A Personal Narrative of 6 Pre-University Research Projects Over 7 Years (2009-2015) Yielding 19 Manuscripts

Abstract

Acquisition of research skills, including scientific enquiry, is an important requirement in scientific education, after the acquisition of a body of fundamental knowledge. Working on research projects is a direct means to gain research skills, as well as gaining a firsthand experience of the research environment. Here, I shall narrate my experience and learnings as a research mentor involving 22 pre-university students in 6 research and development projects over a period of 7 years, yielding a total of 19 peer-reviewed manuscripts. I have 3 intentions to this narration - (a) to demonstrate that pre-university students can carry out useful research, (b) summarize my learning experience in this journey, and (c) providing some pointers and encouragement to my fellow mentors and intended mentors. I learnt that (1) pre-university students can produce publishable work but (2) the scope of the projects must be well-defined with specific and measurable end-points, (3) the involvement of the mentor is substantial both in project formation and project management, (4) quality work can be achieved when students understand the background and context of their work, (5) enduring working relationship between students and mentors requires time and efforts to build, and (6) the students can pleasantly surprise the mentor with their quality of writing and reasoning. Upon reflection, I feel that I gained as much as the mentees / protégés, if not more, and this encouraged me.

Keywords: Pre-university; High school; Research projects; Publications; Critical success factors

Maurice HT Ling*[1,2]
[1]*Colossus Technologies LLP, Singapore*
[2]*School of BioSciences, The University of Melbourne, Australia*

*Corresponding author: Maurice HT Ling, Colossus Technologies LLP, Singapore; School of BioSciences, The University of Melbourne, Australia,
Email: mauriceling@acm.org

Introduction

Scientific enquiry is important in scientific education [1], in addition to acquiring a body of knowledge. Upon acquiring a body of knowledge, the student is expected to be able to apply the knowledge in new settings. A common tool for teaching scientific enquiry is to engage students in project work or problem-based learning [2]; which had been carried out at all levels of scientific education, from high school [3-6] to undergraduate studies [7-13]. Such research skills are exceptionally crucial in bioinformatics as bioinformatics is primarily a hands-on, minds-on endeavour [1]. This emphasis on bioinformatics research and problem-solving skills are supported by significant number of bioinformaticists [14-16].

Several studies have enumerated various benefits of incorporating research experience into scientific education [3,6,8,9,11,13-15]. Krasny [3] studied high-ability students and found that students gained a better understanding of the research process while graduate students acting as mentors gained their first mentoring experience. Moreover, students with research experience are more likely to enter [6,8,13] and maintain a career in science [6]. This may be due to increased motivation, independence, and active learning [12], which may be the result of increased confidence and role-modeling of the research mentor [11]. Research experience is crucial in providing a platform to develop data handling and analytical skills, and gaining a deeper understanding of the theoretical principles [15].

In this article, I account my experience as a research mentor over 7 years (2009-2015) using a multi-case [17-20] personal narration [21-25] approach involving 22 pre-university students in 6 research and development projects with several intentions - (a) to demonstrate that pre-university students can carry out useful research, (b) summarize my learning experience in this journey, and (c) providing some pointers and encouragement to my fellow mentors and intended mentors. Personal narration, sometimes also known as storytelling [21], is autobiographic in nature with rich emotional and contextual cues [22] that are required to describe the human experience [25]. Thus, personal narration puts the human into the story [26,27]. Personal narration also overlaps with participatory action research where reflective self-inquiry is carried out with the intention to improve subsequent actions [28-30]. In a series of cases over a period, it is highly unlikely that learning from preceding cases does not impact on succeeding cases even though this is not intentional. Despite so, there are elements of participatory action research spanning across my narration. Multi-case study is an extension of single case study [31,32] where more than one cases are presented. Thus, multi-case study offers the ability to compare between different cases [32-37] to unravel underlying patterns while providing the rich and in-depth contextual analysis in single case studies [31,32].

