

Figure 7-16: Ten-thousand wells, each with 80 nL of fluorescein can be excited and detected, suggesting that the 10,000 capillary array DNA fluorescence detection requirement is within our present capability.

7.7.4 Fluorescence detection: 10,000 capillary emulation

We also performed an experiment to verify that a 100 mm $\times$ 100 mm spatial area with 10,000 discrete fluorescent sources could be imaged. Figure 7-16 shows 10,000 80 nL wells each filled with 1 μM/L fluorescein.

7.8 Conclusions

A fluorescence detection technology has been developed which can be scaled in area to greater than 5000 capillary channels. The lenslets dramatically increase the intensity of the excitation radiation, the emission collected, and the spatial magnification. The lenslets were shown to increase the SNR by 38$\times$ and the spatial magnification by 100$\times$. End-column fluorescence detection has been demonstrated using a single capillary for a static solution (SNR=10 for 200 nM/L solution). In Chapter 8, we present

detection of charged molecules undergoing electrophoresis as well as the sensitivity of the detection technology. The number of capillaries that can be accommodated is limited only by the area of the LED array, number of pixels in the CCD, and the number of lenslets. This technology could usher in a new era of low cost and rapid mutation scanning in large populations.

Chapter 8

DNA detection experiments

Having systematically described the subsystems of the UTMS, a series of experiments is presented that convey the integrated functionality of the instrument. This Chapter focuses on results from the UTMS for detecting fluorescently-labeled DNA undergoing electrophoresis. The DNA samples used for these experiments vary from 20 base single stranded primer molecules to several hundred base duplex PCR products containing homoduplex/heteroduplex mixtures.

In this Chapter, three systems, one commercially available by Amersham, one research-grade single capillary instrument, and the UTMS, were compared in a series of electrophoretic experiments. After we define the sensitivity functional requirements, we show that the UTMS is comparable to existing technology, with advantages in throughput and cost. We present mathematical descriptions of the processes of electrokinetic injection to understand factors that contribute to loading efficiency as well. Multi-capillary detection results show UTMS end-column detection of fluorescently-labeled DNA undergoing electrophoresis in 100 capillaries, as well as channel to channel variations (e.g., phase-shift, amplitudes, diffusion) in another multi-capillary array experiment. In detail, we present mutation detection results from a melanoma cell line for the gene fragment *BRAF* exon 15, suspect in skin cancer. Results from the three instruments are compared, and they show that all instruments are capable of detecting mutants at a fraction of only 0.005, the lowest possible quantity for our pangenomic mutation scan endeavor. Additional parameters variations are studied

including fluorescent enhancement with intercalating dyes, temperature sensitivity, speed of migration, length of denaturing zone, and non-DC electrophoresis.

8.1 Experimental setup

Three distinct experimental apparatuses were used for these experiments. The first is the UTMS as described previously, with all subsystems integrated: thermal control system, gel loader, buffer reservoirs, capillary array, end-column detector. The second system is designed by Prof. Thilly and colleagues [169] to perform single capillary constant denaturant capillary electrophoresis (CDCE), as shown in Figure 8-1. In this discussion, this system will be referred to as Prof. Thilly's instrument.

Lastly, our collaborators (P.O. Ekstrøm *et al.*) have an Amersham MegaBACE 1000 (Piscataway, NJ), which is an automated 96 capillary array sequencer with modified software to operate as a mutational spectrometer. The system uses side-column laser-induced fluorescence detection with an argon ion 10 mW laser, spectrally filtered to 488 nm center wavelength. While the UTMS and Prof. Thilly's instrument use water to apply the denaturing conditions, this system uses air. This limits the maximum temperature than can be attained, but enables rapid thermal cycling, useful for a denaturing capillary electrophoresis technique called cycling temperature capillary electrophoresis (CTCE), discussed in Chapter 1, Section 1.7.3. Typically, the temperature in this instrument is oscillated with an amplitude of 1.5 °C at 1/120 Hz.

These three instruments were used complementarily to perform DNA electrophoresis experiments. Prof. Thilly's instrument and the MegaBACE served as a reliable benchmark systems, from which variations on the technique could be explored. The UTMS experiments were either reproductions or extensions of experiments performed on them for direct comparison.

8.1.1 Loading and running conditions

In these experiments a wide variety of samples, concentrations, loading and running conditions were used. The detailed conditions of each experiment are included to the

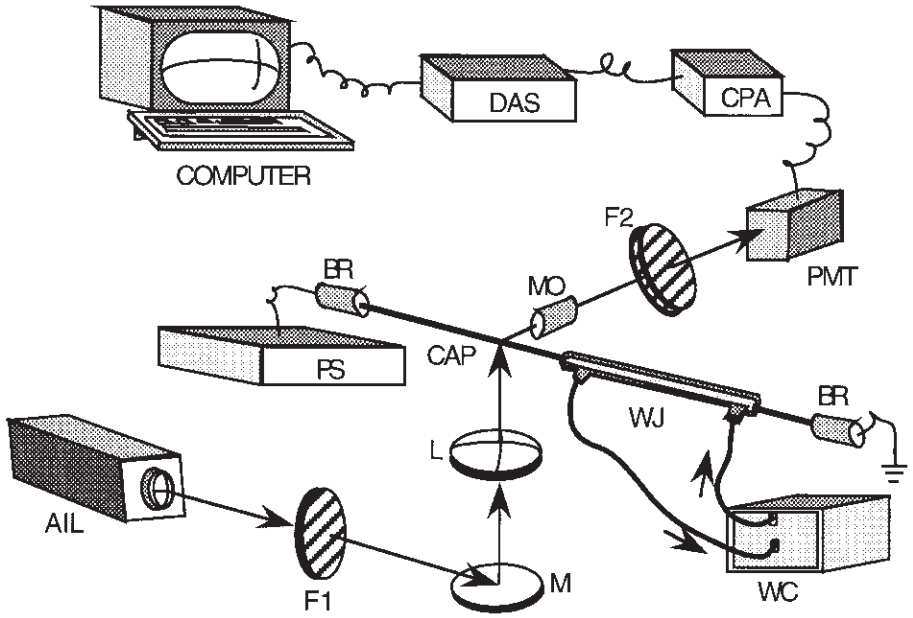

Figure 8-1: Diagram of a single capillary CDCE apparatus, "Prof. Thilly's instrument," as designed by Thilly and colleagues [169]. The power supply (PS) delivers up to 30 kV DC and 100 μA. (Spellman, CZE 1000R-2032). A laser beam from an argon ion laser (AIL) (Ion Laser Technology, 5425ASL) is filtered through a 488 nm narrow bandpass filter, F1 (Corion, 10 nm bandwidth), reflected by a 45° mirror, M (Newport), and focused by a plano-convex glass lens, L (Newport) onto a detection window of a horizontally positioned fused-silica capillary, CAP (Polymicro). Fluorescence emitted from the sample is collected by a 60× microscope objective, MO (Newport) and passes through a 540 nm and a 530 nm filter, F2 (Corion) into a photomultiplier, PMT (Oriel Instruments). The signal from the photomultiplier is amplified (10^7 or 10^8 V/A) by a current preamplifier, CPA (Oriel), recorded by a data acquisition system, DAS (Strawberry Tree Computers, ACM2-16- 8A/T51B), and transmitted to a computer. The buffer reservoirs (BR) are positioned at the two ends of the capillary. At the cathodic end, the capillary is inserted into a water jacket (WJ), the temperature of which is controlled by a constant temperature water circulator (WC). (reproduced from [169])

fullest extent possible to permit reproduction. The samples are further described in Appendix B. Details on the reagents used (e.g., gel, buffer) are included in Appendix A.

For electrokinetic loading, the conditions quoted are on a per capillary basis. Loading requires application of a potential difference (on the order of kV) across the electrodes, which creates a current (on the order of μA) in the system. This current is applied for a duration on the order of tens of seconds. To facilitate comparison between experiments and instruments, we will always present the Coulombs, Q (C), of injection required, defined as

$$Q = I\Delta t \tag{8.1}$$

where I is the injection current (A) and Δt is the injection time (s). There are 6.24×10^{18} electrons/C, by definition.

There are limits to the electric field strength that can be applied for electrokinetic loading. Most commonly, 5 V/mm is used [170], but there have been reports of 13.3-25 V/mm in the literature [121, 170]. Some have even used electric filed strengths higher than 25 V/mm [171]. The electric field breakdown strength has been measured in a variety of LPA-based gels to be about 33 V/mm.

The samples were electrophoresed at currents ranging from 1-10 μA/capillary, depending on the instrument, experiment, etc. In all cases, the power supply driving the electrophoresis was current limited, so that the voltage varied to drive the electrophoretic current.

UTMS settings

The UTMS was operated with the following settings, unless otherwise noted:

- LED array driven at 7 A total (1 A per LED) at ∼22 V, cooled with serpentine heat exchanger to 7 °C. Three hours was allowed between turning on the heat exchanger and LED array electrical current to thermally equilibrate.

- CCD array cooled to −10 °C for and repeatedly capturing images for three hours to thermally equilibrate. The sampling rate was ∼0.29 Hz from consecutive 3

s integration times. This sampling rate permitted sufficient peak sampling at electrophoretic currents of 2 μA or less.

- Capillaries used in these experiments had inner diameter, d_i=75 μm, outer diameter, d_o=360 μm. Their length was typically 300 mm. The window for excitation and detection was at the end of the capillary, as required. The electric field strength was typically V=10 V/mm.

- A water heater jacket was used inside of the thermal control chamber to apply denaturing conditions over a length of 130 mm.

Prof. Thilly's instrument settings

- The PMT gain was set to 10^6-10^7 V/A and time constant set to "medium" and the PMT high-voltage power supply was set to 500-700 V.

- The data acquisition system (Labview, BNC 2110 $\rightarrow$ NI 6023E) was set to sample at 10 Hz.

- Capillaries used in these experiments had inner diameter, d_i=75 μm, outer diameter, d_o=360 μm. Their length was typically 320 mm. The window for excitation and detection was 250 mm from the injection end. The electric field strength was typically V=10 V/mm.

- The electrophoretic current was set to 8 μA.

- A water heater jacket was used to apply denaturing conditions over a length of 130 mm.

Amersham MegaBACE settings

- Capillaries used in these experiments had inner diameter, d_i=75 μm, outer diameter, d_o=360 μm. Their length was typically 620 mm. The window for excitation and detection was 400 mm from the injection end. The electric field strength was typically V=10 V/mm.

- The data acquisition system sampled at 1.75 Hz.

- Heated air was used to apply denaturing conditions over a length of approximately 600 mm.

8.2 Sensitivity of end-column detector

The sensitivity of the UTMS is defined as the signal-to-noise ratio (SNR) as a function of the number of molecules present in the electrophoretic peak.* This is perhaps the most importance performance specification for the instrument.

For mutation detection applications, we seek to measure the mutant fraction, M_f, in a pooled DNA sample of N persons, satisfying

$$M_f \geq \frac{1}{2N}. \tag{8.2}$$

The factor of 2 is present because a person can be heterozygous for the mutation. Thus, we seek to detect a mutant fraction as low as 1 mutant in $2N$ gene copies.† The largest value of M_f possible is 0.5 in a population pooling experiment, as a larger value would make that genotype wild (preponderant), not mutant.‡ Empirically, such as from the electropherogram shown in Figure 1-3, one can asses the measured mutant fraction from a separation using denaturing capillary electrophoresis as

$$M_f = \frac{A_{\text{mutant homoduplex}} + \frac{1}{2}A_{\text{heteroduplexes}}}{A_{\text{wildtype homoduplex}} + A_{\text{mutant homoduplex}} + A_{\text{heteroduplexes}}}, \tag{8.3}$$

where $A_{\text{heteroduplexes}}$ is the area under the two neighboring heteroduplex peaks and $A_{\text{wildtype homoduplex}}$ and $A_{\text{mutant homoduplex}}$ are the areas under the respective wildtype and mutant homoduplex peaks. If we reasonably assume that, when rean-

*SNR is defined as the electrophoretic peak amplitude divided by the standard deviation of the noise.

†Each individual carries 2 autosomal gene copies–one from their mother and one from their father.

‡The mutant fraction can exceed 50% in other types of mutation detection studies. For example, a tumor tissue sample from an individual, in which all or most of the cells contain a mutation, will have a mutant fraction approaching 100%.

nealing, the complementary and single-base mismatched strands have equal probability of hybridizing, then $A_{\text{mutant homoduplex}} = \frac{1}{2}A_{\text{heteroduplexes}}$ and Equation 8.3 reduces to

$$M_f = \frac{A_{\text{heteroduplexes}}}{A_{\text{wildtype homoduplex}} + A_{\text{mutant homoduplex}} + A_{\text{heteroduplexes}}}, \quad (8.4)$$

which is much simpler to apply to an actual electropherogram. Some authors [172, 26], encountering situations where the wildtype fraction greatly exceeds the mutant (e.g., M_f<5%), assume that $A_{\text{mutant homoduplex}}$ is negligible (or unmeasurable practically speaking relative to the wildtype homoduplex is close proximity), so Equation 8.3 simplifies to

$$M_f = \frac{\frac{1}{2}A_{\text{heteroduplexes}}}{A_{\text{wildtype homoduplex}} + A_{\text{heteroduplexes}}}. \quad (8.5)$$

In our experiments to be described, we utilize Equation 8.4 unless the mutant fraction is very low (e.g., M_f<5%), wherein we utilize Equation 8.5.

Were n molecules loaded into the capillary, a spatially distinct peak will contain p molecules in the range of

$$\frac{n}{2N} \leq p \leq n. \quad (8.6)$$

Thus p can vary from merely one heterozygous person in a pool, or $n/2N$, to n if all sequences loaded are identical. Larger N is advantageous because it directly reduces the number runs in a pangenomic scans of P people. For example, doubling N halves the number of runs to scan P people. We initially set N=100 individuals in a pooled sample to expedite the pangenomic scanning.

Further, of the p molecules in an electrophoretic peak, the instrument design is such that only f of them will actually be detected within the lenslet focal volume. Our goal is to make f=p, but that requires single molecule detection capability, a formidable challenge.

Through a series of experiments, the SNR as a function of p was determined for the UTMS, as shown in Figure 8-2. In these experiments, 20 base length primer

molecules were electrokinetically loaded into a 300 mm capillary from a 5 μL aliquot in a micro-well. After loading, the primers were electrophoresed to the end of the capillary (2 μA, $\sim$40 min). Using the end-column detector, fluorescence emission was detected just prior to the time that the molecules were eluted. After elution, the molecules migrated through the electrolyte buffer solution towards the cathode.

In the data, the SNR is observed to decrease proportionally with the number of molecules, as expected, to a limit-of-detection (LOD) of 10^7 molecules, or 16 attomole. The sensitivity of side-column laser-induced fluorescence detection, represented by both the Amersham MegaBACE experiments and Mathies's publication [144], is typically $\sim$1.6-4$\times$ more sensitive than the UTMS. Prof. Thilly's instrument sensitivity [84] (data not shown) is about 30-37$\times$ more than the UTMS. This is attributable to the overlapping excitation and emission paths (termed epi-illumination) in the less sensitive instruments. Prof. Thilly's instrument, while highly sensitive with it's excitation and emission paths oriented perpendicularly (dia-illumination), is not amenable to multi-capillary adaptations.

This stated UTMS sensitivity enables one to reasonably detect p=10^7 molecules or more (up to 10^{12} tested). So assuming N=100, a mutant fraction, M_f, of 1/200 would require that at least n=2×10^9 molecules be loaded onto the column. The mass of DNA corresponding to this number of molecules a useful parameter. To compute this, we consider the molecular weight of nucleotide bases as 330 g/M. Since a primer molecule is 20 bases, single-stranded, its molecular weight is 20$\times$330 g/M = 6600 g/M. Thus the mass of a primer molecule is 6600 g/M / Na = 1.1×10^{-11} ng, where Na is Avogadro's number equal to 6.02×10^{23} molecules/M.

So the required n=2×10^9 molecules corresponds to 0.02 ng of primers (20 bases single-stranded) or conceivably 0.4 ng of wildtype/mutant duplex DNA from a PCR reaction (assuming 200 bases double-stranded). Li *et al.* [170] have shown that 20 ng, 50 ng, and even 80 ng can be loaded onto 75 μm bore capillary with sufficiently aggressive electrokinetic injection conditions while still maintaining resolution adequate to detect single base mutations by CDCE. Thus, we have no doubt that the demonstrated detection sensitivity is sufficient to detect $M_f \geq 0.005$. Eighty ng of du-

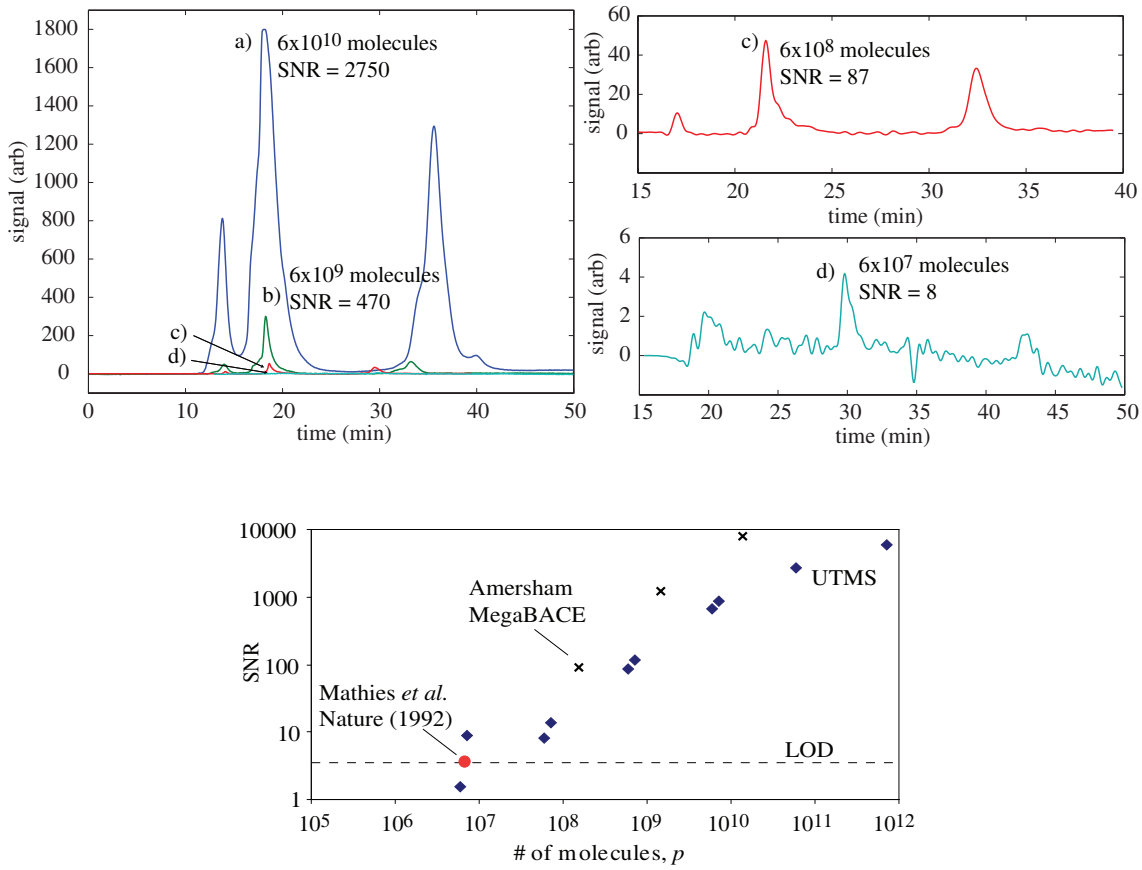

Figure 8-2: Sensitivity of the UTMS end-column detector as compared to side-detection using common excitation/emission paths. In all experiments, fluorescently-labeled primer molecules (20 base single-stranded DNA) were electrokinetically loaded and electrophoresed through a linear (poly)acrylamide sieving matrix to the detector. The sensitivity of the UTMS, with limit-of-detection (LOD) of 10^7 molecules or 16 attomole, is comparable with conventional laser induced fluorescence, with LOD=10 attomole [144]. The UTMS can accommodate 10,000 capillaries, while the others are limited to less than 1,000. In addition, the UTMS detector costs about an order of magnitude less, and is less complex.

plex DNA 200 bp in length corresponds to 3.6×10^{11} molecules, more than two orders of magnitude more that the detection threshold for $M_f \geq 0.005$.

