Springer Theses

Recognizing Outstanding Ph.D. Research

Aims and Scope

The series "Springer Theses" brings together a selection of the very best Ph.D. theses from around the world and across the physical sciences. Nominated and endorsed by two recognized specialists, each published volume has been selected for its scientific excellence and the high impact of its contents for the pertinent field of research. For greater accessibility to non-specialists, the published versions include an extended introduction, as well as a foreword by the student's supervisor explaining the special relevance of the work for the field. As a whole, the series will provide a valuable resource both for newcomers to the research fields described, and for other scientists seeking detailed background information on special questions. Finally, it provides an accredited documentation of the valuable contributions made by today's younger generation of scientists.

Theses are accepted into the series by invited nomination only and must fulfill all of the following criteria

- They must be written in good English.
- The topic should fall within the confines of Chemistry, Physics, Earth Sciences, Engineering and related interdisciplinary fields such as Materials, Nanoscience, Chemical Engineering, Complex Systems and Biophysics.
- The work reported in the thesis must represent a significant scientific advance.
- If the thesis includes previously published material, permission to reproduce this must be gained from the respective copyright holder.
- They must have been examined and passed during the 12 months prior to nomination.
- Each thesis should include a foreword by the supervisor outlining the significance of its content.
- The theses should have a clearly defined structure including an introduction accessible to scientists not expert in that particular field.

More information about this series at http://www.springer.com/series/8790

Bin Zhuang

Development of a Fully Integrated "Sample-In-Answer-Out" System for Automatic Genetic Analysis

Doctoral Thesis accepted by Tsinghua University

Springer

Author
Bin Zhuang
National Engineering Research
Center for Beijing Biochip Technology
Capitalbio Corporation
Beijing, China

Supervisor
Peng Liu
Tsinghua University, Beijing, China

ISSN 2190-5053 ISSN 2190-5061 (electronic)
Springer Theses
ISBN 978-981-10-4752-7 ISBN 978-981-10-4753-4 (eBook)
DOI 10.1007/978-981-10-4753-4

Library of Congress Control Number: 2017945919

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

To CongDong Lin (1924–2002)
My grandpa

Supervisor's Foreword

I was always bothered with tedious biochemistry experiments since the days of my graduate student career. There was a qualified lab with various instruments on the desks or floors and bottles of reagents on the shelves or in the containers. The labs were always run with some standard protocols for security reasons like contamination prevention or hazardous reagents. Lab fellows like me worked there and swam in the ocean of science. It did take time to work things out, especially those beyond the limits of forefathers. Despite the scientific exploring, the lab jobs in life science and related biomedical areas share some common techniques: cell culturing, DNA and protein extraction, purification of extractives, amplification under different conditions, electrophoresis, and chromatography, just to name but a few. Different labs may search for different biomarkers for different diseases or explore different functions of different genes; nevertheless, they all used the same technology to extract, amplify, and detect nucleic acid. In scientific research, DNA was tested to see if specific foreign gene was successfully transfected or find out the common loci all the screened clusters shared. In drug screening, nucleic acid was tested to determine targeting site of certain drugs. And in forensic science, genome information was detected to find out which suspects left their tissues in crime scenes. That notwithstanding, there are various applications in many other areas: to determine the types of pathogen, to judge the drug resistance of bacteria, to see if patient was sensitive to certain drugs, to estimate the risk of congenital disease, and so on. All of these were performed in various biolabs like mine. The classic workflow for gene test in the labs was summarized as the "golden three steps" of extraction, amplification, and detection.

For example, the whole raw blood sample was first centrifuged to separate white cells mixed with protease K solution for cell lysis, which took about 6–12 h under an incubation temperature of about 50 °C. DNA purification after cell lysis was usually performed following protocol of commercialized reagent kits. The commercialized protocol usually started with oscillating mixing ethanol or isopropyl alcohol (isopropanol) with samples and followed by a filtering-collection process