Personal Narration

In this section, I will narrate the 6 research and development projects and within each project, the narrative will be in chronological order to illustrate the learning process and progression within each project. A timeline and publication achievements of the 6 projects and 22 students in question is given in Figure 1 below.
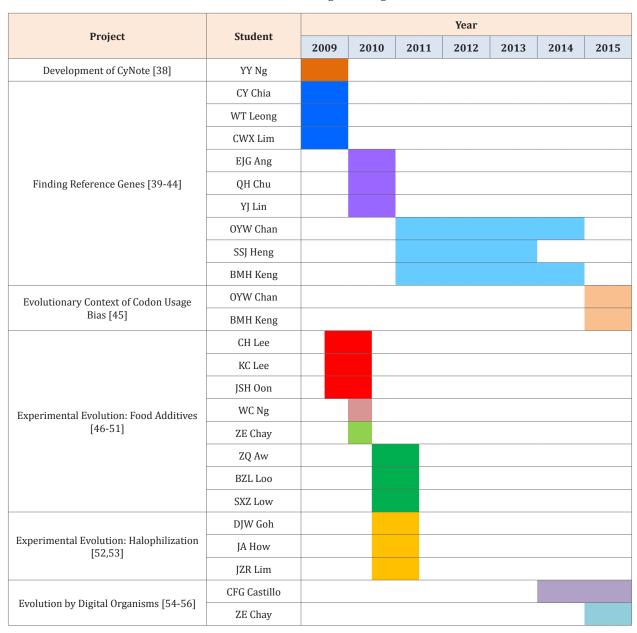
Project	Student	Year						
		2009	2010	2011	2012	2013	2014	2015
Development of CyNote [38]	YY Ng	■						
Finding Reference Genes [39-44]	CY Chia	■						
	WT Leong	■						
	CWX Lim	■						
	EJG Ang		■					
	QH Chu		■					
	YJ Lin		■					
	OYW Chan			■	■	■	■	
	SSJ Heng			■	■	■		
	BMH Keng			■	■	■	■	
Evolutionary Context of Codon Usage Bias [45]	OYW Chan							■
	BMH Keng							■
Experimental Evolution: Food Additives [46-51]	CH Lee	■						
	KC Lee	■						
	JSH Oon	■						
	WC Ng	■						
	ZE Chay	■						
	ZQ Aw		■					
	BZL Loo		■					
	SXZ Low		■					
Experimental Evolution: Halophilization [52,53]	DJW Goh		■					
	JA How		■					
	JZR Lim		■					
Evolution by Digital Organisms [54-56]	CFG Castillo						■	■
	ZE Chay							■

Figure 1: Timeline of Projects.

Development of electronic notebook (CyNote)

My doctoral thesis [57] is on bioinformatics. Instead of using regular paper-based laboratory notebooks, I used an electronic laboratory notebook (ELN) [58]; jointly developed by Oak Ridge National Laboratory, Lawrence Berkeley National Laboratory, and Pacific Northwest National Laboratory, under U.S. Department of Energy's DOE2000 Collaboratory Program; to manage my scripts and analysis output. I envisioned an ELN for biologists and bioinformaticists where common bioinformatics tools can be made available via the ELN and the analysis results be stored. I have been using Python programming language for all my work since my honours and postgraduate years; hence, a default programming language to base my subsequent work on.

The opportunity came when I was a lecturer in the School of Chemical and Life Sciences (CLS) at the Singapore Polytechnic, and was approached by a colleague in School of Media and Information Technology to consider accepting an internship student, YY Ng. A semester long internship is a mandatory requirement for the

diploma that YY Ng was pursuing. The director of CLS approved of this internship.

YY Ng worked in the library made it a point to learn Python programming. At that point in time, my intention was to develop a web-enabled ELN rather than a standalone system as many of the bioinformatics functions provided by BioPython [59], a Python library for computational biology, requires web access. Our original intention was to develop on CherryPy [60], a Python web application framework, but quickly found that CherryPy is too "low leveled", which is useful for complete customization. Web2Py [61], on the other hand, made many assumptions and provide a set of application structure (even though in-depth customization is still possible) to enable rapid application prototyping, which is suitable for CyNote development. I met up with YY Ng every morning in the library to discuss the issues he faced the preceding day and development for the day. YY Ng was encouraged to ask for help in Web2Py Google Group where I will supplement his questions in the group. Massimo, the creator of Web2Py, and other active members of the group are patient to our questions and rendered significant help. Once YY Ng realized that an ELN is essentially a stricter version of a blog application (the common ground), together with the application framework of Web2Py, YY Ng made great strides in development – where notebook entries are synonymous with blog entries, and supplementary notations on each entry are synonymous to comment entries within a blog entry. With each progress and solved issue, YY Ng's confidence and enthusiasm increased. At the end of his internship, we have a working CyNote where I can add bioinformatics tools on. Subsequently, we managed to present CyNote as a 21 CFR Part 11 [62] compliant ELN in the inaugural PyCon Asia Pacific as a peer-reviewed conference paper [38], with YY Ng as co-author. YY Ng passed his repeated subject, awarded "A" for his internship (which significantly improved his grade point average as internship is equivalent to one semester of coursework), and a peer-reviewed article in his name.