The multiple peaks observed in Figure 8-2 are due to inter- and intra- primer binding, and are inconsequential for the sensitivity analysis. About 60% of the total area under the peaks is contained in the second, labeled peak. To accurately compute the labeled number of molecules in this peak, we have multiplied p by 0.6.

In the sensitivity plot, the number of molecules, p, on the abscissa is challenging to directly measure. We have performed a series of experiments to accurately measure the number of molecules in the peak by serially injecting from the same sample well, electrophoresing those samples, and detecting their SNR. By measuring the decay rate of the peak SNR as a function of injection number, we can back-compute the number of molecules injected on the first injection from the sample concentration. This data is shown in Figure 8-3 for both the UTMS and a commercial MegaBACE 96 capillary instrument. The decay rate, or the fraction of sample injected, into the MegaBACE instrument in these experiments is 27%, while it was measured at an average of 24% for the two UTMS experiments shown.

The samples used for these experiments are AlexaFluor 488-labeled primers diluted in Millipore H_2O. Experiments on the Amersham MegaBACE instrument were conducted as follows: a 10 μL solution at 0.33 nM/L (2×10^8 molecules/μL) was injected for 5 s at 10 μA (50 μC), at 10 kV. Electrophoresis was run at 9 kV (9 μA) at 44 °C in a solution of 8% PVP matrix with pH 8 and 7 M urea. On the UTMS, a 5 μL solution of concentration 20 nM (12×10^9 molecules/μL) was injected for 15 s at 2 μA (30 μC), at $\sim$1 kV. Electrophoresis was run at 1 kV (2 μA at 30 °C in Spectrumedix PCR-QC gel, an $\sim$5% LPA solution (see Appendix A).

8.2.1 Loading efficiency

If 25% of the 6×10^{10} 20 base single-stranded primer molecules in a 5 μL, 20 nM solution can be electrokinetically loaded into a sieving medium, or viscous gel, within the capillary at 30-50 °C, what does that imply for the loading of 200 bp duplex DNA, the typical mutant analyte? This question is not straightforward but has been

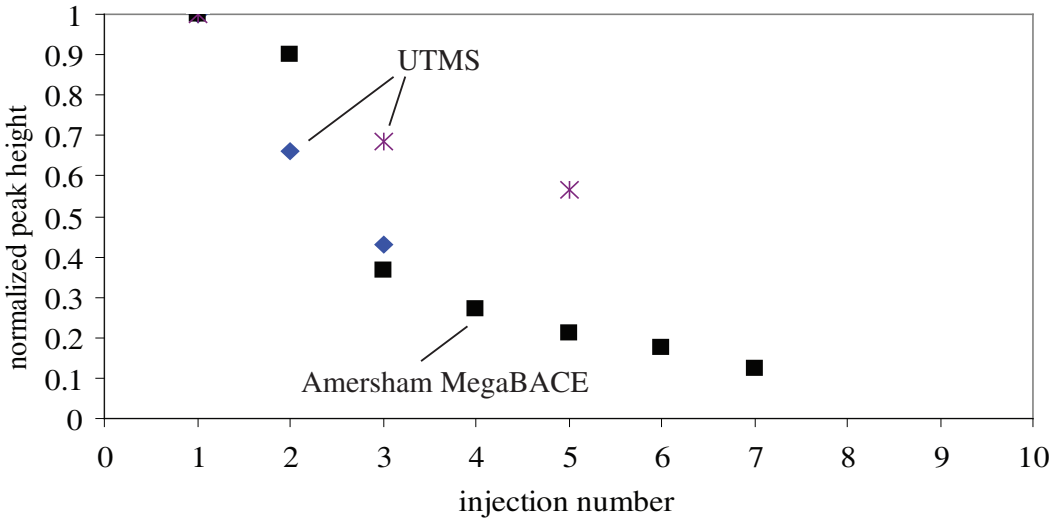

Figure 8-3: The amplitude of electrophoretic peaks decays exponentially with repeated injections from the same sample vial. Measuring this decay rate for given injection conditions can be used to back compute the fraction of molecules injected in the first injection from a virgin sample solution. In these experiments, a 30-50 μC electrokinetic injection loads about 25% of the sample into the capillary's sieving matrix.

addressed in the literature [173, 174, 175, 176]. As described by Boček [173], the mass flux of molecules, m, loaded onto the capillary during electrokinetic loading can be expressed as

$$m = c\frac{\mu}{\kappa} I \Delta t, \tag{8.7}$$

where c is the concentration of the solutes, μ is the mobility of the fragments, κ is the specific conductivity, I is the electric current, and Δt is the time applied. The specific conductivity is further defined as

$$\kappa = \frac{I}{SE}, \tag{8.8}$$

where S is the effective capillary cross-section and E is the electric field strength (V/mm). In free solution (e.g., pure water), Viovy *et al.* [177] has observed that the mobilities of different lengths of DNA are the same, equal to 36.7×10^{-9} $\mathrm{m^2V^{-1}s^{-1}}$. These mobilities are size independent because of the constant total charge-to-length ratio of DNA. This implies that when the sample is injected from a free solution of

DNA, there is no size-dependent injection bias [174, 175], regardless of differences in the mobilities of the DNA inside sieving medium within the capillary.

For our typical case, about 1% of the molecules from the sample well should be electrokinetically loaded according to Equation 8.7, given E=1250 V/300 mm= 4.17 V/mm, S=π(37.5$\times10^{-6}$ m)^2=4.4$\times10^{-9}$ m^2, I=2 μA, Δt=15 s, and c=1.2$\times10^{10}$ molecules/μL, . Our observation of $\sim$25% is not unreasonable but could be discrepant because of a combination of locally higher electric field strength, or different mobility from the ideal free solution.

In the case of loading PCR products, one can assume that the loading efficiency will be lower than for 20 base single-stranded primers at the same concentration. The PCR products will contain typically contain 10$\times10^{11}$ molecules/μL in 20 μL solution (as polymerized duplex DNA or unincorporated primers), but also contain chloride salts, and unincorporated single nucleotide bases (dNTP), phosphates, and template DNA which will affect the concentration of charged carriers.

In addition, in the presence of electroosmotic flow (EOF), a retrograde sieving matrix flow due to the applied electric field, will introduce an electromigration bias [174, 175] of about a factor of 2 for the larger molecules relative to the small. In other published reports, primers were injected at twice the rate of duplex DNA molecules around 200 bp in size.

In our experience, the SNR, and hence number of molecules, injected under identical conditions from a raw PCR solution (without dilution or cleanup), is as much as 40$\times$ less than for primers at the same concentration (e.g., c=10^{11} molecules/μL). Furthermore, our collaborators [Personal communication with P.O. Ekstrøm, Massachusetts Institute of Technology, November 24, 2006] have suggested that primers in a 1$\times$ buffer are injected more than 100$\times$ less efficiently than in H_20. Thus, removing the extraneous charged molecules and suppressing EOF were recognized as critical factors in our experimental setup.

To suppress EOF, we utilized a "dynamically coating" gel, described further in Appendix A. To reduce the concentration of charged carriers except for the analytes of interest, we typically cleaned up the PCR products using a commercially available

technique (Qiagen, QIAquick PCR Purification Kit) that involves binding the DNA analytes to a column, washing away the others, and then eluting the analytes into a dilute buffer solution. Instead of purification, we also explored simple PCR product dilution in H_2O by factors ranging from 2:1 to 10:1. (see Appendix B for more details).

Additionally, we increased the number of molecules injected, n, by injecting more aggressively for the DNA duplex solution than was required for the primers. Discussions with collaborators [Personal communication with P.O. Ekstrøm, Massachusetts Institute of Technology, November 24, 2006], suggested limits approaching 9 μA (at 12 kV) for 60 s (540 μC). Typically, we injected for 120 s at 2 μA, $\sim$1 kV (240 μC). Under these conditions we were able to inject n=5$\times$10^9 molecules (1.5 ng) yielding SNR=256, or twice the amount required for M_f=0.005. Undoubtedly, additional sample preparation exploration can yield higher injection efficiencies, approaching the limits of 80 ng reported by Li *et al.* [170]. Comparing aggressive loading conditions in the UTMS (120 s $\times$ 2 μA = 240 μC) to standard loading conditions in Prof. Thilly's instrument (15 s $\times$ 2 μA = 30 μC), the UTMS still has a lower SNR, but only by 4.6$\times$.

8.2.2 Conversion between concentration and number of molecules

Since we have also used static fluorescent solutions rather than electrophoretic peaks to measure the performance of the end-column detection system, it is useful to be able to convert between them. We would like to know what is the equivalency between the static BSA-AlexaFluor solution concentration and the number of molecules in the peak, p. Also, for a given p, what is the number of molecules detected within the lenslet's focal volume, f?

To convert from concentration to number of molecules in the peak, p, one needs to know the volume occupied by the peak. Our experience and previous experiments [84, 121] show that the peak is approximately 2.5 mm long since it is observable for 30 s using point fluorescence detection while traveling at 5 mm/min. Assuming a

concentration of DNA, c_{DNA}	p	f R_c=1.5 mm	f R_c=2.5 mm
	V_p=11 nL	V_f=100 μm^3	V_f=750 μm^3
0.002 nM/L	1.3×10^4	0.1	1
0.02 nM/L	1.3×10^5	1	9
0.15 nM/L	10^6	9	68
0.2 nM/L	1.3×10^6	12	90
1.5 nM/L	10^7	91	682
2 nM/L	1.3×10^7	120	903
15 nM/L	10^8	900	6800
20 nM/L	1.3×10^8	1200	9000
150 nM/L	10^9	9000	68000
200 nM/L	1.3×10^9	12000	90000

Table 8.1: For a DNA peak with concentration, c_{DNA}, or number of molecules, p, the number of molecules detected, f, using two different lenslets can be computed, using Equation 8.10 and Equation 8.11.

uniform radial distribution in the 75 μm diameter capillary, the total volume occupied by the peak is

$$V_p = \pi r^2 l = 11 \times 10^6 \ \mu\mathrm{m}^3. \tag{8.9}$$

We have previously (Chapter 7, Section 7.4.1) found the focal volumes, V_f of lenslets with radii of curvature, R_c, of 1.5 mm and 2.5 mm, with apertures of 1 mm to be 100 μm^3 and 750 μm^3, respectively.

So the number of molecules detected, f, can be given by

$$f = c_{\mathrm{DNA}}\mathrm{Na}V_f, \tag{8.10}$$

where Na is Avogadro's number equal to 6.02×10^{23} molecules/M and c_{DNA} is the concentration of fluorophores in the sample (M/L). Similarly, p will be given by

$$p = c_{\mathrm{DNA}}\mathrm{Na}V_p. \tag{8.11}$$

In Table 8.1, the conversions between n and p for these lenslets is presented, along with the conversion to concentration equivalent.

Notably, the lenslets only sample a small fraction of the number of molecules in

the peak, as

$$10^{-5} < \frac{p}{n} = \frac{V_f}{V_p} < 10^{-6}. \tag{8.12}$$

For the LED intensity, these lenslets give the maximum SNR. There is promise for a higher SNR for the sample fluor concentration using higher intensities, such as from a 10 W cw argon ion laser. Our use of BSA-AlexaFluor solutions containing concentrations of 0.2-200 nM typically accurately simulates electrophoretic peaks containing 10^6-10^9 molecules.

8.3 Multi-capillary detection

Having demonstrated electrophoretic detection of primers with adequate SNR in single capillaries, we set out to demonstrate the scalability of the instrument. The heat exchanger, buffer reservoir, DNA microwells, gel loader, capillary array, and most importantly the detection system were all designed for multi-capillary capability, so we simply scaled them and began testing. In this Section we report on successful electrophoresis and end-column detection from arrays of 25-100 capillaries, as well as preliminary experiments with detection from 10,000 discrete fluorescent sources using the detection system.

8.3.1 25 capillary array

In our first multi-capillary experiment, we sought to electrophorese and detect labeled primers (see Appendix B), through an array of five capillaries staggered along the diagonal of a hypothetical 25 capillary array with 1 mm pitch. Experimental questions concerned the relative DNA speed (or electrical resistance) of the capillaries to each other and the amount of "cross-talk" or undesirable detection of fluorescence in neighboring lenslet channels.

The samples used for these experiments are AlexaFluor 488-labeled primers diluted in Millipore H_2O. Five sample vials were prepared, each with a 5 μL solution of concentration 16.6 nM (10^{10} molecules/μL). Each capillary was electrokinetically

loaded from its respective sample vial for 15 s at 2 μA (30 μC), at $\sim$1 kV. The capillaries were not loaded simultaneously, but instead the injections were serial, interspersed with a "drive-in" electrophoresis for 30 s at 8 μA each, or 40 μA total (8 μA $\times$ 5 capillaries = 40 μA total). After capillary 1 was loaded, all capillaries were electrophoresed for 30 s at 40 μA. Then capillary 2 was loaded, and all capillaries were electrophoresed for 30 s at 40 μA, and so on. In this way, we imposed a phase shift to the electrophoretic peaks of 240 μC. After the loading and driving-in, all capillaries were electrophoresed for an additional 3.5 minutes at 40 μA, then the current was reduced to 10 μA (2 μA/capillary) for detection. The sieving matrix was in Spectrumedix PCR-QC gel, an $\sim$3% LPA solution (see Appendix A). Elution began at 10-20 minutes after the current was reduced to 10 μA. The results are shown in Figure 8-4.

Approximately 1,000 12-bit CCD images were acquired at 0.29 Hz. Each image was divided into 25 uniformly sized, uniformly spaced regions corresponding to each lenslet (5$\times$5 lenslet array shown in Figure 8-4 (*upper right*)). The counts for the pixels in each region were averaged, background subtracted, and an image-varying bias was applied to compensate for drift.

These average intensities are shown as a function of time for the five capillaries and one region with two active neighbors (Figure 8-4 (*top*)). This resembles the common electropherogram for each capillary. Raw CCD images are included at several interesting times during the DNA elution (Figure 8-4 (*bottom*)).

Detection results indicate SNR varies from 50 to 800 for the electrophoretic peaks. Elution occurred 10-40 min after injection. So from the initial 5 μL DNA solution at 16.6 nM/L, we injection $\sim 10^{10}$ molecules into each capillary (see Figure 8-2).

These results show that the capillary array is functional, that the capillaries can be loaded with sieving matrix hydrodynamically and then loaded with DNA electrokinetically, and that DNA can be detected at the other end after electrophoretic capillary traversal. We can also conclude that the capillaries' inner diameters are aligned sufficiently well with for fluorescence excitation and collection. There is negligible cross-talk between the capillaries, as evidence by the lack of signal in the blank

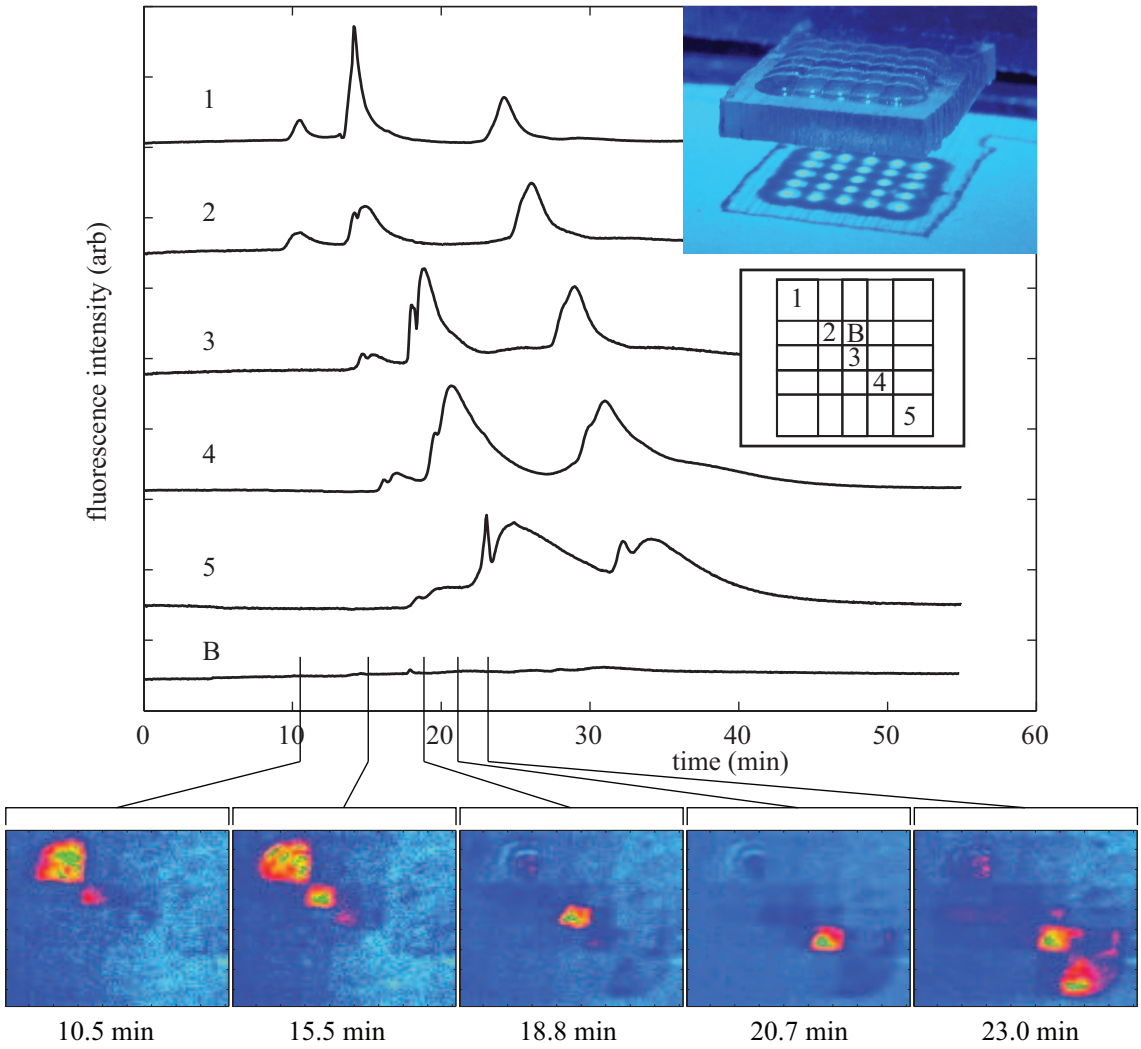

Figure 8-4: Five capillaries along the diagonal of a hypothetical 25 capillary array with 1 mm spacing were loaded with primers at 16.6 nM (10^{10} molecules/μL) for 15 s at 2 μA (30 μC). Injections were performed with a 2 minute phase-shift (at 2 μA) and electrophoresed to detection, from upper-right to lower-left. The lenslet array (*upper right*) was used to excite and collect fluorescence from the capillary tips. Regions of the raw CCD images (*bottom*) were averaged to yield common electropherograms. A blank sample, "B," with two active neighbors is also shown.

capillary, "B," with two active neighbors. Furthermore, there appears to be no residual fluorescence after the DNA has exited the capillary. The effective two minute temporal phase shift between the capillaries is somewhat preserved, but the variation suggests the utilization of a length calibration standard within each capillary would be prudent, as is standard practice in the field. The multiple peaks observed from a single-stranded DNA primer are possibly due to intra- and inter-primer binding, and do not negate these conclusions.

Additionally, the electrophoretic peaks appear to diffuse more in higher numbered capillaries. Apparently, the additional time that the DNA primers reside in the higher numbered capillaries proportionally affects the quality (peak capacity, SNR, etc.) of the separation. This suggests that the electrophoresis should be conducted as quickly as possible, within the sieving matrix electric field breakdown strength, and subsequently slowed for detection. Using this strategy, the entire run can be conducted in 10-30 min. The electric field breakdown strength has been measured in a variety of LPA-based gels to be about 33 V/mm.