on filter column. Then elution was used to redissolve the DNA captured on the filter column. Several steps of centrifugal operation were required in this filtering-collection process to separate DNA from other impurities. In order to acquire templates of higher purity, the process is repeated several times. The acquired templates were then quantified before PCR amplification, which started with Master Mix preparation. The templates were added into the prepared Master Mix and loaded onto thermal cycler along with control groups for amplification. Gel electrophoresis was usually used for product detection. Agarose gel was needed to be prepared before electrophoresis, and the preparation process was always agarose powder weighing, buffer dispensing, heating, bubble removal, and annealing. Then the products were loaded onto the gel and electrophoresis was run. The whole process was about to take one whole day to complete. I reached my lab at 8:00 am and started with centrifuging blood samples. And a couple hours after lunch, the DNA purification began, and thermal cycles started just before my supper, after when I should start preparing agarose gel so that electrophoresis could be run just after PCR was finished. The results usually come out about 9:00 pm. With this whole process, there is no doubt why genetic analysis charged higher and took days to give results when it came to practical applications like clinical diagnosis and forensic identification. Although this tradition approach took time and labor to practice, it did give stable, reliable, and repeatable results if there was no operation miss. The protocols were summarized by former researchers and have already been practiced for decades. People trusted it and its results so I have no intension to change it. All I imagine was just integrating the protocols onto microfluidic platform so that they could be performed automatically.

The trail toward integration started in 2012 when I was hired as principal investigator in Tsinghua University. Bin Zhuang and Wupeng Gan were my first students, one with engineering background and the other with biochemistry background. And they were distributed with jobs matching their background. To let the machine do man's jobs, we were first required to refine our experimental operations. The funny thing at that time was that we presented our works declaiming an automated genetic analysis process while we were actually doing all experiments manually. Zhuang's first job was to construct an instrument for on-chip capillary electrophoresis, while Gan was trying to find some extraction approaches that may have a chance for on-chip integration. The most difficult parts of the instrumentation were adjustment of optical system and the data collection for software. The compact dimension of instrument limited the space for manual adjustment of optical components, and we had tried many times to decide the adjustable free degree of each component. Due to the large data size, the data collection took us so much time, and every data was required to be processed real time so that the calculated electrophoresis data would not be delayed. The instrument for electrophoresis was completed in autumn of 2012; meanwhile, we decided to extract DNA by filter paper and amplify nucleic acid by *in situ* PCR. Gan used Fusion 5 to extract DNA and optimized the PCR conditions off-chip so that we could use commercial RT-PCR system to quantify the extracted DNA. Once the electrophoresis platform

and extraction approach were settled, it came to the most arduous part of the research: the on-chip PCR amplification. It took Zhuang about 9 months to figure out the correct temperature curve and the suitable temperature calibration protocol. During that time, the sealing problem of on-chip valve was set aside temporarily. And after our first successful on-chip singlet PCR, on-chip valve started bothering us. By repeated improvement, the pre-bonding approach (seen in Chap. 3) was decided, and before that, we had tried various temperature and pressure parameters for one-step bonding. The integrated sample preparation module was completed in summer 2014, when it can only amplify simple, common, and short fragment like β-actin. It took another several months for further optimization on system and reagents to perform multiple PCR. The afterward development in combination of two instruments went more smoothly. Puncturing the chip to transfer PCR products was an impressive idea and shaping it like a feather pen was more artful design. In March 2015, the whole integrated system was fully completed, and another 3 months was taken for further verification, including the pharmacogenetic typing of warfarin-related polymorphisms. The original dissertation was submitted earlier, and that made it a pity that the LOD verification and the application on warfarin-related pharmacogenetic typing were not included. So, when I heard from Zhuang that there is a chance to publish the book, I strongly recommended him to rewrite Chap. 4 with related verification and application added. So, in this book, the end of Sect. 4.5 and the whole of Sect. 4.6 were not included in original dissertation. I hope the added content will help readers with a better understanding of this system.

In 2014, we started to cooperate with CapitalBio Technology to industrialize the system. At first, the industrialization only focused on the electrophoresis platform. However, when the sample preparation system was published, the goal was readjusted to commercialize the whole integrated system, including the instrument and the chips. Zhuang fastened the development process by bringing in other researchers and engineers to join the group. And we discussed together to determine our first application aiming at in-field forensic identification. So, once Zhuang achieved his degree, he joined the corporation for further industrialization of the system, as did Gan. Of course, the industrialization in corporation is totally different from the research in the university. Although we developed the system from industrial requirements, there was still a long way to develop a real product. Some components and parts of instrument were needed to be reappraised in order to lower the cost; the software should be reprogrammed on C++ platform for improved adaptability; beltline for chip fabrication was needed to be established and the mass flow for mass production required further optimization; and the expiry date of on-chip stored reagents should be verified. Besides, the reliability, repeatability, and anti-contamination ability should all be verified in a longer term of period before the product was ready for the market. What we see in this book will appear quite different from the final product. As far as I know, a cassette has already been designed to mount PMMA chip and glass chip together for improved user experience. All in our lab are looking forward for the industrialized system being delivered to the market.