Finding reference genes

This project started as a project under the Science Mentorship Programme (SMP, www.moe.gov.sg/education/programmes/gifted-education-programme/special-programmes/science mentorship-programme), managed by the Gifted Education Branch at the Ministry of Education, Singapore. Its aim is to provide a platform for "developing pupils' interest and talent in scientific research" where "they provide opportunities for pupils to deepen their knowledge in a particular area, to learn about the latest research developments, and to be challenged by the intellectually stimulating process of scientific research." Only year 3s and 4s (15 to 16 years old) students in gifted education and recommended by their school, can participate in SMP as project students. A group of 2 or 3 students is assigned to a project for 6 months (February to July), 3 hours weekly. Hence, time is a significant issue as the students must juggle their regular school work and other academic and social commitments. The students must write up a term paper at the end of the project and present their poster at the Youth Science Conference in September of that year. Singapore Polytechnic is a participant in SMP and I chose to participate in this programme as a mentor in 2009.

My project offering is a bioinformatics project on finding reference genes for mouse liver. Reference genes are genes whose expressions ideally do not vary across different experimental conditions, which enables them to be used as expression benchmarking genes in transcriptomic studies [63-65]. CY Chia, CWX Lim, and WT Leong; 15 year old students from Nanyang Girls' High School; are my first SMP group in 2009. In preparation, I had downloaded 8 mouse liver microarray datasets from NCBI Gene Expression Omnibus. Every gene in each of the 8 datasets was analyzed for their coefficient of variation (CV), which is defined as the quotient of the standard deviation to its arithmetic mean. Hence, low CV corresponds to expressional stability; thus, suitable as reference genes. After analyzing each dataset separately, we found that Microtubule affinity-regulating kinase 3 (MARK3) to have low CV (CV < 0.08) across all 8 datasets, making MAPK3 the most suitable reference gene for mouse liver. More importantly, we found that commonly used reference genes, such as Glyceraldehyde-3-phosphate dehydrogenase (GAPDH), showed large expression variation (CV = 0.47) in one of the datasets, making it unsuitable as a reference gene for mouse liver. A lot of time was devoted into writing their term paper as I saw it as consolidation of their understanding. I also witnessed examples of theoretical misunderstanding or uncertainties as they were tasked to put their thoughts and arguments on paper; thereby, providing me with a chance to correct them. As a result, their SMP term paper was submitted to IUBMB Life and accepted for publication [39], which encouraged me continue my participation in this programme and further strengthened my belief that even high school students are able to contribute to the scientific community provided that the project is scoped appropriately. I was also awarded Outstanding Mentor by the Ministry of Education.

In the following year (2010), I proposed to find reference genes in mouse endocrine glands. This project was taken up by QH Chu, EJG Ang, and YJ Lin; 15 year old students from Raffles Institution. In this project, 15 microarray datasets were used and analyzed using the same method as Chia et al. [39]. Similarly, we found that GAPDH showed large expression variations (CV as high as 0.98), making GAPDH probably one of the worst choices for reference genes, but what was more interesting was that MARK3 that was found to be a suitable reference gene in mouse liver [39] was found to be within the list of suitable mouse endocrine reference genes. In addition, this study reflected the findings of Chia et al. [39] that ribosomal proteins are relatively stable in expression. However, the biggest drawback of this study is that the microarray datasets in question are from a variety of microarray chipset, resulting in difficulties in a proper comparative study as different microarray chipset can have different sets of probes. Hence, this project is not suitable for publication in its own rights even though they are required to write up a term paper. Nonetheless, this project forms a suitable case example to show that common housekeeping genes, such as GAPDH, may not be suitable reference genes in every case. This is supported by studies showing that GAPDH has stable expression profile in reticulocytes [66] but varies in expression across different developmental stages of rat fetal brain [67]. Thus, this project forms one of the case studies in a peer-reviewed book chapter on reference genes [40]; thereby, generating a publication

out of their work. Looking back, this project would be substantially more rigorous should datasets from the same microarray chipset were used and we realized the problem too late into the project. This also taught me that the mentor must consider the project more thoroughly, including running through the entire project mentally, as pre-university students will not have the experience to anticipate such issues.