8.3.2 100 capillary array

To further demonstrate the scalability of the UTMS instrument, we attempted to demonstrate electrophoresis and detection in 100 simultaneous channels. In these experiments a 100 capillary array and associated instrumentation were utilized. In Figure 8-5, the 100 lenslet array is shown aligned to the tips of 100 capillaries, bathed in the blue excitation radiation. The 100 capillary array is shown alone (with polyimide coatings for contrast) as well.

The experimental conditions were largely identical to those used in the 25 capillary sub-array experiment of Section 8.3.1, with a few minor differences. Approximately 1 mL of 20 nM (1.2×10^{10} molecules/μL), Millipore H_2O diluted, AlexaFluor 488-labeled primers were prepared. The capillaries were loaded from the 100 well microwell array described in Chapter 4, each well containing 300 nL of solution. Since the UTMS power supply (Spellman, CZE 1000R-2032) was only capable of 100 μA output, electrokinetic loading was performed at 1 μA/capillary for 30 s (30 μC), at

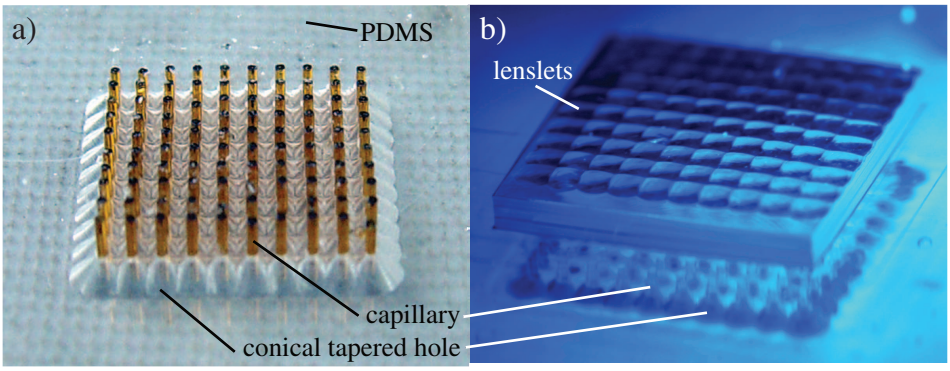

Figure 8-5: Experimental setup for 100 capillary array electrophoresis and end-column detection. (*a*) One-hundred capillaries, inserted into the buffer reservoir, are aligned for detection. (Note: the capillaries' polyimide coatings are removed before the experiment.) (*b*) The lenslet array is easily aligned to the capillary array using a pair of guide pins. Bathed in blue radiation, the system is ready for end-column fluorescence detection.

~0.5 kV. The capillaries were loaded simultaneously, and afterward electrophoresis was conducted at 1 μA $\times$ 100 capillaries = 100 μA at ~0.5 kV for 100 min, as limited by the power supply. The sieving matrix was in Spectrumedix PCR-QC gel, an ~3% LPA solution (see Appendix A).

To make things more interesting and assess cross-talk, only certain capillaries were loaded. Using the capillary array like a dot-matrix printer, we were able to create letters, as shown in Figure 8-6. This data is the concatenation of three non-consecutive experiments in which only some of the 100 wells were loaded; each letter was individually created on the 10×10 array. The concatenated frames shown in Figure 8-6 were selected for their uniform brightness, but of course the electrophoretic peaks are transient, thus the letters appear and then fade over the course of several minutes at this small current (1 μA/capillary).

The variation is brightness across a letter is due to the variation in electrical resistance and hence, DNA speed, in the capillaries. These are raw images that have not been processed, save for background subtraction and bias correction. There is some cross-talk evident. This motivated the development of a new buffer reservoir, in which the DNA must flow in a plane below the lenslet array focal plane to get to the electrode, which solved this minor problem.

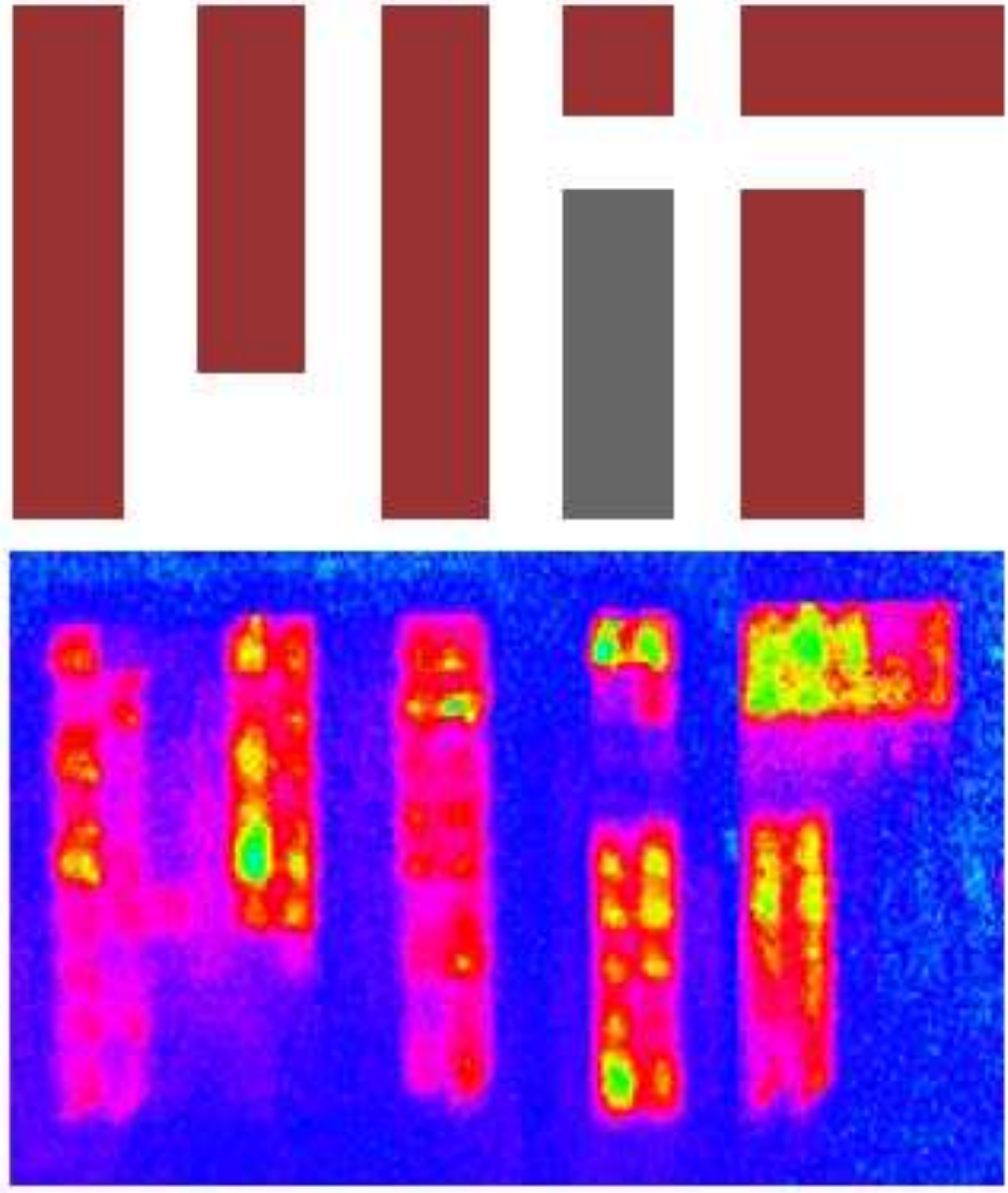

Figure 8-6: Demonstration of 100 capillary electrophoresis capability. This MIT logo is the concatenation of three experiments, one per letter.

8.3.3 10,000 capillary array simulation

Detection of optical emission from the 10,000 capillary array has been shown in Chapter 6, Section 6.3.2, imitating the conditions of fluorescence emission. We also performed an experiment to verify that a 100 mm $\times$ 100 mm spatial area with 10,000 discrete fluorescent sources could be imaged. This was reported in Chapter 7, Section 7.7.4.

8.4 Mutation detection

Mutation detection using end-column detection with SNR sufficient for identification of $M_f \geq 0.005$ is the *coup de grace* for this phase of UTMS research. Numerous conditions, configuration, samples, etc., were attempted before success. We start by very briefly summarizing what didn't work and then presenting the mutation UTMS mutation detection data relative to experiments on Prof. Thilly's system and the Amersham MegaBACE.

Briefly, attempts at mutation detection using the Spectrumedix PCR-QC matrix were foiled because of its poor spatial resolution; peaks overlapped to an extent that made measurement of the mutant fractions, M_f, impractical. The PVP gel system that has been used successfully by our collaborators [178] worked, though with suppressed fluorescence, on Prof. Thilly's system, and did not yield end-column fluorescence data despite persistent efforts. Results with a home-made linear poly acrylamide (LPA) matrix, were encouraging on Prof. Thilly's system, but the lack of availability and extremely high viscosity inhibited their application to the UTMS. Eventually, we settled on Spectrumedix Reveal High Resolution mutation detection matrix (Spectrumedix, MREV-HR-225-001, V#:3) for our mutation detection experiments in both Prof. Thilly's instrument and the UTMS. Subsequently, a successful combination of temperature, DNA concentration and injection conditions, electrophoretic running conditions were determined.

8.4.1 DNA sample and melting map

Samples containing wildtype/mutant (wt/mt) DNA homo/heteroduplexes were prepared or received from collaborators. These solutions contained various mutant fractions, M_f. The sample described herein was isolated from a cell line of human melanoma (skin cancer), THX cells, after efforts to prepare samples from recombinant DNA in *Escherichia Coli* (E. Coli) host were unsuccessful. The gene fragment was isolated from exon 15 of gene *BRAF*, which is of interest for its suspect role in carcinogenesis.[§] This gene is comprised of 18 exons spanning a region of 190,284 bp on chromosome 4. The specific gene fragment, isolated and amplified with PCR is as follows:

> (TTCCTTTACTTACTACACCTCAGatatatttcttcatgaagacctcacagtaaaaata ggtgattttggtctagctacag**T**gaaatctcgatggagtgggtcccatcAGTTTGAAACAGTT GTCTGGA)<u>CGGGCGGGGGCGGCGGGACGGGCGCGGGGCGCGGCG GGCG</u>-6FAM.

The sequence is written 5'→3'. Parentheses denote the natural gene fragment. Bold denotes the mutation T→A (wt→mt). Underlined denotes the GC clamp required for CDCE. The sample was labeled with 6-FAM on the 3' end, after the GC clamp, as shown. Capitalization denotes the primer regions. The cell line contains cells that are heterozygous for the mutation as well as wildtype homozygous. This leads to the presence of a mixture of wildtype and mutant gene fragments, and a corresponding mutant fraction, M_f. More details about the sample preparation can be found in Appendix B and have been published by Hinselwood *et al.* [179].

To separate mutant from wildtype DNA by denaturing capillary electrophoresis, we must apply denaturing conditions to the capillaries. As discussed in Chapter 2, Section 2.1, this is most easily achieved by externally applying heat. To determine the temperature required to optimally separate mutant from wildtype DNA, we use an algorithm written by L.S. Lerman *et al.* (originally [80]), which is based on the

[§]Mutations in the *BRAF* gene can prevent apoptosis, or programmed cell death, which can lead to uncontrolled proliferation–cancer.

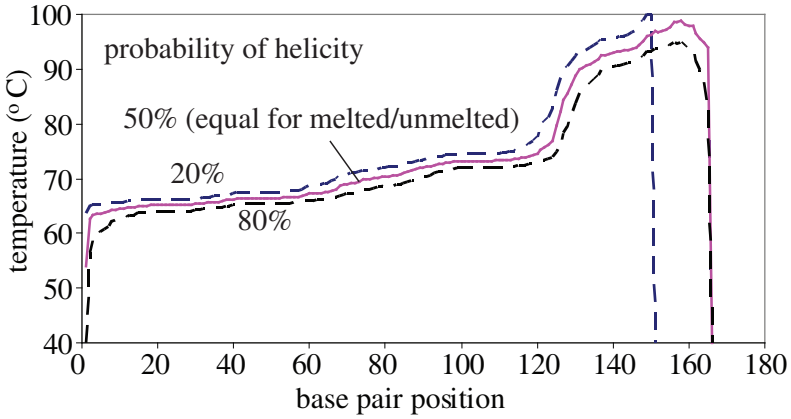

Figure 8-7: Melting map of *BRAF* exon 15, computed using the algorithm written from L.S. Lerman *et al.* (originally [80]). The melting map represents the equilibrium temperature for equal probability of the helical and melted (denatured) states, and is an important tool for establishing the experimental conditions for denaturing capillary electrophoresis.

theory of Poland, Fixman, and Freire [81, 82, 83]. The so-called "melting map," or denaturing temperature as a function of base pair position, represents the equilibrium temperature for equal probability of the helical and melted (denatured) states. For our 167 bp sequence shown above, the melting map is shown in Figure 8-7. In addition to the curve for 50% helicity, the curves showing the melting temperatures for 80% and 20% probability of helicity are included. From this data, we choose the melting temperature of 69 °C as an initial experimental setting–the mean melting temperature of the 127 bp native sequence (excluding the GC clamp). Thus, at this temperature, there is a 50% probability that the wildtype homoduplex will be denatured. Since the mutant homoduplex will have a different probability of helicity, it will have a different migration velocity through the denaturing zone. We experimentally converged on 70 °C for the optimal CDCE denaturing temperature, 1 °C above the theoretical optimum. The difference can be explained by heat loss between the water recirculator and capillary water jacket, or a temperature sensing error in the recirculator. Such a systematic error can be easily compensated once identified.

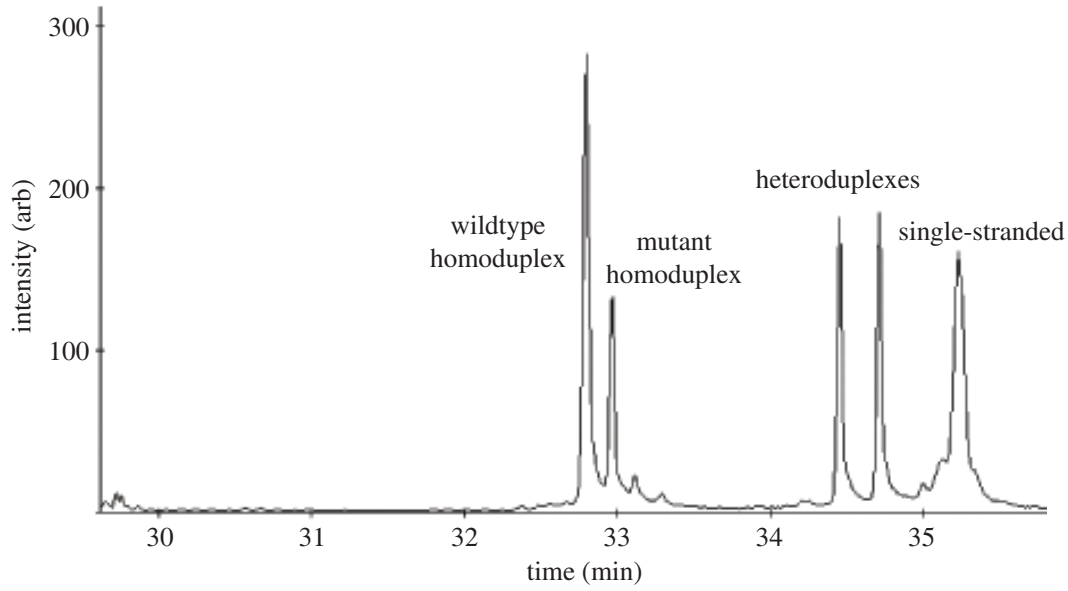

Figure 8-8: Denaturing capillary electrophoresis (using CTCE) separation of *BRAF* exon 15 mutant from wildtype on Amersham MegaBACE instrument. Running conditions: Amersham MegaBACE, MegaBACE long read matrix, 45-48 °C cycled 20×, 600 mm length of heated zone, 1× raw PCR products before shipment, 200 μC injection, run at 10 μA.

8.4.2 Experimental results

The sample was run on the Amersham MegaBACE under cycling temperature capillary electrophoresis (CTCE) using a commercial LPA matrix with 7 M/L urea, which reduces the melting temperature required by 3 °C/M/L. The temperature was cycled between 45-48 °C at approximately 1/120 Hz. Electrokinetic injection was performed at 10 μA for 20 s (200 μC), at 10 kV from raw PCR products. Electrophoresis was run at 9 μA (at 9 kV). The results are shown in Figure 8-8.

After verifying the products using the MegaBACE, the sample was shipped across the Atlantic Ocean from our collaborators (P.O. Ekstrøm *et al.*) for our experiments. During shipping the sample was unintentionally and undesirably modified, most likely by heating it. The damage was repaired by incubating it with a high-fidelity enzyme, Pfu. This procedure and its effects are described in Appendix B.

After this treatment, the sample was run on Prof. Thilly's instrument using Spectrumedix Reveal High Resolution mutation detection matrix. The results are shown

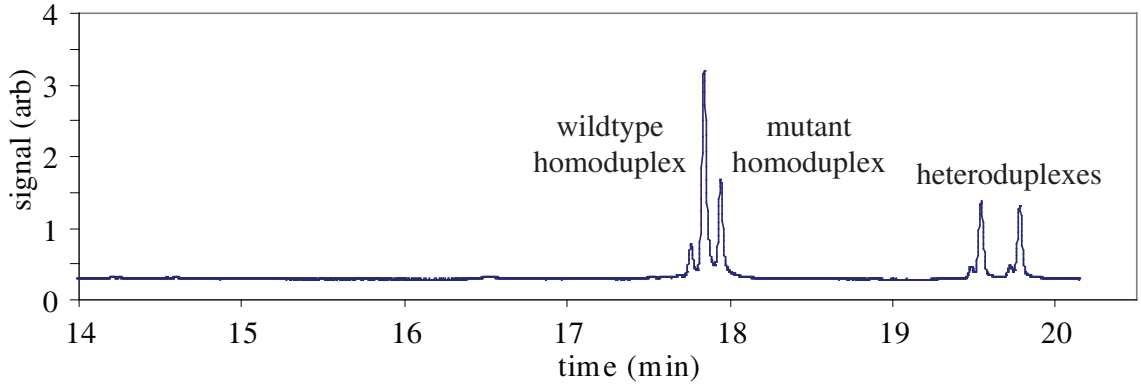

Figure 8-9: Denaturing capillary electrophoresis (using CDCE) separation of *BRAF* exon 15 mutant from wildtype on Prof. Thilly's instrument. Running conditions: Prof. Thilly's Instrument, Spectrumedix HR gel, 70 °C, 130 mm length of heated zone, 0.1× raw PCR products after incubation with Pfu, 30 μC injection, run at 8 μA.

in Figure 8-9. For the experimental conditions, the raw PCR products after Pfu treatment were diluted 10× in Millipore H_2O. From a 5 μL aliquot, samples were electrokinetically injected onto the 320 mm long capillary column on Prof. Thilly's instrument for 15 s at 2 μA (30 μC), at ∼1 kV. The sample was electrophoresed at 8 μA, at ∼5 kV. After passing through a 130 mm denaturing zone heated by a water jacket to 70 °C, the sample passed by a window located 250 mm from the injection end, where fluorescence emission was measured at 10 Hz from a PMT with gain settings of 0.65 kV and 10^7 V/A.

The experiment with this sample was then repeated on the UTMS using Spectrumedix Reveal High Resolution mutation detection matrix. The results are shown in Figure 8-10. The experimental conditions are identical to the run on Prof. Thilly's instrument, except for a more aggressive injection of 120 s at 2 μA (240 μC), at ∼1 kV, and electrophoresis was slowed to 2 μA (from 8 μA) for detection. Identical capillary length, water jacket length and temperature, sample loading volume, and sample concentration were used. The time axis has been shifted for the data from Prof. Thilly's instrument to align the results for comparison.