This book presents an approach toward integrated "sample-in-answer-out" genetic analysis system, mainly from engineering aspect. Zhuang and I both hope the book would be able to help readers and researchers with some inspiration toward further fully integrated microfluidic systems. We also hope that our research would help more people and benefit the society.

Enjoy the book.

Tsinghua University, China By Peng Liu
2017/3/9

Abstract

A fully integrated and automated genetic analysis system has been developed based on the microfluidic technologies. The traditional three steps—DNA extraction, PCR amplification, and CE—were integrated together to form an automated genetic analyzer which is able to collect genetic information directly from blood sample. The work of this dissertation includes:

1. A general platform for on-chip capillary electrophoresis (CE) has been developed. This platform, with a dimension of only $48 \times 35 \times 18$ cm^3, is much more compact than the traditional ABI systems and can be either desktop or vehicle-type. The platform includes a reusable CE chip, a confocal optical system, a scanning stage for multichannel parallel detection, four high-voltage modules, several A/D and filter modules for data collection and processing, and a mainboard with related software. The platform is used for electrophoretic separation and detection of reacted products.
2. An integrated DNA extraction and amplification process was achieved by constructing an automated extraction and PCR module. The module is formed by a disposal extraction and PCR chip, a fluidic control system and a thermocycler. We bonded two plastic plates together to form the disposal chip, and such a method is more accessible for industrialization. Once blood sample was added into the system, the DNA template will be twisted on the filter paper, which is the principle used here for DNA extraction, and then an *in situ* PCR will be carried out without elution. The whole process took less than 100 min and as little as 0.3 μL blood was able to be tested when 59-bp β-actin gene was amplified. The DNA extraction and amplification module was used for congenital hearing loss diagnosis and screening. The module received blood samples from subjects and gave out products. We were able to tell homozygous and heterozygous apart from wild type once the products were detected.
3. A fully integrated and automated genetic analyzer has been constructed by combining the DNA extraction and amplification module and the general CE platform. A really automated blood-in-result-out process was achieved. In this

system, the amplification is carried out on a disposal plastic chip, while the glass CE chip is reusable. The composite chip design not only makes our system more accessible for industrial manufacture but also guarantees us the reliability and stability of every module. The whole automated genetic analysis, from DNA extraction to capillary electrophoresis, can be completed in 2 h by our instrument.

Keywords Fully integrated; Capillary electrophoresis; Congenital hearing loss; STR; Composite chip

Preface

Miniaturization-based system has helped a lot in reducing the size of the computer, which occupies great space in the laboratory, to a notebook. A fully automated bioanalytical system involving sample preparation, biochemical reaction, and result detection has always been a dream for industry for over one hundred years. In the mid-1990s, the concept "Laboratory on a Chip" was first proposed, and the world started to see where the hope lies. I still remember in 1998 I published a research article in *Nature Biotechnology* with "Lab on a Chip" as the cover story together with my colleagues at Nanogen in San Diego. But the system we put together has not been really fully integrated for nucleic acid analyses as the amplification was missing. Since then almost twenty years have gone unnoticed and much hard work has been devoted to achieve the goal of "Lab on a Chip" by research labs around the world including my lab at Tsinghua University in Beijing, China.

The thesis published here is written by my PhD student Bin Zhuang. In this book Dr. Zhuang tells his story of how to build a completely integrated genetic analyzer which involves the extraction of DNA from a blood sample, the PCR amplification of the purified DNA template, and the Sanger sequencing or the short tandem repeat-based DNA profiling. The entire process was made possible by employing the modular microfluidic chips and takes only two and half hours. This system has been applied for rapid genetic analysis of inherited hearing loss and the pharmacogenetic typing of multiple warfarin-associated single nucleotide polymorphisms. These experiments demonstrate both the universal applicability and the reliability of this system. To build such a complicated system, a combined background of biomedical science, electromechanical engineering, materials science, and chemistry is a must. Bin had an undergraduate training in precision instrument and a graduate training in biomedical science. With the help from another PhD student of mine Wupeng Gan and my colleague Dr. Peng Liu who are strong at chemistry and material science, the dream team has successfully built the "sample-to-answer" system for DNA analysis and achieved the goal of "Lab on a Chip."