I left Singapore Polytechnic at the end of 2010 and through a series of event; I got my final SMP group in 2011. They were 15 year old students from Raffles Institution whom had been unsuccessful in securing a SMP project. Their school acted upon this to run their SMP as school-based option in which the school had to either find an external mentor or assign one of their teachers as project mentor. Due to my involvement in QH Chu's SMP, I was asked if I was interested to take on a project which I agreed. This enabled a flexible working schedule and we will meet up every weekend or alternate weekends at a café to discuss project progress and to clear any doubts. In this project, I took the lesson from QH Chu's project and ensured that I had all the datasets from the same microarray chipset. In addition, I decided to narrow into finding reference genes in *Escherichia coli* instead of mammalian organs. As BMH Keng recounted their project formation years later:

I remember we failed to get any of the SMP projects we applied for during the first two rounds of project selection, which was pretty disappointing. Fortunately, our teacher-mentor managed to get hold of a few more options for us, of which the E. coli project was one. I thought it sounded good because it seemed to have the most relevance to biology, which I was interested in. The workload also didn't seem too demanding, which was nice because we were fairly heavily involved in core curricular activities and other activities at the time. So we came to a consensus that the E. coli project was the most appealing.

It was early 2011 that I had a meeting with Dr. Too Heng Phon (Department of Biochemistry, National University of Singapore) and he made a casual comment on Chia et al. [39] saying that CV may not be stringent enough. Hence, in this new project with Heng et al. , we used CV as the first screening. From the 4 datasets obtained from NCBI GEO, 3 datasets came from *E. coli* MG1655 and 1 came from *E. coli* W3110. We treated the datasets separately – 10% CV was used as threshold to potential reference genes from each of the *E. coli* MG1655 datasets and 39 genes were common between all 3 potential gene lists, which were then subjected to NormFinder [68]. We found that Glucan biosynthesis protein G (mdoG) is a suitable reference gene for both MG1655 and W3110 strains of *E. coli*, and found that GADPH showed high expressional variation and was not in the potential list of reference genes. Their SMP term paper was published almost in verbatim [41] and years later, both BMH Keng and OYW Chan felt that it was still an important step in their scientific pursuits – "*Looking back, I'd say that publishing made this research experience stand out from my other science-related pursuits*" and "*I definitely think the publication of that first paper was not to be overlooked.*"

After the successful completion of their SMP, SSJ Heng, OYW Chan, and BMH Keng were interested to move onto the next

project even though I had taken up a position in South Dakota State University, USA. At the end of the *E. coli* project, I was wondering that if it is unlikely for an organism-specific reference gene (a reference gene suitable for various organs within the same organism), can it be possible for organ-specific reference gene (a reference gene for a specific organ across various organisms)? Two microarray datasets [69], from the livers of *Spermophilus lateralis* (ground-roaming golden-mantle ground squirrel) and *Spermophilus tridecemlineatus* (tree habiting 13-lined ground squirrel) each, under different states of feeding and hibernation had been studied using the same microarray chip, were found in NCBI GEO. The experimental conditions were the same in these 2 datasets, making them a valuable resource for examining the possibility of organ-specific reference genes; or more specifically in this case, genus-organ or Spermophilus liver specific reference genes. Due to geographical and time zone differences (USA vs Singapore), progress was not as soon as desired – as OWY Chan recounted:

"*We did emailing and instant messaging to keep in touch with Maurice, but some clarifications are just best done over pen and paper in person. We asked many dumb questions multiple times, which Maurice was patient with, but was clear evidence of the lowered efficiency. We never thought of abandoning it though. We were sure that we'd complete it eventually - it was just a question of how much progress we could discipline ourselves into accomplishing that week.*"