After this encouraging result, we increased the concentration to 0.5× raw PCR products with all other conditions identical to the previous experiment. This resulted

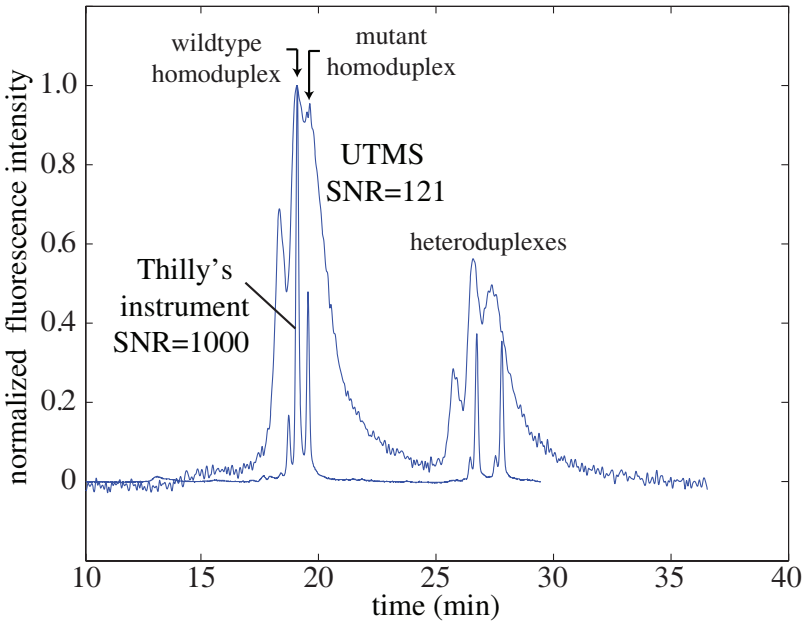

Figure 8-10: Denaturing capillary electrophoresis (using CDCE) separation of *BRAF* exon 15 mutant from wildtype on both UTMS and Prof. Thilly's instrument. Running conditions: Prof. Thilly's Instrument: Spectrumedix HR gel, 70 °C, 130 mm length of heated zone, 0.1× raw PCR products after incubation with Pfu, 30 μC injection, run at 8 μA. UTMS: Spectrumedix HR gel, 70 °C, 0.1× raw PCR products after incubation with Pfu, 240 μC injection, run at 2 μA during detection, 8 μA during separation.

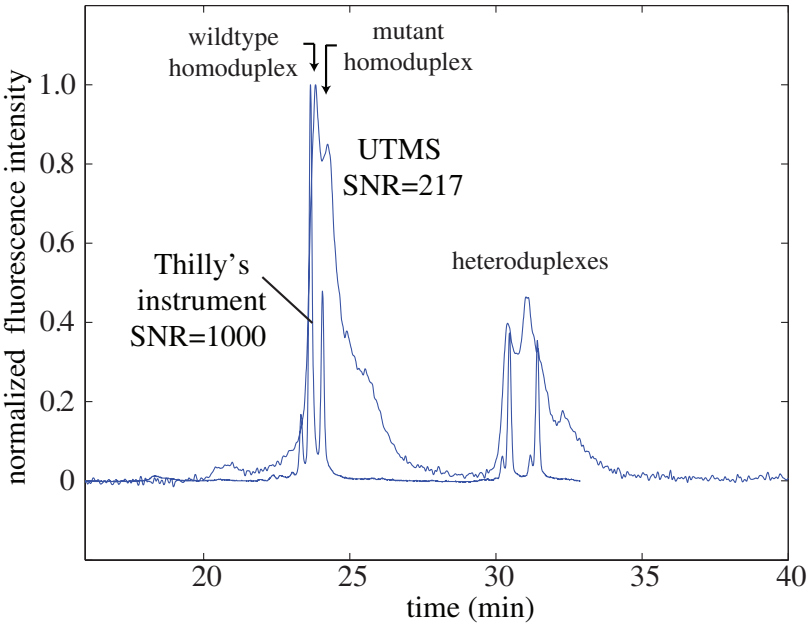

Figure 8-11: Aggressively loaded denaturing capillary electrophoresis (using CDCE) separation of *BRAF* exon 15 mutant from wildtype on both UTMS and Prof. Thilly's instrument. Running conditions: Prof. Thilly's Instrument: Spectrumedix HR gel, 70 °C, 130 mm length of heated zone, 0.1× raw PCR products, 30 μC injection, run at 8 μA. UTMS: Spectrumedix HR gel, 70 °C, 0.5× raw PCR products after incubation with Pfu, 240 μC injection, run at 2 μA during detection, 8 μA during separation.

in further improved SNR for the UTMS measurement, as shown in Figure 8-11. This result, in addition to demonstrating the higher SNR possible, illustrates the repeatability of the experiment. Higher concentrations of raw PCR products do not generally improve the SNR beyond this concentration however, because of the factors discussed in Section 8.2.1.

From this experiment and runs in Figure 8-8 and Figure 8-9, we tabulated the SNR of the wildtype homoduplex (homo) and one of the identical mutant heteroduplexes (hetero) and calculated the mutant fraction from the electropherograms according to Equation 8.4. The results are shown in Table 8.2. The three values of M_f are in good agreement; we did not have a way to verify the true M_f independently. In this type of experiment with a skin cancer cell line, M_f can be greater than 50%, as this fraction represents the percentage of the cells that contain the mutant gene fragment.

Were the mutant fraction in this experiment only 0.5%, the SNR on all three

	Signal-to-noise ratio SNR homo/hetero	Mutant fraction M_f
MegaBACE	350/162	35.2%
Prof. Thilly's Instrument	1000/500	34.2%
UTMS	217/101	35.2%

Table 8.2: Results from mutation detection experiments. The signal-to-noise ratio (SNR) from the UTMS is compared with two other instruments. The mutant fractions measured by the three instruments agree well. Furthermore, the SNR on all three is sufficient to detect a $M_f \geq 0.005$.

instruments would be sufficient to detect it. This was determined by attenuating the mutant heteroduplex peaks, adding back the noise, and computing M_f according to Equation 8.4. The SNR for M_f=0.005 is 3.75 in the UTMS using this data processing technique. Experiments were not run with real samples containing M_f=0.5%. The generally accepted threshold for detection of a signal is SNR=3.

We also notice that the UTMS peaks, measured with the end-column detector, are substantially broader than those detected by side-column fluorescence detection, as is the case in Prof. Thilly's instrument (dia-illumination), as well as the Amersham MegaBACE (epi-illumination). This is undesirable, but does not reduce our ability to measure M_f in this experiment.

The broadening could be caused by differences in the length of the fluorescence measurement zone between the instruments. More specifically, the UTMS will collect fluorescence over a depth-of-focus, of 25 μm, for these 1 mm aperture, 5.1 mm focal length lenslets. Prof. Thilly's instrument interrogates fluorescence across only the width of the point spread function, $\sim$6 μm for a 3 mm diameter laser beam focused at 30 mm. However, this does not entirely explain the difference.

In Figure 8-9, the peaks on Prof. Thilly's instrument pass by the detector in about 3 s. Based on the time for to reach the detector, we know they are traveling at average speed of 13 mm/min (216 μm/s) at 8 μA. Thus the peaks are approximately 650 μm in physical extent within the capillary. Thus the side-column detector, with its 6 μm wide interrogation zone, is virtually a point detector. In the UTMS, the

peaks should be the same length, traveling at 25% of the speed in Prof. Thilly's instrument (at 2 μA). Thus they should be visible for only 12 s. But from Figure 8-10, for example, the peaks are visible for 120-240 s. The 25 μm interrogation length alone can not explain this. The measured peak duration is only possible if fluorescence is being excited and detected when the DNA is outside of (below and above) the focal volume. Indeed, excitation light traveling axially through the lenslet can excite fluorescence and emission traveling axially from that stimulation can reach the CCD detector. We see a 10-20$\times$ broadening of the peaks using the end-column detector, as compared with a point, or side-column detector. Broadening of this description is systematic of this design and has been observed in all data, confounding isolation of the homoduplexes from each other as well as heteroduplexes from each other. Such separation resolution is not required, but can be useful in interpreting the electropherogram. This is a major challenge to overcome in future work. Ideas to alleviate this problem should not compromise SNR, as it is already marginally acceptable. Initial suggestions include a pseudo-side-column detection that will enable 2-D packing density without axial illumination, or extending the work of Zhang *et al.* [139], described in Chapter 7, albeit with extremely costly and complex optics required.

8.5 Parameter variation

Many parameters were varied to optimize the mutation detection experiments on the UTMS. Among the ones already discussed in this Chapter include DNA loading time and current and concentration of DNA and other charged molecules in the solution. In this Section we address a few other parameters whose variation may promise increased SNR or separation efficiency, which could potentially reduce run times, volume of reagents or analytes, or ambiguity in the determination of M_f.

8.5.1 Fluorescence enhancement: intercalating dye

In an attempt to increase fluorescence of the duplex DNA gene fragments, we incubated the DNA with an intercalating dye. An intercalating dye, such as SYBR Green, binds to double-stranded DNA non-specifically at a ratio of 5 bp/dye molecule, permitting sensitive fluorescence detection, even on isolated single DNA molecules.

We diluted SYBR from 10,000× in DMSO, a solvent, to 1× in 0.01× in Spectrumedix Reveal High Resolution Running Buffer (see Appendix A). Five μL of raw 6-FAM-labeled PCR products containing mutant and wildtype homo/hetero-duplexes were incubated with 1× SYBR solution.

Using Prof. Thilly's instrument, this solution was injected on the capillary column as usual (15 s at 2 μA = 30 μC), and ran them at the appropriate denaturing conditions. Though the fluorescence was about 2× brighter, the separation quality was far worse. Thus this course was discontinued.

8.5.2 Temperature

The optimal denaturing temperature predicted by the Lerman algorithm discussed in Section 8.4.1 was 69 °C. We explored the proximal temperatures to observe their effects of separation. As has been observed by Khrapko *et al.* [84], we observed about a 1 °C range of good separation. Below this temperature, the homoduplex peaks are overlapping, Above, the heteroduplex peaks are not well separated from the homoduplexes and each other. An example of our results illustrating this is shown in Figure 8-12. After observing these electropherograms, we settled on 70 °C as the optimal CDCE separation temperature for this exonic gene fragment. In practice, those skilled in the art can select an adequate operating temperature from the Lerman algorithm prediction without experimental exploration. The predicted optimal temperatures for all ∼250,000 exonic gene fragments in the human genome were presented in Figure 2-1 in Chapter 2.

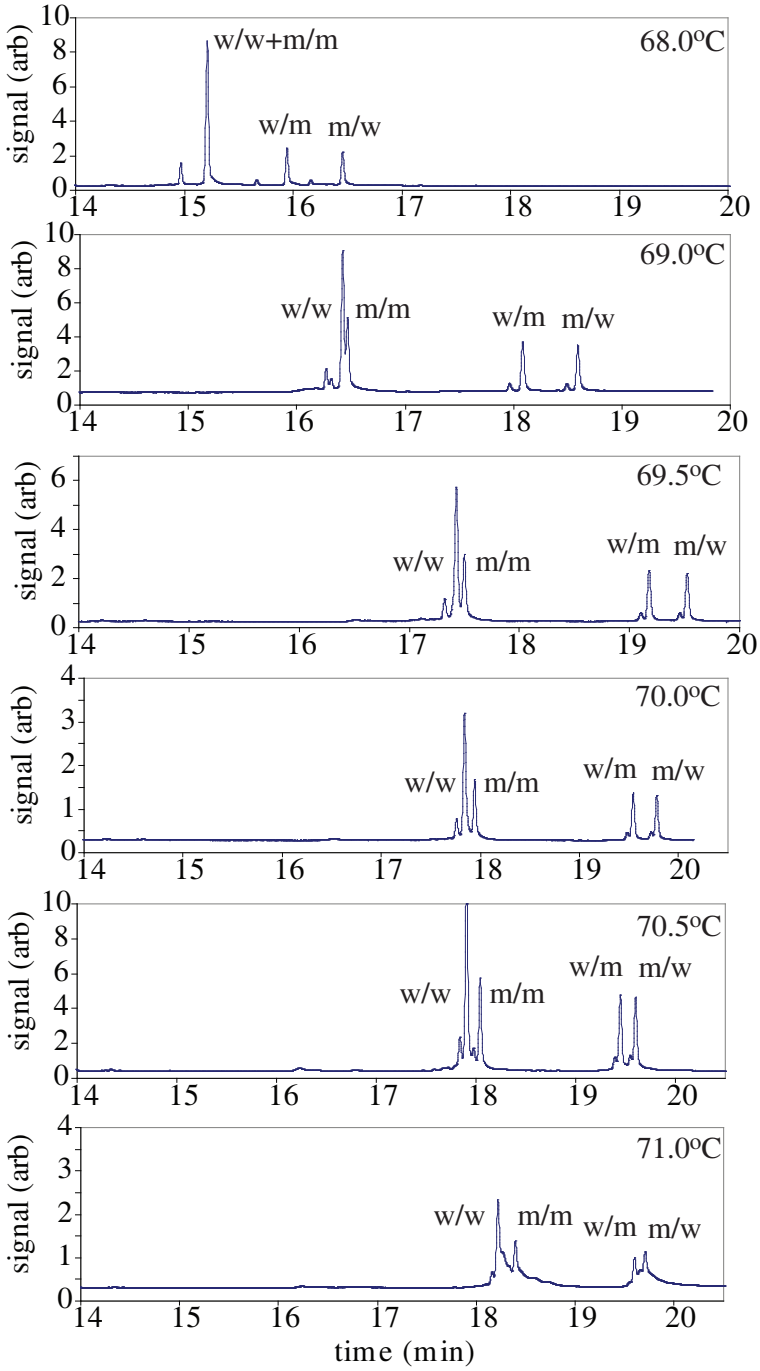

Figure 8-12: Denaturing capillary electrophoresis (using CDCE) separation of *BRAF* exon 15 mutant from wildtype on Prof. Thilly's instrument for several temperatures in the range of 68-71 °C. Running conditions: Prof. Thilly's Instrument: Spectrumedix HR gel, 68-71 °C as labeled, 130 mm length of heated zone, 0.1× raw PCR products after incubation with Pfu, 30 μC injection, run at 8 μA.

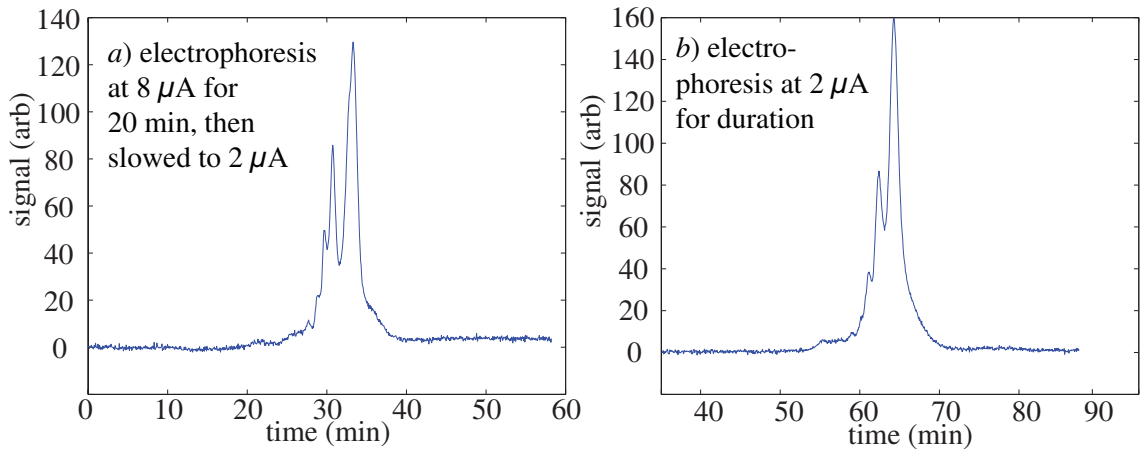

Figure 8-13: Typical speed (a) and intentionally slowed (b) denaturing capillary electrophoresis (using CDCE) separation of purified but possibly degraded PCR products on UTMS instrument at 68 °C. Slowing the DNA traversal through the denaturing zone does not improve the separation efficiency; in fact, diffusion reduces it. Running conditions: UTMS: Spectrumedix PCR-QC gel, 68 °C, 130 mm length of heated zone, 1× purified PCR products eluted in running buffer, 240 μC injection.

8.5.3 Speed of migration

Up to the dielectric breakdown strength of the sieving matrix, the electrophoretic current, and hence the speed of migration of the DNA, can be increased. In practice we typically electrophorese at a current limited to 1-10 μA (3-26 V/mm). At 8 μA, ~200 bp duplex DNA migrates at an average speed of 12-16 mm/min, with the speed depending on the gel viscosity (i.e., slower values measured in Spectrumedix Reveal High Resolution mutation detection matrix, faster values measured in Spectrumedix PCR-QC matrix). The speed is proportional to current. We sought to investigate whether slowing the speed of migration through the denaturing zone had a beneficial effect on separation, keeping the gel matrix constant. Figure 8-13 illustrates the effects of reduced speed on the quality of separation. There appears to be no benefit to reducing the speed during electrophoresis. Rather, diffusion effects probably degrade the quality of separation at slower speeds.

For the experimental conditions, purified PCR products (see Appendix B) in a concentration of 10^{11} molecules/μL were electrokinetically injected for 120 s at 2 μA (240 μC), from a 5 μL aliquot. The samples were run at 68 °C, the theoretical

optimum for these gene fragments. In Figure 8-13*a*, the sample was electrophoresed at 8 μA (at $\sim$5 kV) until a time just before elution of the peaks then slowed to 2 μA (at $\sim$1 kV). For the experiment in Figure 8-13*b*, the sample was electrophoresed at 2 μA (at $\sim$1 kV) for the entire separation run. The capillary was 320 mm long and the heater jacket was 130 mm long. The sieving matrix was Spectrumedix PCR-QC, and consequently the separations are of poor resolution. Additionally, we suspect PCR product degradation explains the uncommon peak amplitudes/locations. Regardless, the conclusions regarding migration speed are valid.

8.5.4 Length of capillary and denaturing zone

The heater length used in both Prof. Thilly's instrument and the UTMS for experiments in Section 8.4 was 130 mm. We were interested to know if the separation of the peaks could be improved with increased heater length. Since migration speeds of the mutant/wildtype duplexes differ during traversal of the denaturing zone, their separation is proportional to the length of the denaturing zone. However, the peaks will broaden due to diffusion and increased capillary length, thus increasing the peak widths. We ran two different sets of experiments to explore these competing effects. The results showed that increasing the heated, denaturing length has a net benefit.

On Prof. Thilly's instrument, we ran three consecutive separation runs, each with effectively increasing heater jacket lengths. The heater jacket was always 130 mm, but to increase its effective length, the capillary was translated during the separation. Capillaries were cut to 320, 420, and 520 mm lengths. The effective heater lengths were, respectively, 130, 230, and 330 mm. The results are shown in Figure 8-14. To display the runs in this comparative manner with aligned peaks, we scaled the time axis by 2 and 4.4 respectively, for the successively longer capillaries and heaters. The results indicate that increasing the heater length does improve the resolution, but the improvement is not entirely proportional to heater length. Additionally, the broader peaks due to diffusion have a lower SNR, which could be problematic in the UTMS in which our SNR is marginally acceptable.