The star system is now being commercialized by CapitalBio Technology in Beijing. We all are hoping that the industrial version of this system can soon come to the market to make a good story for translational medicine.

By Jing Cheng, Ph.D.
Member of the Chinese Academy of Engineering
Cheung Kong Professor
Tsinghua University School of Medicine
Beijing, China
15th April, 2017

Publications

[1]. Gan WP,# Zhuang B,# Zhang PF, Han JP, Li CX, Liu P. A filter paper-based microdevice for low-cost, rapid, and automated DNA extraction and amplification from diverse sample types. Lab Chip 2014;14:3719–28.

[2]. Zhuang B, Gan WP, Wang SQ, Han JP, Xiang GX, Li CX, Sun J, Liu P. A fully automated sample preparation microsystem for genetic testing of hereditary hearing loss using two-color multiplex allele-specific PCR. Anal Chem. 2015;87(2):1202–9.

[3]. Zhuang B, Han JP, Xiang GX, Gan WP, Wang SQ, Wang D, Wang L, Sun J, Li CX, Liu P. A fully integrated and automated microsystem for rapid pharmacogenetic typing of multiple warfarin-related single-nucleotide polymorphisms. Lab Chip 2016;16(1):86–95.

[4]. Han JP, Sun J, Zhuang B, Liu P, Zhao XC, Li WS, Ji AQ, Ye J, Liu Y, Li CX. The electrophoresis on microfluidic chips and its application in forensic science. Prog Biochem Biophys 2015;42(12):1128–35.

[5]. Han JP, Sun J, Wang L, Liu P, Zhuang B, Zhao L, Liu Y, Li CX. The optimization of electrophoresis on a glass microfluidic chip and its application in forensic science. J Forensic Sci., 2017; just accepted.

# Joint first authors with equal contributions.

Acknowledgments

Sincere thanks go to my supervisor, Prof. Jing Cheng, from Tsinghua University for his careful guidance on my research and enthusiastic concern on my development.

Thanks go to Prof. Peng Liu for his guidance and help on my research and for his support and encouragement that helped me through all the difficulties in the research. The achievements would never be this far without his support.

Thanks also go to Prof. Yuxiang Zhou, Prof. Wanli Xing, Prof. Guoliang Huang, Assoc. Prof. Yong Guo, Assoc. Prof. Ran Liu, Dr. Youchun Xu, and Dr. Lan Xie from Tsinghua University Medical Systems Biology Research Center for their concern, support, and help.

Thanks to Director Dong Wang, Director Lei Wang, and Director Guanxin Xiang from CapitalBio Technology for their valuable advices and generous help in the research.

Thanks also to my colleagues Wupeng Gan, Shuaiqin Wang, Junping Han, Yin Gu, and Ying Zhang for their help on the experiments and to Shengtai Bian, Pengfei Zhang, Shisheng Su, Tongzhou Wang, Na Sun, and Xiangrui Meng for valuable discussions.

Thanks go to Xiang Chen, Xinying Zhou, Liang Bai, Xiao Lin, Yuting Sun, and Yan Liu from the Institute of Engineering and Technology, CapitalBio Technology, for their help.

Thanks also go to my family for supporting me for all these years. I love them.

The project was supported by the National High-tech R&D Program (863 Program) and the Natural Science Foundation of China, which are hereby acknowledged.

Contents

Abbreviations

DNA	Deoxyribonucleic acid
RNA	Ribonucleic acid
EB	Ethidium bromide
PCR	Polymerase chain reaction
STR	Short tandem repeat
SNP	Single nucleotide polymorphisms
CE	Capillary electrophoresis
EOF	Electroosmotic flow
OFC	Oxygen-free copper
μTAS	Micro total analysis system
PDMS	Poly-dimethylsiloxane
PMMA	Poly-methylmethacrylate
PEG	Polyethylene glycol
BSA	Bovine serum albumin
LPA	Linear poly-acrylamide
PCB	Printed circuit board
PMT	Photomultiplier tube
RTD	Resistance temperature detector
DAQ	Data acquisition
LOD	Limit of detection
HR	High resistance
A/D	Analog/digital
D/A	Digital/analog