However, I will strongly attribute this to the determination of the group. Eventually, our results suggested that Spermophilus liver specific reference genes is not likely as suitable reference genes in *S. lateralis* liver are not suitable for *S. tridecemlineatus* liver and vice versa. Since *S. lateralis* and *S. tridecemlineatus* are from the same genus and the livers are harvested at the same feeding conditions; thereby, enabling rigorous comparison which is the intention of Williams et al. [69]; we extrapolated to genus-organ-specific reference genes are not likely. By further extension, it suggests that lineage-organ-specific reference genes are not likely, and there is a constant need to identify reference genes for each organ and in each species. These findings are published [42]. In a certain sense, this [42] is a negative result as we failed to identify Spermophilus liver specific reference genes but it presented a good case that negative results can be published if the project scope is clear and the scientific rigors are maintained, as BMH Keng recounted – "*At first I thought it quite weird that we were publishing negative results, so the process was quite eye-opening for me.*"

One of the issues with available tools for finding reference genes from expression datasets; such as, NormFinder [68], geNorm [70], and BestKeeper [71]; is the inability to cope with large datasets containing thousands of genes, as in the case of microarray datasets. For example, geNorm [70] can only process 102 genes. This required transcriptome sized datasets to be filtered into candidate reference genes before analyses by these tools [68,70,71]. Hence, I was considering the development of a method that is comparable to these tools but with the ability to handle transcriptome sized datasets. Therefore, I proposed the idea of a method development project to BMH Keng and

OWY Chan in 2013 (SSJ Heng opted to leave the group after the spermophilus project), which excited them, as BMH Keng recounted – "*The fact that we had succeeded in publishing two papers gave us confidence in starting another one. I think we had gained a much better understanding of the field at the time, plus the premise of the third project sounded quite interesting - not just using available tools to find reference genes like the first two projects, but actually developing our own algorithm.*" We selected a subset of microarray data for analysis using NormFinder [68], geNorm [70], and BestKeeper [71]. The outputs were correlated – with the correlation between NormFinder's output and geNorm's output being the highest ($r^2 = 0.987$) while the coefficient of determination (r^2) between NormFinder/BestKeeper and geNorm/BestKeeper being 0.489 and 0.459 respectively. These results are consistent with study [72] showing that the identified reference genes by NormFinder [68], geNorm [70], and BestKeeper [71] vary even on the same dataset. However, these coefficients of determination set a minimum correlation between our algorithm(s) – we need to achieve a coefficient of determination of at least 0.489 when correlated to NormFinder's output. This was an exploratory project and OYW Chan and BMH Keng were encouraged to use their creativity to come up with new methods by building on existing methods; hence, it was of great fun. Eventually, we found 2 algorithms that achieved coefficient of determination of 0.892 when correlated to NormFinder's output. In addition, these 2 algorithms perform linearly to the number of data points. For example, if 10 samples of 1000 genes (10,000 data points) took 2 seconds to complete processing, 100 samples of 5000 genes (500,000 data points) is estimated to take 100 seconds to process. Two publications, one publishing the results and algorithm [43] and the other publishing the tool [44], resulted from this work. For the sake of ownership and fun, we had decided to name the tool as OLIVER (OLIgonucleotide Variable Expression Ranker), after OYW Chan.

Evolutionary context of codon usage bias

By the end of OLIVER [43,44], OYW Chan and BMH Keng were in the middle of their first year in a Junior College (equivalent to Year 11 and 12) and were due to take their General Certificate of Education (Advanced Level) at the end of the following year. Hence, I was not expecting a fourth project but they were enthusiastic about it, as BMH Keng recounted the scene then – "*Well I guess we were just talking about it at Coffeebean, and somebody said "ok why not" and the other person said "yeah let's do it, I'm sure we can manage". I don't think we really pondered too much into it.*" I was surprised and did not expect it. As a result, I was almost scrambling to think of a new project. At that point in time, I had just completed a NIH-funded project in South Dakota State University, examining evolution of antisense transcripts [73]. Hence, I have a deep interest in applying statistical genomics into the study of evolution, and since my previous employment as a Senior Scientist in Life Technologies, I had nurtured an interest in codon usage bias (CUB) [74]. After a few days of mental brewing, I decided to propose a project to examine the null hypothesis that CUB is not evolutionarily conserved, using the same statistical techniques used in my antisense expression study [73]. Using 18 codon usage tables across both eukaryotic and prokaryotic species from Codon Usage Database (www.kazusa.or.jp/codon), we found

that CUBs is correlated within species of the same class and the CUBs between closely related organisms are higher correlated compared to more distant organisms, suggesting that CUBs corroborates evolutionary time. This is the same conclusion from my antisense expression study [74] and that of Jordan et al. [75]. In addition, we showed that visualizing CUBs on a dendrogram fits our currently known phylogeny and we published this work in Asia Pacific Journal of Life Sciences [45].