For the experimental conditions, raw PCR products (see Appendix B) were elec-

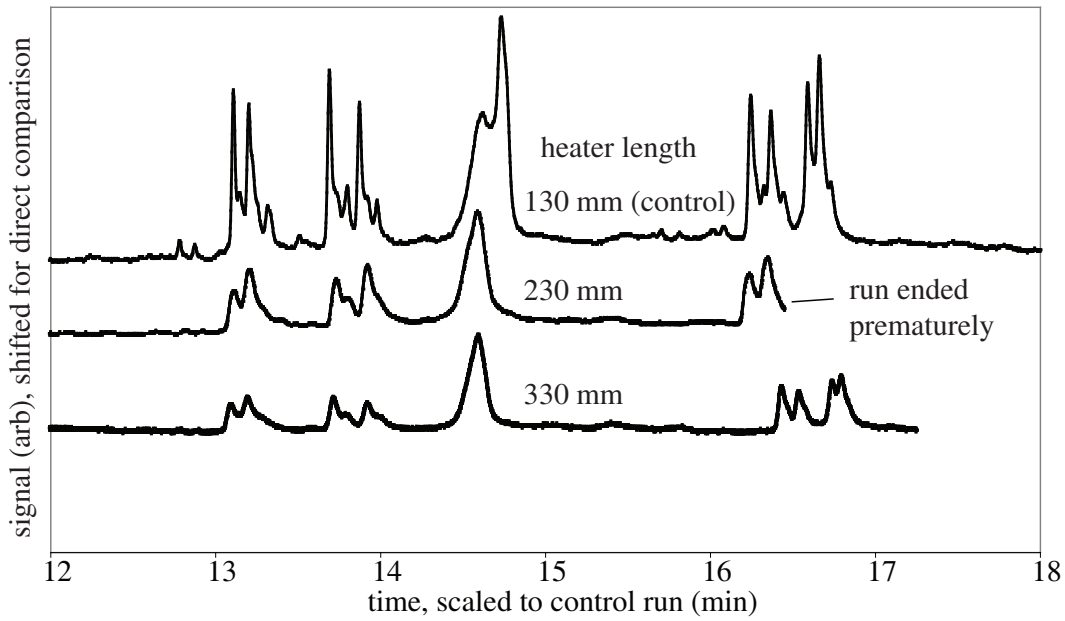

Figure 8-14: On Prof. Thilly's instrument, denaturing capillary electrophoresis (using CDCE) separation using three different lengths of denaturant zone. Capillaries: 320, 420, 520 mm; Denaturing zone: 130, 230, 330 mm, respectively. The time axes were scaled to align the peaks for direct comparison. The vertical axis are directly comparable. Though increasing the denaturing zone does improve the separation efficiency, there is a corresponding loss of SNR. Running conditions: Prof. Thilly's Instrument: Spectrumedix HR gel, 72 °C as labeled, 0.1× raw PCR products before incubation with Pfu, 60 μC injection, run at 8 μA.

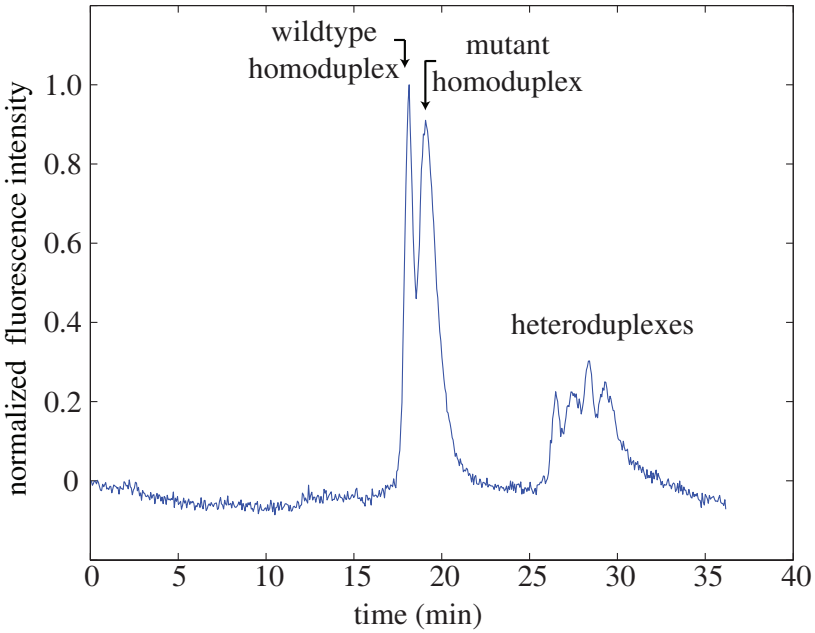

Figure 8-15: On the UTMS instrument, denaturing capillary electrophoresis (using CDCE) separation of *BRAF* exon 15 mutant from wildtype on UTMS with longer denaturing zone (180 mm instead of 130 mm). There is substantial improvement in separation resolution, as compared with Figure 8-10; elution occurs at a later time due to a longer denaturing zone. Running conditions: UTMS: Spectrumedix HR gel, 70 °C, 180 mm length of heated zone, 0.1× raw PCR products after incubation with Pfu, 240 μC injection, run at 2 μA during detection, 8 μA during separation.

trokinetically injected at for 30 s at 2 μA (60 μC), from a 5 μL aliquot. The samples were electrophoresed at 8 μA for the entire separation run, with intermittent pauses to translate the capillary as required. The sieving matrix was Spectrumedix Reveal High Resolution mutation detection matrix. Duplex peaks are observed because this sample was not incubated with Pfu enzyme after shipment (see Appendix B). After passing through the denaturing zone heated by a water jacket, the sample passed by a window located 70 mm from the elution end, where fluorescence emission was measured at 10 Hz from a PMT with gain settings of 0.65 kV and 10^7 V/A.

An experiment was also performed on the UTMS for comparison with Figure 8-10, which used a denaturing zone length of 130 mm. Using the same capillary length of 320 mm, the water jacket was translated along the capillary, increasing its effective length to 180 mm. The results are shown in Figure 8-15.

There is clearly substantial improvement in resolution as compared with Figure 8-10. The thermal control system described in Chapter 2 will apply heat over a length of approximately 270 mm, for a 300 mm capillary. Thus, we can expect results at least as good as these for the UTMS instrument run with the UTMS thermal control system, rather than the water jacket of Prof. Thilly's instrument.

8.5.5 Non-DC voltage electrophoresis

The efficiency of DNA mutant separation by denaturing capillary electrophoresis depends on a variety of factors including gel type (e.g., percent acrylamide), electric field strength, and DNA fragment length. Another parameter which is not often exploited is electric field modulation. Rather than a constant field, what if the driving potential and consequently, the electrophoretic current, is modulated? What amplitudes and frequencies might be useful? This technique, known in the literature as pulsed-field gel electrophoresis, has been utilized to separate lengthy (i.e., kB), DNA molecules according to length, but not for short, several hundred base fragments as we are using (e.g., [180, 181, 182, 183]). Further, its applicability to denaturing capillary electrophoresis has not been explored. The optimal pulsed-field frequency versus DNA molecule size is shown in Figure 8-16 for length based separations, adapted from work by Demana *et al.* [180]. The optimal frequency appears to rise logarithmically towards an asymptote at approximately 100 Hz for 5-10 kB length DNA.

As a side investigation, we built a software and hardware interface for the electrophoretic power supply to apply a 1 Hz square wave. Using our existing power supply, modulation at higher frequencies was not possible. In general, modulating kV scale voltages at tens to hundreds of Hz is a challenging electrical engineering problem. The results for DNA mutation detection with a different, but similar, wild-type/mutant sample are shown in Figure 8-17. In this figure, we have overlaid the results from our 1 Hz square wave drive voltage with a DC experiment, after normalizing the time to the DC experiment. There is virtually no improvement in DNA separation efficiency at this frequency, but this remains an interesting area for investigation in the future.

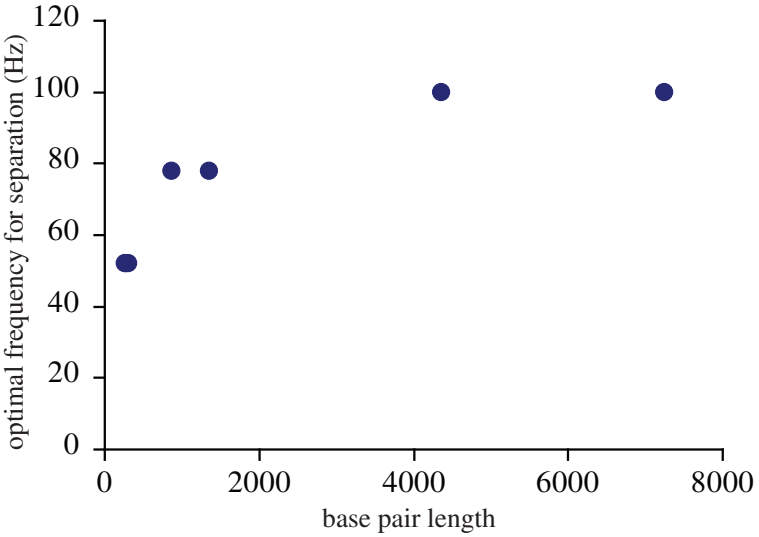

Figure 8-16: Optimal pulsed-field, non-AC, drive voltage frequency as a function of the DNA fragment length. This optimization is for length based separations, not sequence based separations, as in denaturing capillary electrophoresis, but the results should be generally applicable. Interpreted from Demana *et al.* [180].

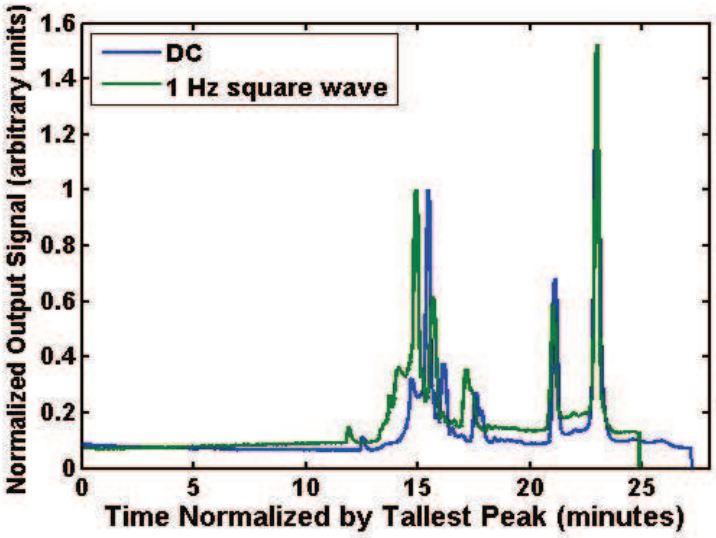

Figure 8-17: Experimental investigation into the effects of low frequency (1 Hz) drive voltage on CDCE separation versus typical DC. The separation efficiency difference is negligible for this frequency comparison. Higher frequencies (e.g., 10-50 Hz), could be interesting for future work, despite the challenges of switching high voltages (kV) rapidly.

Chapter 9

Conclusions and future work

In this thesis we have described progress towards the creation of a "mutational spectrometer" instrument containing up to 10,000 capillary channels and enabled with subsystems for loading, separating, and detecting fluorescently-labeled DNA. This effort was greatly aided by a team of 14 undergraduate research assistants whom I directly supervised, and several collaborators, critically Drs. Cathy Hogan and Per Ekstrøm, in addition to my committee members.

A variety of subsystems were created for this instrument, which were eventually integrated for electrophoresis experiments described in the preceding Chapter. In use, a capillary array of 1-10,000 capillaries is assembled (Chapter 6) and installed in the thermal control chamber (Chapter 2) which heats them to the appropriate DNA denaturing temperature. A viscous polymer matrix is then hydrodynamically loaded (Chapter 3) into this array of capillaries. Next, DNA solutions are placed into an array of 250 nL wells (Chapter 4), and this well array is interfaced to the capillary array to electrokinetically inject DNA into all capillaries simultaneously. After injection, the well array is replaced by a reservoir containing buffer solution (Chapter 5), which, like the well array, interfaces to the capillary array at both ends using a micro-fluidic connector. The optical detection system (Chapter 7) is aligned to the capillaries, and electrophoresis is conducted. Under electrophoresis, the DNA migrates through the length of the capillary array, until it is detected at the other end, just prior the time when it elutes from the capillaries and migrates towards the

electrode. Detection and sequestration of the eluting DNA in independent wells is also possible by replacing one of the buffer reservoirs with a buffer-filled well array.

9.1 Major contributions

The major contributions in this work are the following:

- Design and demonstration of an end-column fluorescence detection technology that permits tightly packed capillary arrays as required for ultra-high throughput genetic analysis. This system has a limit of detection of 10^7 molecules, comparable to the state-of-the-art capillary array electrophoresis detectors that utilize laser-induced fluorescence. Our system has the advantages of 100× more capillary detection capability and 10× reduction in cost, but suffers from reduced, though adequate, temporal resolution in mutation detection applications. (Patent pending, application 60/651,681)

- Invention of a milling/microforging fabrication technique for microlens arrays that is equal or superior to all other manufacturing technologies for such arrays in aspects such as cost, time, size capability, figure error, roughness, or fill factor.

- Development and implementation of capillary array manufacturing technologies for assembly and constraint of arrays of 1-10,000 capillaries with accuracy and repeatability on the order of tens of micrometers in three axes. All capillaries in the array can be replaced for refurbishment.

- Development and demonstration of a hybrid MEMS manufacturing technique for micro-well array fabrication that utilizes electrical discharge machining of highly doped silicon and subsequent polishing to readily integrate these parts into the full suite of semiconductor fabrication processes including deposition and bonding.

- Design, manufacture, and implementation of a microfluidic connector technology, which we have shown to work for interfacing arrays of 100 capillary ports repeatably, with fluidic sealing, and fluidic and electrical connectivity.

- Integration of a suite of subsystems to make a functional 100 capillary array electrophoresis instrument, with all aspects scalable or demonstrated at 10,000 capillary capacity.

9.2 Future work

There is still a great deal of work to be done to create an easy to use, repeatable and reliable, 10,000 capillary ultra-high throughput mutational spectrometer. Throughout the body of this thesis, comments about the directions for recommended future work were included. In this section, we briefly mention the more general areas for activity.

The thermal control system and gel loader are capable currently of operating at the 10,000 capillary capacity, but could be modified to make them more rapidly accessible for repeated experiments. The microwell and buffer reservoirs array, at 100 well capacity, will need to be remade with 10,000 wells and ports, using the described processes.

The detection system can be readily upgraded with optical components for 10,000 capillary capacity; it is currently limited to approximately 1,000. Before this undertaking, further study of the temporal resolution is suggested, and redesign may be necessary to reduce or eliminate band broadening and increase SNR as much as possible.

The capillary array is the biggest concern and open question presently. Assembly and constraint was demonstrated, but some challenges were encountered at the 10,000 capillary level that warrant redesign. In particular, the constraint device should at least be modified to suit this capacity by reducing the stiffness of the compressible layer. Partial reassembly and testing, though tedious, can expose continued design weaknesses. Additional guides at the capillary tips may be required to ensure position

accuracy across the full array.

The other major concern is interfacing to the array. The 100-port microfluidic connector works well, but interfacing repeatedly to the 10,000 capillary array will undoubtedly lead to fractured or displaced capillaries, quickly degrading instrument performance. Especially on the detection end, where the capillary coating have been removed to expose bare fused silica, fragility warrants alternative design considerations. A new design approach that minimizes or eliminates the need for repeated capillary array interfacing at one or both ends should would be superior to the current design.

Electrophoresis in 1,000 or 10,000 capillaries simultaneously would require, respectively, 8 and 80 mA, at up to 10 kV. At these voltages, these currents are quite deadly and safety precautions should be taken before this can be tested. A separate room with interlocks to the power supply should be implemented so that no person can be nearby when the power is on.

As the highest throughput yet obtained with capillary array electrophoresis instrumentation is 768 channels, one should prudently proceed to the 1,000 capillary level and demonstrate mutation detection with this capacity before proceeding to the full 10,000 capillary capability.

9.3 Alternative applications

The ultra-high throughput mutational spectrometer instrument can be useful in a range of applications. An example, we introduce genomic and proteomic imaging. In this application, we envision spatial mapping of biomarkers from tumor biopsies for enhanced identification and scientific understanding of gene expression in tissue. Currently, several tissue samples are taken from a tumor and/or nearby healthy tissue for diagnosis [184]. Gene expression (RNA) from these tissues can be analyzed with arrays of cDNA, or gene chips (e.g., Affymetrix, Santa Clara, CA). Increasing the number of biopsies in the tumor site and surrounding tissue to 50-1000 with spatial separation of 1-3 mm has been recognized as advantageous for successful tumor de-

tection and elimination of false positives [185, 186]. Using our capillary arrays and coupled detection system, we could analyze thousands of tissue locations at the spatial frequency matched to the lens array. A tissue section could be diced, purified, fluorescently-labeled, loaded, separated, and imaged with direct spatial correspondence for high resolution mapping of protein concentration or RNA markers.

l

Appendix A

Consumables

Denaturing capillary electrophoresis instrumentation requires a variety of consumables. In this appendix, we define the specific gel matrices, buffer solutions, and washes that have been used along with recipes when available (non-proprietary). In addition, we present techniques for removing capillary coatings.

A.1 Gel matrices

A variety of gel matrices were used for these experiments:

1. Spectrumedix PCR-QC matrix (Spectrumedix, MPCR-500-004).
2. Spectrumedix Reveal High Resolution mutation detection matrix (Spectrumedix, MREV-HR-225-001, V#:3).
3. Home-made linear poly acrylamide (LPA) matrix.
4. PVP matrix.
5. MegaBACE long read matrix (GE Healthcare, US79676).

The Spectrumedix products were proprietary, so the recipe is not known. Based on discussions with the company and analysis, we can conclude only that the products are approximately 3% and 5% LPA (for PCR-QC and Reveal mutation detection

matrices, respectively), without urea or dye. The PCR-QC matrix is advertised to be capable of separating PCR products or DNA fragments in the range of 100-800 bp, while the Reveal matrix is advertised as capable of single base mutation detection for fragments up to 500 bp in length.

The home-made LPA was prepared as follows: 5% weight/volume acrylamide solution in TBE was deoxygenated by argon bubbling in an ice bath for 10 min and any contact of the solution with air was avoided until polymerization was complete. TEMED and ammonium persulfate were added to 0.03% and 0.003%, respectively. The polymerizing solution was immediately dispensed into 10 mL glass syringes and left at 2 °C for several days. The matrix was dispensed from the 10 mL syringes into a high-pressure loading device for filling the capillaries.

The PVP matrix was prepared as follows [178]: Granulated polymer powder of PVP with molecular mass of 360 kDa was added 6% w/v to the PVP buffer (described in Section A.2), and 7 M urea.

The MegaBACE long read matrix recipe is proprietary, but is known to contain 3% LPA and 7 M/L urea.

A.1.1 Capillary coating

All of the gel matrices are dynamically coating except for the home-made LPA, which requires a pre-treatment. In this treatment, capillary coatings that covalently link chains of the polymer to the silano groups of the silica are required before use [105, 106] to suppress the detrimental effects of electroosmotic flow [107]. This procedure has also been described by Thilly *et al.* [187]. In the dynamically coating matrices, such as those based on polyvinylpyrrolidone (PVP) [108], and poly(dimethylacrylamide) [109], the inherent chemical composition works to shield the capillary walls, suppressing electroosmotic flow [110, 111].

A.1.2 Urea

Urea can reduce the DNA melting temperatures by 3 °C/M/L, but the resolution of the separation can be somewhat degraded. The PVP gel used by our collaborators typically contains 7 M urea, which lowers the melting temperature by 21 °C.

A.2 Buffer solutions

In conjunction with the gel solutions listed in Section A.1, the following buffer solutions were used (with numeric correspondence):

1. Spectrumedix Reveal Running Buffer, 5× concentration(Spectrumedix, BRUR-500-002).

2. Spectrumedix Reveal High Resolution Running Buffer w/o dye, 5× concentration (Spectrumedix, BRUR-500-004), pH=8.5 at 1×.

3. 1×Tris-Boric acid-EDTA buffer (TBE$^{\frac{1}{2}}$, pH=7.8–8.1).

4. PVP buffer.

5. MegaBACE buffer.

To prepare the Spectrumedix buffers, they were diluted in Millipore filtered H_2O (Millipore Corporation, Billerica, MA) to 1×. To make 1 L of 10× PVP buffer containing 300 mM Tris, 1M TAPS and 10mM EDTA add the following:

- Tris base (Tris (hydroxymethyl-)aminomethane) (121.1 g/M) 36.33 g.

- TAPS (243.3 g/M) 243.3 g.

- EDTA (372.2 g/M) 20 mL of a 0.5 M/L solution.

Add Millipore H_2O to make 1 L of buffer. Test the pH of the solution and add NaOH until pH reaches 8.0-8-5. Sterile filtrate or autoclave the solution before use. The chemicals used for the buffer were of analytical grade (Sigma-Aldrich).

A.3 Wash solutions

In between all capillary runs, except for those using the PVP matrix, the capillaries were rinsed with a wash solution. The wash was a surfactant aqueous solution, used at 1× concentration (Spectrumedix, Capillary Wash, WASH-500-002).