This was the last project I did with OYW Chan and BMH Keng as they had to prepare for their milestone examination the following year. In total, they had worked with me for 4 years and generated 5 peer-reviewed journal articles [41-45] and 1 peer-reviewed book chapter [50]. In preparation for this narrative, I did ask them if they would go through all these all over again and would they encourage their juniors to do the same – their responses are positive.

Bryan: *Overall, I enjoyed the process of working on this series of papers, and if given the opportunity, I would do it again. Doing multiple interrelated projects in the same field really helped to solidify my understanding and made me better able to assume greater ownership of the work we did. I think the process was a lot more meaningful, in terms of personal growth as a researcher, as compared to completing just a single project. …… I feel that completing multiple projects is useful in honing the skills necessary to reach the eventual publication of a manuscript.*

Oliver: *This research experience taught me a great deal about the way science is done and the rationale behind it. As a student, I learnt much about a specific set of bioinformatics fields and as a growing adult I learnt much about myself. Student research provides opportunities that are uncommonly replicated elsewhere – independent work, depth of inquiry, exposure to a field expert. …… Yes. Please embark on a project. Don't stop at one paper because a good working relationship takes time to build and so does a base of usable field knowledge.*

Experimental evolution: food additives

As a lecturer in the Singapore Polytechnic, I was required to take 1 or 2 final year projects (FYP) with 3 diploma final year students (about 19 years old) per project. As Diploma in Biotechnology (a 3-year full time course) emphasized on practical skills (by the time each student enters final year, they would have accumulated about 300 hours of laboratory training, excluding 8 weeks of full time internship), there was an unspoken rule that most FYPs have substantial experimental component. Having completed my honours project and doctoral thesis in the Department of Zoology (The University of Melbourne, Australia), the importance of evolution was engrained into me, as Theodosius Dobzhansky would say, "*Nothing in biology makes sense except in the light of evolution*" [76]. Another philosophy that I hold dear is the words of Professor Ringertz of Karolinska Institute who said the following to Dr. Barbara McClintock, during her Nobel Prize citation, "*your work is encouraging because it shows that great discoveries can still be made with simple tools.*"

Hence, in formulating a FYP, I needed something that is doable within the capabilities of the students, requiring simple tools due to budgetary constraints (maximum of SGD 5000 per project),

with the end goal of producing a peer-reviewed manuscript. During my graduate days, I came to realize the advantage of having one or more peer-reviewed manuscript(s) during the "fight" for scholarships. Furthermore, compared to the time I was a student in Singapore Polytechnic (from 1997 to 1999); community services, leadership positions in school clubs, and testimonials had become commonplace. Hence, I reckoned that the only edge I can afford my project group is peer-reviewed manuscript(s).

As I was mooting a FYP, I came across the work by Richard Lenski's group at Michigan State University [77-80]. Their core protocol, which is bacterial sub-culturing, seems simple enough for diploma students; yet, complex enough to yield interesting insights. I thought of examining the adaptation of *E. coli* to food additives and I decided to use a sequenced strain, ATCC 8739, as the presence of sequenced genome is instrumental in the design of primers for DNA fingerprinting. A study treated pigs with ampicillin, a common antibiotic, demonstrated a significant increase in the occurrence of ampicillin-resistant *E. coli* from 6% to more than 90% after 7 days [81] as incomplete intestinal may led to sub-therapeutic doses of antibiotics in the faeces, resulting in evolutionary pressure for intestinal bacteria such as *E. coli* towards antibiotic resistance [82,83]. I reckoned that the same may happen in the presence of food additives. Three additives; sodium chloride (salt, common flavouring), monosodium glutamate (MSG, common flavor enhancer), and benzoic acid (common preservative); each with 2 concentrations (high and low) were used. In addition, 3 low concentrations and 3 high concentrations were combined into a combined high and low concentration treatments were tested, giving a total of 8 treatments. Sub-culturing was performed 3 days a week on Monday, Wednesday, and Friday, where cell density readings were taken. At every 3rd sub-culture, generation time was measured. At every 5th to 7th sub-culture, cells from one culture were swapped to another culture and generation time was measured to estimate fitness in new environments after potential adaptation. At every 12th sub-culture, DNA fingerprinting was carried out to give indication whether any observed phenotypical adaptations (evident by decreasing or increasing generation time) had a genetic basis. Hence, this project was experimentally simple but logistically and operationally demanding.