A.4 Capillary coating removal

Each capillary is coated with a ∼10 μm thick polyimide sheath. This sheath interferes with fluorescence detection because it is autofluoresent in the 480-520 nm band. For our experiments with fluorescence detection, the sheath was removed 3-5 mm from the tips of the capillaries using a variety of methods, including hydrogen flame, ethyl alcohol flame, laser ablation, mechanical scraping, and immersion in boiling sulfuric acid. Each of these will be briefly discussed with advantages and drawbacks, and amenability to parallelism.

Hydrogen flame: removed coating cleanly, but serial process and can melt fused silica capillary walls if extreme care is not used.

Ethyl alcohol flame: coating removal technique of choice for this phase of the research. Leaves a black residue (charred polyimide) that must be mechanically removed with a wet tissue, but is reliable and simple, though generally serial.

Laser ablation: performed by Polymicro Technologies (Polymicro Technologies, LLC, Phoenix, AZ). Very clean and recommended for UTMS, cost permitting.

Mechanical scraping: razor blade removal technique can work, but is a sensitive, manual, serial process. Not recommended.

Immersion in boiling sulfuric acid: Simple technique for coating removal from a capillary array *en masse*. Several support structures were built to facilitate dunking the capillaries in this bath. Capillary array flushing with H_20 is recommended afterwards. Array uniformity can be a challenge and the base of the removed sheath can be less clean than with the other techniques.

Appendix B

Sample preparation techniques for DNA and fluorescent solutions

For denaturing capillary electrophoresis, we require a 250 nL-5 μL DNA sample solution from which to load the capillaries. In this Appendix, we introduce the biotechniques used to prepare the various sample solutions used in this thesis.

B.1 Fluorescent solutions

As a simulant of the transient fluorescent emission expected from electrophoresis of charged, fluorescently-labeled DNA, a solution of bovine serum albumin (BSA) conjugated to AlexaFluor 488 (Synthegen, Houston, TX) with a ratio of 6.4 was prepared with a solvent of phosphate-buffered saline (PBS). The concentrations were varied from 200 nM/L to 0.002 nM/L in decrements of 10×. The solution was loaded into empty capillaries by means of capillary action. To prevent evaporation from affecting measurements, one end of the capillary was immersed in a virtually infinite supply of solution to replenish evaporated water from the capillaries.

In addition, a solution of Fluorescein was prepared at 2 μM/L concentration to test the spatial resolution and field-of-view of the fluorescence detection system.

B.2 Single-stranded DNA primer solution

Primers are generally used as probes for initiating and terminating the polymerase chain reaction (PCR) at specific, complementary locations in genomic DNA. In this work, we also utilized them as simple test material for electrophoresis experiments. The DNA segment used for electrophoresis was a 20 base primer 5′-fluorescently-labeled (5′GACAAACCATGCCACCAAGC) (Sigma Genosys, Woodland, TX) in various concentrations, typically 2 nM/L-200 nM/L.

B.3 Duplex DNA wildtype/mutant solutions

A variety of DNA samples were prepared for our experiments. In this section we present two representative recipes, one of which was less successful but educational, and a second from which all the mutation detection experiments in this thesis can be reproduced.

B.3.1 HaeIII hemoglobin sample preparation

Dr. Cathy Hogan provided several *Escherichia Coli* (E. Coli) bacteria samples containing extra-chromosomal, plasmid, recombinant DNA and her kind assistance was provided in the preparation of these samples. These circular plasmids contained a 272 bp HaeIII hemoglobin (hbglb) fragment (bp 348-620 genbank orientation and numbering).

The bacteria, initially frozen in glycerol, were streaked on media with ampicillin selection. Single colonies were then grown in Luria-Bertani broth, aliquotted, lysed, and sieved according to standard procedures. Using a QIAprep Spin Miniprep Kit (Qiagen, Valencia, CA), we purified the plasmid DNA.

The purified plasmid DNA was amplified using labeled forward and unlabeled reverse primers, as well as pairs of unlabeled primers with SYBR green intercalating dye. PCR reactions were performed in 20 μL capillary tubes (Roche Diagnostics, LightCycler Capillaries) on a Roche Lightcycler 2.0 (Roche, Indianapolis, IN).

PCR was performed by mixing 2 μL of purified plasmid DNA with 9.6 μL H_2O, 2.4 μL $MgCl_2$, 2 μL reaction mix, and 2 μL each of the forward and reverse primers, according to standard procedures (Roche Lightcycler FastStart DNA Master SYBR Green I), for a total volume of 20 μL per PCR reaction

PCR reactions were preincubated at 95 °C for 5 min followed by 55 cycles at 95 °C for 10 s, annealing at 49 °C for 10 s and extension at 72 °C for 10 s. To form heteroduplexes, the raw PCR products were heated to 95 °C for 1 min followed by a gradual lowering to room temperature. Products were injected raw or diluted in Millipore filtered water prior to electrokinetic injection for denaturing capillary electrophoresis.

Quality control measures included analysis of DNA fragment lengths using slab gel electrophoresis and DNA sequencing. Sequencing revealed unexpected noise in the sequence, due either to mishandling during preparation or pre-existing conditions. Denaturing capillary electrophoresis experiments were unsuccessful, not yielding clean separation of homoduplexes and heteroduplexes, due to this noise.

B.3.2 *BRAF* exon 15 sample preparation

The procedure for these samples has been described in Hinselwood *et al.* [179], but will be briefly restated here. From the cell line (THX) containing human melanoma, cells were collected. Using a Qiagen tissue extraction kit (Qiagen, Valencia, CA), DNA was extracted from the cells and aliquots were transferred into 96-well PCR plates. PCR reactions were performed in 96-well plates on a PTC DYAD DNA engine (MJ Research, Waltham, MA). Primers used to amplify the *BRAF* exon 15 are described in Chapter 8, along with labeling with 6-FAM.

PCR was performed by mixing 50 ng genomic DNA with 25 μM/L of each dNTP, 10× Taq buffer (750 mM/L Tris-HCl (pH 8.8 at 25 °C), 20 mM/L $(NH_4)_2SO_4$, 0.1%v/v Tween 20), 1 unit (U) of 10:1 Taq:PFU polymerase and 5 pM of each primer in a final volume of 20 μL. PCR reactions were preincubated at 96 °C for 5 min followed by 35 cycles at 95 °C for 30 s, annealing at 56 °C for 30 s and extension at 72 °C for 1 min. To form heteroduplexes, the raw PCR products were heated to 95

°C for 1 min followed by 65 °C for 30 min was then followed by a gradual lowering to room temperature. Products were injected raw or diluted in Millipore filtered water prior to electrokinetic injection for denaturing capillary electrophoresis. Gentle vortexing and centrifugation was performed prior to injection to ensure mixing and sample aggregation. Acquisition of all human materials strictly adhered to approved institutional guidelines.

Incubation with Pfu enzyme

The *BRAF* wt/mt samples were shipped across the Atlantic Ocean for testing. During transit, some external factors, most likely heat, caused the appearance of doublet peaks during CDCE analysis. Biologically, the PCR enzyme may have not fully extended the gene fragment during an unintended thermal cycle, leading to the creation of some shorter DNA duplexes. To remedy this problem, we incubated the raw PCR products with native Pfu polymerase enzyme (Stratagene, Pfu DNA Polymerase, 600135) to extend the DNA to blunt end products once again. The "before" and "after" runs under identical conditions, are shown in Figure B-1. The procedure was as follows: Take 10 μL of the PCR product; add 0.4 μL PFU enzyme (1 unit at 10× concentration), place the tube at 72 °C for 30 minutes.

The samples used for these experiments are *BRAF* exon 15 with mutation T→A labeled with 6-carboxyfluorescein (6-FAM). Raw PCR products were diluted 10× in Millipore H_2O. From a 5 μL aliquot, samples were electrokinetically injected onto the 320 mm long capillary column on Prof. Thilly's instrument for 15 s at 2 μA (30 μC), at ∼1 kV. The sample was electrophoresed at 8 μA, at ∼5 kV. After passing through a 130 mm denaturing zone heated by a water jacket to 69 °C, the sample passed by a window at a distance of 250 mm from the injection end, where fluorescence emission was measured at 10 Hz from a the PMT with gain settings of 0.65 kV and 10^7 V/A. The sieving matrix was in Spectrumedix Reveal High Resolution mutation detection matrix (see Appendix A). The experimental conditions and results of Figure B-1 are identical those described in Figure 8-9, except for a 1 °C temperature difference.

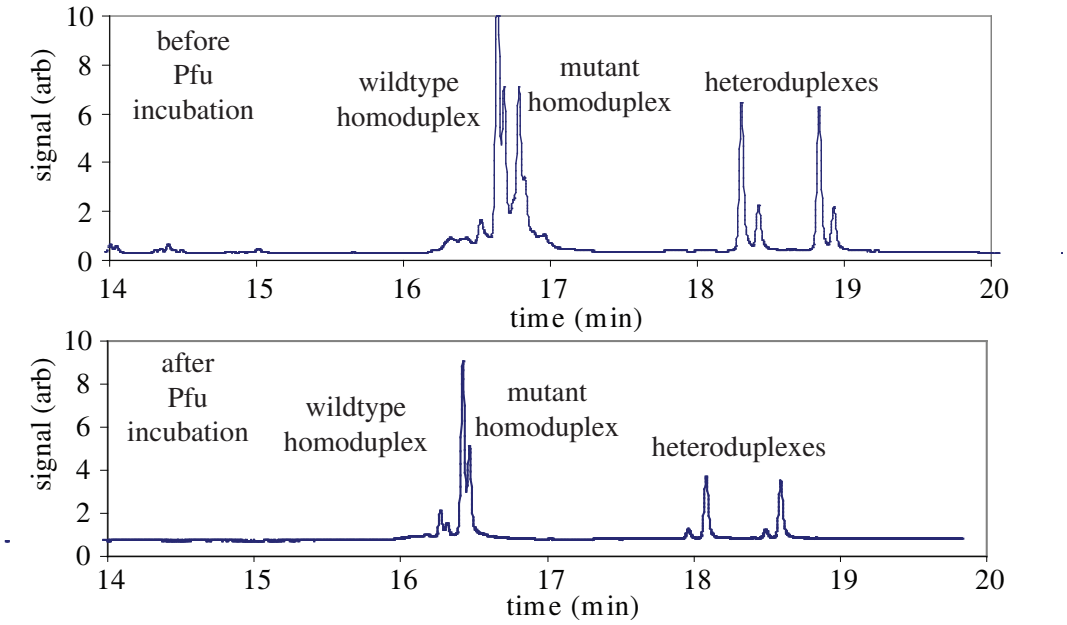

Figure B-1: Effects of Pfu incubation on denaturing capillary electrophoresis (using CDCE) separation of *BRAF* exon 15 mutant from wildtype. The Pfu enzyme extends all off the duplex DNA to blunt end fragments, eliminating "doublet" peaks caused by incomplete polymerization. Running conditions: Prof. Thilly's Instrument, Spectrumedix HR gel, 69 °C, 0.1× raw PCR products, 30 μC injection.

	absorption max (nm)	emission max (nm)	extinction coefficient (cm^{-1} M^{-1})	quantum yield
6-carboxyfluorescein (FAM)	495	520	83,000	0.90
AlexaFluor 488 (Alexa488)	495	519	71,000	0.94

	brightness	stokes shift (nm)
6-carboxyfluorescein (FAM)	74.7	25
AlexaFluor 488 (Alexa488)	66.7	24

Table B.1: Properties of fluorophores AlexaFluor 488 and 6-FAM.

PCR cleanup

Using a commercially available technique (Qiagen, QIAquick PCR Purification Kit), we were able to remove extraneous materials from the raw PCR products. In this procedure, single- or double-stranded DNA fragment ranging from 100 bp to 10 kb are purified from primers, nucleotides, polymerases, and salts. The procedure involves binding the DNA analytes to a column, washing away the others, and then eluting the analytes into a dilute buffer solution. In addition to purification, we also explored simple PCR product dilution in H_2O by factors ranging from 2:1 to 10:1 to reduce salt concentration.

B.4 Fluorophores

Two types of fluorophores were used in this work, 6-FAM and AlexaFluor 488. Some useful properties are listed in Table B.1 [http://home.earthlink.net/ pubspectra/]. In addition to these metrics, 6-FAM is much cheaper than AlexaFluor 488 488, but AlexaFluor 488 is less susceptible to photobleaching, more pH stable (pH range 4-10), and better resistant to conjugation quenching [129].

Bibliography

[1] C.J. Easley, J.M. Karlinsey, J.M. Bienvenue, L.A. Legendre, M.G. Roper, S.H. Feldman, M.A. Hughes, E.L. Hewlett, T.J. Merkel, J.P. Ferrance, and J.P. Landers. A fully integrated microfluidic genetic analysis system with sample-in-answer-out capability. *Proc. Natl. Acad. Sci. U.S.A.*, 103(51):19272–19277, 2006.

[2] The International SNP Map Working Group. A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature*, 409(6822):928–933, 2001.

[3] The International HapMap Project. The International HapMap Project. *Nature*, 426(6968):789–796, 2003.

[4] R. Sladek, G. Rocheleau, J. Rung, C. Dina, L. Shen, D. Serre, P. Boutin, D. Vincent, A. Belisle, S. Hadjadj, B. Balkau, B. Heude, G. Charpentier, T.J. Hudson, A. Montpetit, A.V. Pshezhetsky, M. Prentki, B.I. Posner, D.J. Balding, D. Meyre, C. Polychronakos, and P. Froguel. A genome-wide association study identifies novel risk loci for type 2 diabetes. *Nature*, 445(7130):881–885, 2007.

[5] A. Harmon. Facing life with a lethal gene. *The New York Times*, March 18 2007.

[6] *The Oxford English Dictionary*. 2nd edition, 1989.

[7] International Programs Center U.S. Census Bureau. World PopClock Projection. Washington, D.C. 20233-8800.

[8] N.J. Risch. Searching for genetic determinants in the new millenium. *Nature*, 405(6788):847–856, 2000.

[9] M. Kimura and J.F. Crow. The number of alleles that can be maintained in a finite population. *Genetics*, 49(4):725–738, 1964.

[10] E.S. Lander and D. Botstein. Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics*, 121(1):185–199, 1989.

[11] J.K. Pritchard. Are rare variants responsible for susceptibility to complex diseases? *Am. J. Hum. Genet.*, 69(1):124–137, 2001.

[12] L.L. Calvari-Sforza and W.F. Bodmer. *The Genetics of Human Populations*. San Francisco, 1971.

[13] D.L. Hartl. *A Primer of Population Genetics*. Sunderland, MA, 3rd edition, 2000.

[14] D.E. Reich and E.S. Lander. On the allelic spectrum of human disease. *Trends Genet.*, 17(9):502–510, 2001.

[15] G. Cabrera, M. Cot, F. Migot-Nabias, P.G. Kremsner, P. Deloron, and A.J. Luty. The sickle cell trait is associated with enhanced immunoglobulin G antibody responses to *plasmodium falciparum* variant surface antigens. *J. Infect. Dis.*, 191:1631–1638, 2005.

[16] G.B. Pier, M. Grout, T. Zaidi, G. Meluleni, S.S. Mueschenborn, G. Banting, R. Ratcliff, M.J. Evans, and W.H. Colledge. Salmonella typhi uses CTFR to enter intestinal epithelial cells. *Nature*, 393(6680):79–82, 1998.

[17] E.S. Lander. The new genomics: global views of biology. *Science*, 274(5287):539–539, 1996.

[18] A. Chakravarti. Population genetics-making sense out of sequence. *Nat. Genetics*, 21:56–60, 1999.

[19] S.L. Slager, J. Huang, and V.J. Vieland. Effect of allelic heterogeneity on the power of the transmission disequilbrium test. *Genet. Epidemiol.*, 18(2):143–156, 2000.

[20] K.M. Weiss and S.M. Fullerton. Phenogenetic drift and the evolution of genotype-phenotype relationships. *Theoretical Population Biology*, 57(3):187–195, 2000.

[21] K.M. Weiss and J.D. Terwilliger. How many diseases does it take to map a gene with SNPs? *Nat. Genetics*, 26:151–157, 2000.

[22] A. Collins, C. Lonjou, and N.E. Morton. Genetic epidemiology of single-nucleotide polymorphisms. *Proc. Natl. Acad. Sci. U.S.A.*, 96(26):15173–15177, 1999.

[23] N.J. Schork, L.R. Cardon, and X. Xu. The future of genetic epidemiology. *Trends Genet.*, 14(7):266–272, 1998.

[24] J. Ott. Predicting the range of linkage disequilbrium. *Proc. Natl. Acad. Sci. U.S.A.*, 97(1):2–3, 2000.

[25] S. Benzer and E. Freese. Induction of specific mutations with 5-Bromouracil. *Proc. Natl. Acad. Sci. U.S.A.*, 44:112–119, 1958.

[26] S. Morgenthaler and W.G. Thilly. A strategy to discover genes that carry multi-allelic or mono-allelic risk for common diseases: A cohort allelic sums test (CAST). *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, 615(1-2):28–56, 2007.

[27] P.J. Farabaugh, U. Schmeissner, M. Hofer, and J.H. Miller. Genetic studies of the LAC Repressor VII on the molecular nature of spontaneous hotspots in the LACI gene of *Escherichia Coli. J. Mol. Biol.*, 126(4):847–857, 1978.

[28] K. Khrapko, H.A. Coller, P.C. Andre, X.-C. Li, J.S. Hanekamp, and W.G. Thilly. Mitochondrial mutational spectra in human cells andtissues. *Proc. Natl. Acad. Sci. U.S.A.*, 94(25):13798–13803, 1997.

[29] A. Tomita-Mitchell, B.P. Muniappan, P. Herrero-Jimenez, H. Zarbl, and W.G. Thilly. Single nucleotide polymorphism spectra in newborns and centenarians: identification of genes coding for risk of mortal disease. *Gene*, 223(1-2):381–391, 1998.

[30] K. Sankaranarayanan and R. Chakraborty. Ionizing radiation and genetic risks: XIII. Summary and synthesis of papers VI to XII and estimates of genetic risks in the year 2000. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, 453(2):183–197, 2000.

[31] I. Cascorbi. Pharmacogenetics of cytochrome P4502D6: genetic background and clinical implication. *Eur. J. Clin. Invest*, 33(Supple. 2):17–22, 2003.

[32] R.C. Strange and A.A. Fryer. The glutathione s-transferases: influence of polymorphism on cancer susceptibiltity. *IARC Sci. Publ.*, 148:231–249, 1999.

[33] T. Imanishi, T. Itoh, Y. Suzuki, C. Donovan, S. Fukuchi, K.O. Koyanagi, R.A. Barrero, T. Tamura, Y. Yamaguchi-Kabata, M. Tanino, K. Yura, S. Miyazaki, K. Ikeo, K. Homma, A. Kasprzyk, T. Nishikawa, M. Hirakawa, J. Thierry-Mieg, D. Thierry-Mieg, J. Ashurst, L. Jia, M. Nakao, M.A. Thomas, N. Mulder, Y. Karavidopoulou, L. Jin, S. Kim, T. Yasuda, B. Lenhard, E. Eveno, Y. Suzuki, C. Yamasaki, J. Takeda, C. Gough, P. Hilton, Y. Fujii, H. Sakai, S. Tanaka, C. Amid, M. Bellgard, M. Bonaldo, H. Bono, S.K. Bromberg, A.J. Brookes, E. Bruford, P. Carninci, C. Chelala, C. Couillault, S. Souza, M. Debily, M. Devignes, I. Dubchak, T. Endo, A. Estreicher, E. Eyras, K. Fukami-Kobayashi, G.R. Gopinath, E. Graudens, Y. Hahn, M. Han, Z. Han, K. Hanada, H. Hanaoka, E. Harada, K. Hashimoto, U. Hinz, M. Hirai, T. Hishiki, I. Hopkinson, S. Imbeaud, H. Inoko, A. Kanapin, Y. Kaneko, T. Kasukawa, J. Kelso, P. Kersey, R. Kikuno, K. Kimura, B. Korn, V. Kuryshev, I. Makalowska, T. Makino, S. Mano, R. Mariage-Samson, J. Mashima, H. Matsuda, H. Mewes, S. Minoshima, K. Nagai, H. Nagasaki, N. Nagata, R. Nigam, O. Ogasawara, O. Ohara, M. Ohtsubo, N. Okada, T. Okido, S. Oota, M. Ota, T. Ota, *et al.*

Integrative annotation of 21,037 human genes validated by full-length cDNA clones. *PLoS Biology*, 2(6):e162, 2004.