This project was undertaken by CH Lee, JSH Oon, and KC Lee. Their FYP was supposed to start at the end of April 2009. Many experimental ideas; such as, using live-dead staining; were suggested during our initial discussion and each idea was subjected to cost-benefit analysis as experimental time can be of short supply. Hence, the students had a good grasp of the project before the start of the experiment and despite so, some mistakes were made in the initial stages and the entire culture experiment had to be repeated. I felt that more mistakes, resulting in greater time loss, were likely if in depth discussion and thought process had not been carried out adequately.

I spent a significant amount of my time with my project group, often having lunch and dinner together. I am blessed with students who have the time availability (financially sound and not having to work extensively to make ends meet) and enthusiasm to be willing to stay back after class to discuss their project, analyze the data, and write drafts after drafts of their report. I

had the group start drafting their project report, starting from literature review, barely 3 months into their year-long project. I had chased KC Lee for his draft for nearly a month when CH Lee and JSH Oon were onto their second or third draft. Eventually, KC Lee submitted his draft to me and to my surprise; it was well-written and only required minor edits. I had seen KC Lee's former laboratory reports before and the draft he gave me after more than a month was miles better. This project was tedious but in the end, we found that *E. coli* can adapt to various additive media as generation time decreased over passages. However, our results suggest that the cells adapted to their specific media as swapping the media resulted in increased generation time. This is even for the case where cells in high concentration were swapped into low concentration of the same additive. DNA fingerprinting suggests mutations were involved in adaptation; yet, the fingerprints converged over passages suggesting that there is a presence of global stress response which had been suggested by previous studies [84,85]. The amount of work sustained over extended period resulted in 2 peer-reviewed publications [46,47] from this project.

This project was continued for another academic year with a new group of students – BZL Loo, ZQ Aw, and SXZ Low – extending the sub-culture from passage 70 to 150. The transition was relatively easy as JSH Oon had kindly guided the new group before he was due for his mandatory national service and the protocol had been developed. In fact, BZL Loo, ZQ Aw, and SXZ Low continued the laboratory notebook that CH Lee, JSH Oon, and KC Lee used; which enabled them to have a good sense of the data logging requirements. My personal guidance had been consistent – we had a formal meeting every Friday, when generation time were taken, where we discussed the progress of the project and resolve any issues faced. I recalled that I enjoyed the process of guiding the group to compare between their results and that of the previous group [47]. This work resulted in a succeeding manuscript [50] where two interesting observations emerged. Firstly, the rates of adaptation (in terms of decline in generation time) were reduced compared to pre-70th passage [47], suggesting that the adaptations were approaching plateau. I taught the group to estimate 95% confidence interval of their linear regression (decline in generation time across passages) using Jack-knife resampling. Secondly, the cells adapted to different concentrations of the same chemical additives as though they were different additives rather than different concentrations of the same additives, suggesting possibility of adaptation into an ecological niche (also known as ecological specialization). This was supported by DNA fingerprinting suggesting that the cells diverged genetically.

Near to the end of BZL Loo, ZQ Aw, and SXZ Low's FYP, I was wondering whether there will be any worth continuing this adaptation study as it was becoming obvious that the rate of interesting observations is likely to decrease. I had also thought that should I carry on this project for the next academic year, I might look at loss of adaptation instead. However, that was the point of time that I left the Singapore Polytechnic.