[34] J.L. Rees. The genetics of sun sensitivity in humans. *Genet.*, 75(5):739–751, 2004.

[35] G.S. Hageman, D.H. Anderson, L.V. Johnson, L.S. Hancox, A.J. Taiber, L.I. Hardisty, J.L. Hageman, H.A. Stockman, J.D. Borchardt, K.M. Gehrs, R.J.H. Smith, G. Silvestri, S.R. Russell, C.C.W. Klaver, I. Barbazetto, S. Chang, L.A. Yannuzzi, G.R. Barile, J.C. Merriam, R.T. Smith, A.K. Olsh, J. Bergeron, Jana J. Zernant, J.E. Merriam, B. Gold, M. Dean, and R. Allikmets. A common haplotype in the complement regulatory gene factor H (HF1/CFH) predisposes individuals to age-related macular degeneration. *Proc. Natl. Acad. Sci. U.S.A.*, 102(20):7227–7232, 2005.

[36] R.E. Saunders, T.H. Goodship, P.F. Zipfel, and S.J. Perkins. An interactive web database of factor H associated hemolytic uremic syndrome mutations: insights into the structural consequences of disease associated mutations. *Hum. Mutat.*, 27(1):21–30, 2005.

[37] K.P. Magnusson, S. Duan, H. Sigurdsson, H. Petursson, Z. Yang, Y. Zhao, P.S. Bernstein, J. Ge, F. Jonasson, E. Stefansson, G. Helgadottir, N.A. Zabriskie, T. Jonsson, A. Bjornsson, T. Thorlacious, P.V. Jonsson, G. Thorleifsson, A. Kong, H. Stefansson, K. Zhang, K. Stefansson, and J.R. Gulcher. CFH Y402H confers similar risk of soft drunsen and both forms of advanced AMD. *PLoS Med. 3*, (e5), 2005.

[38] J.N. Hirschhorn and M.J. Daly. Genome-wide association studies for common diseases and complex traits. *Nature Reviews*, 6(2):95–105, 2005.

[39] G. Jimenez-Sanchez, B. Childs, and D. Valle. Human disease genes. *Nature*, 409(6822):853–855, 2001.

[40] J.-P. Hugot, M. Chamaillard, H. Zouali, S. Lesage, J.-P. Cezard, J. Belaiche, S. Almer, C. Tysk, C.A. O'Morain, M. Gassull, V. Binder, Y. Finkel, A. Cortot, R. Modigliani, P. Laurent-Puig, C. Gower-Rousseau, J. Macry, J.-F. Colombel, M. Sahbatou, and G. Thomas. Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. *Nature*, 411(6837):599–603, 2001.

[41] Y. Ogura, D.K. Bonen, N. Inohara, D.L. Nicolae, F.F. Chen, R. Ramos, H. Britton, T. Moran, R. Karaliuskas, R.H. Duerr, J.-P. Achkar, S.R. Brant, T.M. Bayless, B.S. Kirschner, S.B. Hanauer, G. Nunez, and J.H. Cho. A frameshift mutation in NOD2 associated with susceptibility to Crohn's disease. *Nature*, 411(6837):603–606, 2001.

[42] J.D. Rioux, M.J. Daly, M.S. Silverberg, K. Lindblad, H. Steinhart, Z. Cohen, T. Delmonte, K. Kocher, K. Miller, S. Guschwan, E.J. Kulbokas, S. O'Leary, E. Winchester, K. Dewar, T. Green, V. Stone, C. Chow, A. Cohen, D. Langelier, G. Lapointe, D. Gaudet, J. Faith, N. Branco, S.B. Bull, R.S. McLeod, A.M. Griffiths, A. Bitton, G.R. Greenberg, E.S. Lander., K.A. Siminovitch, and T.J. Hudson. Genetic variation in the 5Q31 cytokine gene cluster confers susceptibility to Crohn disease. *Nat. Genetics*, 29(2):223–228, 2001.

[43] M. Stoll, B. Corneliussen, C.M. Costello, G.H. Waetzig, B. Mellgard, W.A. Koch, P. Rosenstiel, M. Albrecht, P.J.P. Croucher, D. Seegert, S. Nikolaus, J. Hampe, T. Lengauer, S. Pierrou, U.R. Foelsch, C.G. Mathew, M. Lagerstrom-Fermer, and S. Schreiber. Genetic variation in DLG5 is associated with inflammatory bowel disease. *Nat. Genetics*, 36(5):476–480, 2004.

[44] H. Stefansson, E. Sigurdsson, V. Steinthorsdottir, S. Bjornsdottir, T. Sigmundsson, S. Ghosh, J. Brynjolfsson, S. Gunnarsdottir, O. Ivarsson, T.T. Chou, O. Hjaltason, B. Birgisdottir, H. Jonsson, V.G. Gudnadottir, E. Gudmundsdottir, A. Bjornsson, B. Ingvarsson, A. Ingason, S. Sigfusson, H. Hardardottir, R.P. Harvey, D. Lai, M. Zhou, D. Brunner, V. Mutel, A. Gonzalo, G. Lemke, J. Sainz, G. Johannesson, T. Andresson, D. Gudbjartsson, A. Manolescu, M.L.

Frigge, M.E. Gurney, A. Kong, J.R. Gulcher, H. Petursson, and K. Stefansson. Neuregulin 1 and susceptibility to schizophrenia. *Am. J. Hum. Genet.*, 71(4):877–892, 2002.

[45] L. Nistico, R. Buzzetti, L.E. Pritchard, B. Van der Auwera, C. Giovannini, E. Bosi, M.T. Larrad, M.S. Rios, C.C. Chow, C.S. Cockram, K. Jacobs, C. Mijovic, S.C. Bain, A.H. Barnett, C.L. Vandewalle, F. Schuit, F.K. Gorus, R. Tosi, P. Pozzilli, and J.A. Todd. The CTLA-4 gene region of chromosome 2Q33 is linked to, and associated with, type 1 diabetes. Belgian diabetes registry. *Hum. Mol. Genet.*, 5(7):1075–1080, 1996.

[46] M.J. Daly and J.D. Rioux. New approaches to gene hunting in IBD. *Inflammatory Bowel Diseases*, 10(3):312–317, 2004.

[47] N. Risch and K. Merikangas. The future of genetic studies of complex human diseases. *Science*, 273(5281):1516–1517, 1996.

[48] M. Cargill, D. Altshuler, J. Ireland, P. Sklar, K. Ardlie, N. Patil, C.R. Lane, E.P. Lim, N. Kalyanaraman, J. Nemesh, L. Ziaugra, L. Friedland, A. Rolfe, J. Warrington, R. Lipshutz, G.Q. Daley, and E.S. Lander. Characterization of single-nucleotide polymorphisms in coding regions of human genes. *Nat. Genetics*, 22(3):231–238, 1999.

[49] M.K. Halushka, J.-B. Fan, K. Bentley, L. Hsie, N. Shen, A. Weder, R. Cooper, R. Lipshutz, and A. Chakravarti. Patterns of single-nucleotide polymorphisms in candidate genes for blood-pressure homeostasis. *Nat. Genetics*, 22(3):239–247, 1999.

[50] D. Botstein, R.L. White, M. Skolnik, and R.W. Davis. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am. J. Hum. Genet.*, 32(3):314–331, 1980.

[51] J.L. Weber and P.E. May. Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reation. *Am. J. Hum. Genet.*, 44(3):388–396, 1989.

[52] M. Litt and J.A. Luty. A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. *Am. J. Hum. Genet.*, 44(3):397–401, 1989.

[53] C.E. Lowe, J.D. Cooper, J.M. Chapman, J. Barratt, J. Twells, E.A. Green, D.A. Savage, C. Guja, C. Ionescu-Tîrgovite, E. Tuomilehto-Wolf, J. Tuomilehto, J.A. Todd, and D.G. Clayton. Cost-effective analysis of candidate genes using htSNPs: a staged approach. *Genes and Immunity*, 5(4):301–305, 2004.

[54] M. Nothnagel and Klaus Rohde. The effect of single-nucleotide polymorphism marker selection on patterns of haplotype blocks and haplotype frequency estimates. *Am. J. Hum. Genet.*, 77(6):988–998, 2005.

[55] D.M. Evans and L.R. Cardon. Guidelines for genotyping in genomewide linkage studies: single-nucleotide-polymorphism maps versus microsatellite maps. *Am. J. Hum. Genet.*, 75(4):687–692, 2004.

[56] The International Multiple Sclerosis Genetics Consortium. Enhancing linkage analysis of complex disorders: an evaluation of high-density genotyping. *Hum. Mol. Genet.*, 13(17):1943–1949, 2004.

[57] S. John, N. Shephard, G. Liu, E. Zeggini, M. Cao, W. Chen, N. Vasavda, T. Mills, A. Barton, A. Hinks, S. Eyre, K.W. Jones, W. Ollier, A. Silman, N. Gibson, J. Worthington, and G.C. Kennedy. Whole-genome scan, in a complex disease, using 11,245 single-nucleotide polymorphisms: Comparison with microsatellites. *Am. J. Hum. Genet.*, 75(1):54–64, 2004.

[58] F.A. Middleton, M.T. Pato, K.L. Gentile, C.P. Morley, X. Zhao, A.F. Eisener, A. Brown, T.L. Petryshen, A.N. Kirby, H. Medeiros, C. Carvalho, A. Macedo, A. Dourado, I. Coelho, J. Valente, M.J. Soares, C.P. Ferreira,

M. Lei, M.H. Azevedo, J.L. Kennedy, M.J. Daly, P. Sklar, and C.N. Pato. Genomewide linkage analysis of bipolar disorder by use of a high-density single-nucleotidepolymorphism (SNP) genotyping assay: A comparison with microsatellite marker assays and finding of significant linkage to chromosome 6Q22. *Am. J. Hum. Genet.*, 74(5):886–897, 2004.

[59] D. Levy, A.L. DeStefano, M.G. Larson, C.J. O'Donnell, R.P. Lifton, H. Gavras, L.A. Cupples, and R.H. Myers. Evidence for a gene influencing blood pressure on chromosome 17: Genome scan linkage results for longitudinal blood pressure phenotypes in subjects from the Framingham heart study. *Hypertension*, 36(4):477–483, 2000.

[60] J.D. Terwilliger and K.M. Weiss. Linkage disequilibrium mapping of complex diseases: Fantasy or reality? *Curr. Opin. Biotech.*, 9(6):578–594, 1998.

[61] N.L. Anderson and N.G. Anderson. Proteome and proteomics: New technologies, new concepts, and new words. *Electrophoresis*, 19(11):1853–1861, 1998.

[62] R. Brent. Functional genomics: Learning to think about gene expression data. *Current Biology*, 9(9):R338–R341, 1999.

[63] R. Brent. Genomic biology. *Cell*, 100(1):169–183, 2000.

[64] R.P. Woychik, M.L. Klebig, M.J. Justice., T.R. Magnuson, and E.D. Avrer. Functional genomics in the post-genome era. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, 400(1-2):3–14, 1998.

[65] R. Waterland and C. Garza. Potential mechanisms of metabolic imprinting that lead to chronic disease. *Am. J. Clin. Nutr.*, 69(2):179–197, 1999.

[66] J.N. Hirschhorn and D. Altshuler. Once and again–issues surrounding replication in genetic association studies. *J. Clin. Endocrinol. Metab.*, 87(10):4438–4441, 2002.

[67] M. Faham, J. Zheng, M. Moorhead, H. Fakhrai-Rad, E. Namsaraev, K. Wong, Z. Wang, S.G. Chow, L. Lee, K. Suyenaga, J. Reichert, A. Boudreau, J. Eberle, C. Bruckner, M. Jain, G. Karlin-Neumann, H.B. Jones, T.D. Willis, J.D. Buxbaum, and R.W. Davis. Multiplexed variation scanning for 1,000 amplicons in hundreds of patients using mismatch repair detection (MRD) on tag arrays. *Proc. Natl. Acad. Sci. U.S.A.*, 102(41):14717–14722, 2005.

[68] G. A. Watterson. On the number of segregating sites in genetical models without recombination. *Theoretical Population Biology*, 7(2):256–276, 1975.

[69] L. Kruglyak and D.A. Nickerson. Variation is the spice of life. *Nat. Genetics*, 27(3):234–236, 2001.

[70] J.C. Venter, M.D. Adams, E.W. Myers, P.W. Li, R.J. Mural, G.G. Sutton, H.O. Smith, M. Yandell, C.A. Evans, R.A. Holt, J.D. Gocayne, P. Amanatides, R.M. Ballew, D.H. Huson, J.R. Wortman, Q. Zhang, C.D. Kodira, X.H. Zheng, L. Chen, M. Skupski, G. Subramanian, P.D. Thomas, J. Zhang, G.L. Gabor Miklos, C. Nelson, S. Broder, A.G. Clark, J. Nadeau, V.A. McKusick, N. Zinder, A.J. Levine, R.J. Roberts, M. Simon, C. Slayman, M. Hunkapiller, R. Bolanos, A. Delcher, I. Dew, D. Fasulo, M. Flanigan, L. Florea, A. Halpern, S. Hannenhalli, S. Kravitz, S. Levy, C. Mobarry, K. Reinert, K. Remington, J. Abu-Threideh, E. Beasley, K. Biddick, V. Bonazzi, R. Brandon, M. Cargill, I. Chandramouliswaran, R. Charlab, K. Chaturvedi, Z. Deng, V.D. Francesco, P. Dunn, K. Eilbeck, C. Evangelista, A.E. Gabrielian, W. Gan, W. Ge, F. Gong, Z. Gu, P. Guan, T.J. Heiman, M. Higgins, R.-R. Ji, Z. Ke, K.A. Ketchum, Z. Lai, Y. Lei, Z. Li, J. Li, Y. Liang, X. Lin, F. Lu, G.V. Merkulov, N. Milshina, H.M. Moore, A.K. Naik, V.A. Narayan, B. Neelam, D. Nusskern, D.B. Rusch, S. Salzberg, W. Shao, B. Shue, J. Sun, Z.Y. Wang, A. Wang, X. Wang, J. Wang, M.-H. Wei, R. Wides, C. Xiao, C. Yan, *et al.* The sequence of the human genome. *Science*, 291(5507):1304–1351, 2001.

[71] N. Patil, A.J. Berno, D.A. Hinds, W.A. Barrett, J.M. Doshi, C.R. Hacker, C.R. Kautzer, D.H. Lee, C. Marjoribanks, D.P. McDonough, B.T.N. Nguyen, M.C. Norris, J.B. Sheehan, N. Shen, D. Stern, R.P. Stokowski, D.J. Thomas, M.O. Trulson., K.R. Vyas, K.A. Frazer, S.P.A. Fodor, and D.R. Cox. Blocks of limited haplotype diversity revealed by high-resolution scanning of human chromosome 21. *Science*, 294(5547):1719–1723, 2001.

[72] L.D. Hurst and N.G.C. Smith. Do essential genes evolve slowly? *Current Biology*, 9(14):747–750, 1999.

[73] J.N. Hirschhorn, K. Lohmueller, E. Byrne, and K. Hirschhorn. A comprehensive review of genetic association studies. *Genet. Med.*, 4(2):45–61, 2002.

[74] E.J.C.G. van den Oord and P.F. Sullivan. False discoveries and models for gene discovery. *Trends Genet.*, 19(10):537–542, 2003.

[75] R.G.H. Cotton. *Mutation Detection.* Oxford University Press, Oxford, 1997.

[76] J. Shendure, G.J. Porreca, N.B. Reppas, X. Lin, J.P. McCutcheon, A.M. Rosenbaum, M.D. Wang, K. Zhang, R.D. Mitra, and G.M. Church. Accurate multiplex polony sequencing of an evolved bacterial genome. *Science*, 309(5741):1728–1732, 2005.

[77] F. Sanger, G.M. Air, B.G. Barrell, N.L. Brown, A.R. Coulson, J.C. Fiddes, C.A. Hutchison, P.M. Slocombe, and M. Smith. Nucleotide sequence of bacteriophage ϕ X174 DNA. 265(5596):687–695, 1977.

[78] E.M. Southern, U. Maskos, and J.K. Elder. Analyzing and comparing nucleic acid sequences by hybridization to arrays of oligonucleotides: Evaluation using experimental models. *Genomics*, 13(4):1008–1017, 1992.

[79] S.G. Fischer and L.S. Lerman. DNA fragments differing by single base-pair substitutions are separated in denaturing gradient gels: Correspondence with melting theory. *Proc. Natl. Acad. Sci. U.S.A.*, 80(6):1579–1583, 1983.

[80] L.S. Lerman and K. Silverstein. Computational simulation of DNA melting and its application to denaturing gradient gel electrophoresis. *Methods Enzymol.*, 155:482–501, 1987.

[81] D. Poland. Recursion relation generation of probability profiles for specific-sequence macromolecules with long-range correlations. *Biopolymers*, 13(9):1859–1871, 1974.

[82] D. Poland and H.A. Scheraga. *Theory of Helix-Coil Transitions in Biopolymers*. Academic Press, 1970.

[83] M. Fixman and J.J. Freire. Theory of DNA melting curves. *Biopolymers*, 16(12):2693–2704, 1977.

[84] K. Khrapko, J.S. Hanekamp, W.G. Thilly, A. Belenkii, F. Foret, and B.L. Karger. Constant denaturant capillary electrophoresis (CDCE): a high resolution approach to mutational analysis. *Nucleic Acids Res.*, 22(3):364–369, 1994.

[85] M. Minarik, L. Minarikova, J. Bjørheim, and P.O. Ekstrøm. Cycling gradient capillary electrophoresis: A low-cost tool for high-throughput analysis of genetic variations. *Electrophoresis*, 24(11):1716–1722, 2003.

[86] J. Bjørheim, M. Minarik, G. Gaudernack, and P.O. Ekstrøm. Evaluation of denaturing conditions in analysis of DNA variants applied to multi-capillary electrophoresis instruments. *Journal of Separation Science*, 26(12-13):1163–1168, 2003.

[87] B. Crane, C. Hogan, L. Lerman, and I.W. Hunter. DNA mutation detection via fluorescence imaging in a spatial thermal gradient, capillary electrophoresis system. *Review of Scientific Instruments*, 72(11):4245–4251, 2001.

[88] X.-C. Li-Sucholeiki, G. Hu, T. Perls, A. Tomita-Mitchell, and W.G. Thilly. Scanning the β-globin gene for mutations in large populations by denaturing capillary and gel electrophoresis. *Electrophoresis*, 26(13):2531–2538, 2005.

[89] J.H. Aborn, S.A. El-Difrawy, M.Novotny, E.A. Gismondi, R. Lam, P. Matsudaira, B.K. Mckenna, T. O'Neil, P. Streechon, and D.J. Ehrlich. A 768-lane microfabricated system for high-throughput DNA sequencing. *Lab on a Chip*, 5(6):669–674, 2005.

[90] C.A. Emrich, H. Tian, I.L. Medintz, and R.A. Mathies. Microfabricated 384-lane capillary array electrophoresis bioanalyzer for ultrahigh-throughput genetic analysis. *Anal. Chem.*, 74(19):5076–5083, 2002.

[91] L. Koutny, D. Schmalzing, O Salas-Solano, S. El-Difrawy, A. Adourian, S. Buonocore, K. Abbey, P. McEwan, P. Matsudaira, and D. Ehrlich. Eight hundred-base sequencing in a microfabricated electrophoretic device. *Anal. Chem.*, 72(14):3388–3391, 2000.

[92] B.L. Karger and R.J. Nelson. Integrated temperature control/alignment system for high performance capillary electrophoretic apparatus. U.S. Patent No. 5,085,757.

[93] B.L. Karger, A. Paulus, A.S. Cohen, and R.J. Nelson. Integrated temperature control/alignment system for high performance capillary electrophoretic apparatus. U.S. Patent No. 4,898,658.

[94] X.-C. Li-Sucholeiki, A. Tomita-Mitchell, K. Arnold, B.J. Glassner, T. Thompson, J.V. Murthy, L. Berk, C. Lange, P.-M. Leong-Morgenthaler, D. MacDougall, J. Munro, D. Cannon, T. Mistry, A. Miller, C. Deka, B.L. Karger, K.M. Gillespie, P.O. Ekstrøm, J.A. Todd, and W.G. Thilly. Detection and frequency estimation of rare variants in pools of genomic DNA from large populations using mutational spectrometry. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, 570(2):267–280, 2005.

[95] A. Zhukauskas and R. Ulinsakas. Efficiency parameters for heat transfer in tube banks. *Heat Transfer Engineering*, 62(2):19–25, 1985.

[96] B.R. Munson, D.F. Young, and T.H. Okiishi. *Fundamentals of Fluid Mechanics*. John Wiley & Sons, Inc., New York, NY, 3rd edition, 1998.

[97] B.A. Proctor, I. Whitney, and J.W.Johnson. The strength of fused silica. *JSTOR*, 297(1451):534–557, 1967.

[98] C.H.K. Williamson. Vortex dynamics in the cylinder wake. *Annual Review of Fluid Mechanics*, 28:477–539, 1996.

[99] L.S.G. Kovasznay. Hot-wire investigation of the wake behind cylinders at low Reynolds numbers. *Proceedings of the Royal Society of London, Series A, Mathematical Physical Sciences*, 198(1053):174–190, 1949.

[100] U. Fey, M. Konig, and H. Eckelmann. A new Strouhal-Reynolds-number relationship for the circular cylinder in the range $47<\mathrm{Re}<2\mathrm{x}10^5$. *Physics Fluids*, 10(7):1547–1549, 1998.

[101] C.H.K. Williamson and G.L. Brown. A series in $1/\sqrt{\mathrm{Re}}$ to represent the Strouhal-Reynolds number relationship of the cylinder wake. *Journal of Fluids and Structures*, 12(8):1073–1085, 1998.

[102] C.F. Beards. *Structural Vibration: Analysis and Damping*. Arnold, London, UK, 1996.

[103] D.N. Heiger, A.S. Cohen, and B.L. Karger. Separation of DNA restriction fragments by high performance capillary electrophoresis with low and zero crosslinked polyacrylamide using continuous and pulsed electric fields. *J. Chromatogr.*, 516(1):33–48, 1990.

[104] W. Goetzinger, L. Kotler, E. Carrilho, M.C. Ruiz-Martinez, O. Salas-Solano, and B.L. Karger. Characterization of high molecular mass linear polyacrylamide powder prepared by emulsion polymerization as a replaceable polymer matrix for DNA sequencing by capillary electrophoresis. *Electrophoresis*, 19(2):242–248, 1998.

[105] S. Hjerten. High-performance electrophoresis: Elimination of electroendosmosis and solute adsorption. *J. Chromatogr. A*, 347:191–198, 1985.

[106] D. Schmalzing, C.A. Piggee, F. Foret, E. Carrilho, and B.L. Karger. Characterization and performance of a neutral hydrophilic coating for the capillary electrophoretic separation of biopolymers. *J. Chromatogr. A*, 652(1):149–159, 1993.

[107] F. Foret, M. Deml, and P. Boček. Quantitative study of the effects of some dispersive processes on the separation efficiency. *J. Chromatogr.*, 452:601–613, 1988.

[108] Q. Gao and E.S. Yeung. A matrix for DNA separation: Genotyping and sequencing using poly(vinylpyrrolidone) solution in uncoated capillaries. *Anal. Chem.*, 70(7):1382–1388, 1998.

[109] R.S. Mandabhushi. Separation of 4-color DNA sequencing extension products in noncovalently coated capillaries using low viscosity polymer solutions. *Electrophoresis*, 19(2):224–230, 1998.

[110] J. Horvath and V. Dolník. Polymer wall coatings for capillary electrophoresis. *Electrophoresis*, 22(4):644–655, 2001.

[111] P.G. Righetti, C. Gelfi, B. Verzola, and L. Castelletti. The state of the art of dynamic coatings. *Electrophoresis*, 22(4):603–611, 2001.

[112] J.R. Scherer, B.M. Paegel, G.J. Wedemayer, C.A. Emrich, J. Lo, I.L. Medintz, and R.A. Mathies. High-pressure gel loader for capillary array electrophoresis microchannel plates. *Biotechniques*, 31(5):1150–1152, 2001.

[113] B.F. Abu-Sharkh, G.O. Yahaya, S.A. Ali, E.Z. Hamad, and I.M. Abu-Reesh. Viscosity behavior and surface and interfacial activities of hydrophobically modified water-soluble acrylamide/N-phenyl acrylamide block copolymers. *J. Applied Polymer Science*, 89(8):2290–2300, 2003.

[114] B. Briscoe, P. Luckham, and S. Zhu. Pressure influences upon shear thickening of poly(acrylamide) solutions. *Rheol. Acta.*, 38(3):224–234, 1999.

[115] T.S. Rushing and R.D. Hester. Low-shear-rate capillary viscometer for polymer solution intrinsic viscosity determination at varying temperatures. *Rev. Sci. Inst.*, 74(1):176–181, 2003.

[116] A.M. Mood, F.A. Graybill, and D.C. Boes. *Introduction to the Theory of Statistics*. McGraw-Hill, 3rd edition, 1974.

[117] M. Minarik, K. Klepárník, M. Gilár, F. Foret, A. Miller, Z. Sosic, and B.L. Karger. Design of a fraction collector for capillary array electrophoresis. *Electrophoresis*, 23(1):23–42, 2003.

[118] T. Irie, T. Oshida, H. Hasegawa, Y. Matsuoka, T. Li, Y. Oya, T. Tanaka, G. Tsujimoto, and H. Kambara. Automated DNA fragment collection by capillary array gel electrophoresis in search of differentially expressed genes. *Electrophoresis*, 21(2):367–374, 2000.

[119] O. Múller, F. Foret, and B.L. Karger. Design of a high-precision fraction collector for capillary electrophoresis. *Anal. Chem.*, 67(17):2974–2980, 1995.

[120] T.S. Kanigan, C.J. Brenan, S. Lafontaine, L. Sosnowski, P. Madden, and I.W. Hunter. Living chips for drug discovery. In *Proceedings of the SPIE*, volume 3926, pages 172–180, 2000.

[121] K. Khrapko, H. Coller, P. Andre, X.-C. Li, F. Foret, A. Belenky, B. L. Karger, and W. G. Thilly. Mutational spectrometry without phenotypic selection: human mitochondrial DNA. *Nucleic Acids Res.*, 25(4):685–693, 1997.

[122] M.S. Bello. Electrolytic modification of a buffer during a capillary electrophoresis run. *J. Chromatogr. A*, 744(1):81–91, 1996.

[123] M. Macka, P. Andersson, and P.R. Haddad. Changes in electrolyte pH due to electrolysis during capillary zone electrophoresis. *Anal. Chem.*, 70(4):743–749, 1998.

[124] H. Corstjens, H.A.H. Billiet, J. Frank, and K.C.A.M. Luyben. Variation of the pH of the background electrolyte due to electrode reaction in capillary electrophoresis: theoretical approach and *in situ* measurement. *Electrophoresis*, 17(1):137–143, 1996.

[125] I. Rodriguez and N. Chandrasekhar. Experimental study and numerical estimation of the current changes in electroosmotically pumped microfluidic devices. *Electrophoresis*, 26(6):1114–1121, 2005.

[126] M.A. Kelly, K.D. Altria, and B.J. Clark. Approaches used in the reduction of buffer electrolysis effects for routine capillary electrophoresis procedures in pharmaceutical analysis. *J. Chromatogr. A*, 768(1):73–80, 1997.

[127] P. Jandik and G. Bohn. *Capillary Electrophoresis of Small Molecules and Ions*. VCH Publishers, New York, NY, 1993.

[128] A. Timperman, S.E. Tracht, and J.V. Sweedler. Dynamic on-column pH monitoring in capillary electrophoresis: Application to volume-limited outlet vials. *Anal. Chem.*, 68(15):2693–2698, 1996.

[129] R.P. Haughland. *Handbook of Fluorescencent Probes and Research Products*. Eugene, OR, 9th edition, 2002.

[130] T.M. Ridgeway, D.B. Hayes, T.P. Moody, T.J. Wilson, A.L. Anderson, J.H. Lavasseur, P.D. Demaine, B.E. Kenty, and T.M. Laue. An apparatus for membrane-confined analytical electrophoresis. *Electrophoresis*, 19(10):1611–1619, 1998.

[131] B.L. Gray, S.D. Collins, and R.L. Smith. Interlocking mechanical and fluidic interconnections for microfluidic circuit boards. *Sensors and Actuators A*, 12:18–24, 2004.

[132] C. Gartner, H. Becker, B. Anton, A. O'Neil, and O. Rotting. Polymer based microfluidic devices - examples for fluidic interfaces and standardization concepts. In *Proceedings of the SPIE*, volume 4982, pages 99–104, 2003.

[133] V. Nittis, R. Fortt, C.H. Legge, and A.J. Mello. A high-pressure interconnect for chemical microsystem applications. *Lab on a Chip*, 1(2):148–152, 2001.

[134] Z. Yang and R. Maeda. A world-to-chip socket for microfluidic prototype development. *Electrophoresis*, 23(20):3474–3478, 2002.

[135] A.H. Slocum. *Precision Machine Design*. Society of Manufacturing Engineers, Prentice-Hall, Dearborn, MI, 1992.

[136] D. Schmalzing, L. Koutney, A. Adourian, P. Belgrader, P.T. Matsudaira, and D.J. Ehrlich. DNA typing in thirty seconds with a microfabricated device. *Proc. Natl. Acad. Sci. U.S.A.*, 94(19):10273–10278, 1997.

[137] A.M. Skelly, J.R. Scherer, A.D. Aubrey, W.H. Grover, R.H.C. Ivester, P. Ehrenfreund, F.J. Grunthaner, J.L. Bada, and R.A. Mathies. Development and evaluation of a microdevice for amino acid biomarker detection and analysis on Mars. *Proc. Natl. Acad. Sci. U.S.A.*, 102(4):1041–1046, 2005.

[138] A. Tsupryk, M. Gorbovitski, E.A. Kabotyanski, and V. Gorfinkel. Novel design of multicapillary arrays for high-throughput DNA sequencing. *Electrophoresis*, 27(14):2869–2879, 2006.

[139] J. Zhang, M. Yang, X. Puyang, Y. Fang, L.M. Cook, and N.J. Dovichi. Two-dimensional direct-reading fluorescence spectrograph for DNA sequencing by capillary array electrophoresis. *Anal. Chem.*, 73(6):1234–1239, 2001.

[140] J.E. Shigley and C.R. Mischke. *Mechanical Engineering Design*. McGraw-Hill, New York, NY, 1989.

[141] R.A. Norwood, B. Brown, J. Holman, and L.W. Shacklette. Polymer gripping elements for optical fiber splicing. U.S. Patent No. 6,266,472.

[142] F.W. Taylor. *The Principles of Scientific Management*. 1st World Library, 1911.

[143] R.J. Zagursky and R.M. McCormick. DNA sequencing separations in capillary gels on a modified commercial DNA sequencing instrument. *Biotechniques*, 9(1):74–79, 1990.

[144] R.A. Mathies and X.C. Huang. Capillary array electrophoresis - an approach to high-speed, high-throughput DNA sequencing. *Nature*, 359(6391):167–169, 1992.

[145] J. Taylor and E.A. Yeung. Multiplexed fluorescence detector for capillary electrophoresis using axial optical fiber illumination. *Anal. Chem.*, 65(7):956–960, 1993.

[146] K. Ueno and E.S. Yeung. Simultaneous monitoring of DNA fragments separated by electrophoresis in a multiplexed array of 100 capillaries. *Anal. Chem.*, 66(9):1424–1431, 1994.

[147] M.A. Quesada and S. Zhang. Multiple capillary DNA sequencer that uses fiber-optic illumination and detection. *Electrophoresis*, 17(12):1841–1851, 1996.

[148] R. Siebert, S. Bottani, H. Rebscher, and L. Valentin. Multiple capillary electrophoresis systems. U.S. Patent No. 6,613,212.

[149] T. Anazawa, S. Takahashi, and H. Kambara. A capillary array gel electrophoresis system using multiple laser focusing for DNA sequencing. *Anal. Chem.*, 68(15):2699–2704, 1996.

[150] X. Lu and E.S. Yeung. Optimization of excitation and detection geometry for multiplexed capillary array electrophoresis of DNA fragments. *Appl. Spectrosc.*, 49(5):695–709, 1995.

[151] H. Kambara and S. Takahashi. Multiple-sheathflow capillary array DNA analyser. *Nature*, 361(6412):565–566, 1993.

[152] S. Takahashi, K. Murakami, T. Anazawa, and H. Kambara. Multiple sheathflow gel capillary-array electrophoresis for multicolor fluorescent DNA detection. *Anal. Chem.*, 66(7):1021–1026, 1994.

[153] J. Zhang, K.O. Voss, D.F. Shaw, K.P. Roos, D.F. Lewis, J. Yan, R. Jiang, H. Ren, J.Y. Hou, Y. Fang, X. Puyang, H. Ahmadzadeh, and N.J. Dovichi. A multiple-capillary electrophoresis system for small-scale DNA sequencing and analysis. *Nucl. Acids Res.*, 27(24):e36, i–vii, 1999.

[154] H.J. Crabtree, S.J. Bay, D.F. Lewis, L.D. Coulson, G. Fitzpatrick, D.J. Harrison, S.L. Delinger, J.Z. Zhang, and N.J. Dovichi. Construction and evaluation of a capillary array DNA sequencer based on a micromachined sheath-flow cuvette. *Electrophoresis*, 21(7):1329–1335, 2000.

[155] S. Bottani, H. Rebschera, R. Sieberta, and L. Valentina. Post-capillary laser-induced fluorescence detection without sheath-flow. *J. Chromatogr. A*, 844(1-2):433–438, 1999.

[156] J.W. Goodman. *Statistical optics*. Wiley, New York, NY, 1985.

[157] E. Hecht. *Optics*. Addison-Wesley, Reading, MA, 1998.

[158] J.W. Goodman. *Introduction to Fourier Optics*. McGraw Hill, Boston, MA, 1996.

[159] M. Born and E. Wolf. *Principles of Optics*. Pergamon Press, Inc., Elmsford, NY, 6th edition, 1980.

[160] Z.D. Popovic, R.A. Sprague, and N. Connell. Technique for monolithic fabrication of microlens arrays. *Appl. Opt.*, 27(7):1281–1284, 1988.

[161] S.-K. Lee, K.-C. Lee, and S.S. Lee. A simple method for microlens fabrication by the modified LIGA process. *J. Micromech. Microeng.*, 12(3):334–340, 2002.

[162] T.R. Jay and M.B. Stern. Preshaping photoresist for refractive microlens fabrication. *Opt. Eng.*, 33(11):3552–3555, 1994.

[163] M. Fritze, M. B. Stern, and P. W. Wyatt. Laser-fabricated glass microlens arrays. *Opt. Lett.*, 23(2):141–143, 1998.

[164] N. F. Borelli and D. K. Morse. Microlens arrays produced by a photolytic technique. *Appl. Opt.*, 27(3):476–479, 1988.

[165] M. Oikawa, K. Iga, S. Misawa, and Y. Kokubun. Improved distributed-index planar microlens and its application to 2-D lightwave components. *Appl. Opt.*, 22(3):441–442, 1983.

[166] D.L. MacFarlane, V. Narayan, J.A. Tatum, W.R. Cox, T. Chen, and D.J. Hayes. Microjet fabrication of microlens arrays. *Photonics Technology Letters, IEEE*, 6(9):1112–1114, 1994.

[167] H. Yang, C.-K. Chao, M.-K. Wei, and C.-P. Lin. High fill-factor microlens array mold insert fabrication using a thermal reflow process. *J. Micromech. Microeng.*, 14(8):1197–1204, 2004.

[168] S. Kalpakjian and S. Schmid. *Manufacturing Engineering and Technology*. Prentice Hall, Upper Saddle River, NJ, 4th edition, 2001.

[169] X.-C. Li-Sucholeiki, K. Khrapko, P.C. Andre, L.A. Marcelino, B.L. Karger, and W.G. Thilly. Applications of constant denaturant capillary electrophoresis/high-fidelity polymerase chain reaction to human genetic analysis. *Electrophoresis*, 20(6):1224–1232, 1999.

[170] X.-C. Li and W.G. Thilly. Use of wide-bore capillaries in constant denaturant capillary electrophoresis. *Electrophoresis*, 17(12):1884–1889, 1996.

[171] K. Klepárník, S. Fanali, and P. Boček. Selectivity of the separation of DNA fragments by capillary zone electrophoresis in low-melting-point agarose sol. *J. Chromatogr. A*, 638(2):283–292, 1993.

[172] P. Keohavong and W.G. Thilly. Fidelity of DNA polymerases in DNA amplification. *Proc. Natl. Acad. Sci. U.S.A.*, 86(23):9253–9257, 1989.

[173] P. Boček and A. Chrambach. Capillary electrophoresis of DNA in agarose solutions at 40 °C. *Electrophoresis*, 12(12):1059–1061, 1991.

[174] K. Klepárník, M. Garner, and P. Boček. Injection bias of DNA fragments in capillary electrophoresis with sieving. *J. Chromatogr. A*, 698(1-2):375–383, 1995.

[175] K. Klepárník, M. Zdena, and P. Boček. Electromigration behaviour of DNA molecules at the free electrolyte-polymer solution interface. *J. Chromatogr. A*, 772(1-2):243–253, 1997.

[176] O. Salas-Solano, M.C. Ruiz-Martinez, E. Carrilho, L. Kotler, and B.L. Karger. A sample purification method for rugged and high-performance DNA sequencing by capillary electrophoresis using replaceable polymer solutions. Quantitative determination of the role of sample matrix components on sequencing analysis. *Anal. Chem.*, 70(8):1528–1535, 1998.

[177] J.L. Viovy, T. Duke, and F. Caron. The physics of DNA electrophoresis. *Contemp. Phys.*, 33:25–40, 1992.

[178] P.O. Ekstrøm and J. Bjørheim. Evaluation of sieving matrices used to separate alleles by cycling temperature capillary electrophoresis. *Electrophoresis*, 27(10):1878–1885, 2006.

[179] D.C. Hinselwood, T.W. Abrahamsen, and P.O. Ekstrøm. *BRAF* mutation detection and identification by cycling temperature capillary electrophoresis. *Electrophoresis*, 26(13):2553–2561, 2005.

[180] T. Demana, M. Lanan, and M.D. Morris. Improved separation of nucleic acids with analyte velocity modulation capillary electrophoresis. *Anal. Chem.*, 63(23):2795–2797, 1991.

[181] A.S. Dukhin and S.S. Dukhin. Aperiodic capillary electrophoresis method using an alternating current electric field for separation of macromolecules. *Electrophoresis*, 26(11):2149–2153, 2005.

[182] T.A.J. Duke, R.H. Austin, E.C. Cox, and S.S. Chan. Pulsed-field electrophoresis in microlithographic arrays. *Electrophoresis*, 17(6):1075–1079, 1996.

[183] M.J. Navin, T.L. Rapp, and M.D. Morris. Variable frequency modulation in DNA separations using field inversion capillary gel electrophoresis. *Anal. Chem.*, 66(7):1179–1182, 1994.

[184] L. Liberman, D.D. Dershaw, P.P. Rosen, A.F. Abramson, B.M. Deutch, and L.E. Hann. Stereotaxis 14-gauge breast biopsy-how many core biopsy specimens are needed? *Radiology*, 192(3):793–795, 1994.

[185] C.S. Stewart, B.C. Leibovich, A.L. Weaver, and M.M. Leiber. Prostate cancer diagnosis using a saturation needle biopsy technique after previous negative sextant biopsies. *J. Urol.*, 166(1):86–92, 2001.

[186] Biopsy technique takes more samples, detects more prostate cancer. *American Cancer Society, ACS News Center*, Sep. 2001.

[187] K. Khrapko, H. Coller, and W.G. Thilly. Efficiency of separation of DNA mutations by constant denaturant capillary electrophoresis is controlled by the kinetics of DNA melting equilibrium. *Electrophoresis*, 17(12):1867–1874, 1996.