In addition, two side projects had emerged out of this adaptation project. As I was searching for methods to measure genetic distance based on DNA fingerprinting, I came across a list

of other potential distance measures and I had implemented most, if not all, of them into a Python module [86]. Subsequently, another student, ZE Chay (from the same batch as CH Lee), approached me indicating that he wanted to learn Python programming and work on something while waiting for his time for enlistment into national service (mandatory conscription for all Singaporean males). I got him to learn Python programming online and given my interests in statistics his first project (a practice project) was to implement statistical testing routines in Python, which resulted in a code paper [48]. Upon gaining more confidence with Python programming, I got ZE Chay to examine 20 other distance measures for their applicability in DNA fingerprinting using the DNA fingerprinting data from CH Lee, JSH Oon, and KC Lee's FYP where Nei and Li's distance measure [87] was used. Nei and Li's distance [87] is identical to Dice's distance [88]. We found that certain distance measures; such as Simpson's distance [89], have negligible correlation (r = 0.008) with Dice's distance [88]; hence, not suitable as substitute for Dice's distance [88]. Russel and Rao's distance measure [90], on the other hand, is highly correlated (r = 0.998) with Dice's distance [88]; thus a suitable substitute for Dice's distance [88]. A manuscript was published on these findings [49].

Another side project originated at the early phases of the main adaptation project. One of the key "amazement" of Lenski's long-term experimental evolution is the number of generations that *E. coli* had undergone [77-80] and measuring the number of generations is crucial for this very purpose. We used optical density as an indirect measure of cell density. However, Sezonov et al. [91] studied *E. coli* K-12 in Luria-Bertani broth (LB) and found that the average cell mass decreases after optical density (OD 600) more than 0.3. After which, a decreasing growth rate was observed, which the authors [91] attributed the depletion of fermentable sugars. However, LB is able to sustain extended periods of cell growth but at a slower growth rate due to the high availability of amino acids, which the cells could utilize as a carbon source. This resulted in the need to correct for decreased cell size when optical density is used as proxy for measuring cell numbers, such as in this project. However, this raised two questions. Firstly, is depletion of sugars the cause of reduced cell growth in other common media; such as, nutrient broth (NB); which was used in this project? Secondly, if there is decreased cell mass at high optical density, is it possible to use DNA content as a proxy to estimate cell numbers?

We performed media exhaustion experiment by repeated harvesting and re-inoculation of fresh cultures into LB, NB, and brain heart infusion (BHI), for 11 weeks and found that the sustainable cell counts (by spread plate counting) declined steadily to plateau by 8th week after peaking at 4th week. The pH of the media were neutral after 11 weeks (7.1 < pH <7.7). However, addition of glucose into LB, NB, or BHI rescued the media; thereby, confirming that lack of sugar is the major reason for growth decline [91]. We also found low correlations between optical density and viable plate counts ($r^2 < 0.2$), suggesting a change in cell size [91]. Using the genomic sequence of *E. coli* ATCC 8739, the genomic DNA content per million cells can be estimated at 515.35 nanograms (ng). By extracting genomic DNA from different aliquots of cells with known cell density and quantifying the extracted DNA, we extracted an average of 527

ng of genomic DNA per million cells, with a standard error of 10.605 ng, which is not significantly different from the estimated genomic DNA content. With the amount of extracted genomic DNA, we are able to estimate the number of cells and this is highly correlated ($r^2 = 0.93$) to the results from viable plate count. Hence, we demonstrated that lack of sugars results in growth depletion in LB, NB, and BHI; and cell size changes at different cell densities; and DNA content can be a proxy to estimate cell numbers; thus, publishing a manuscript in the process [51]. The attraction of this side project is the time contributed is minor compared to the amount of time contributed to the main project. In fact, this was considered low priority project that we nearly dumped. Looking back, I am glad that we persisted on.

Experimental evolution: Halophilization

This project stemmed from the first food additive project by CH Lee et al. At that point in time, I was quite certain that *E. coli* is able to adapt to various environmental conditions should sufficient time be given. Hence, I was considering a FYP project to stress the cells. A few options were considered; including, increased temperature, reduced oxygen, and increased preservatives (such as, benzoic acid). However, all these choices were not feasible due to resource constraints. For example, the laboratory for running FYPs is also used as a teaching laboratory; hence, it is not possible for me secure an incubator at a constantly higher temperature, neither was reduced oxygen a feasible option. Increased benzoic acid seems feasible but my source of benzoic acid was a gift from my fellow faculty in food technology; hence, purchasing or requesting for more benzoic acid could be logistically challenging. The most feasible option appeared to be salt (sodium chloride, NaCl) stress as sodium chloride is economically sustainable and is also a common food preservative. Hence, this project is called "halophilization". The intention was gradual increase of salt concentration, from 1% NaCl (by CH Lee et al.), at the rate of about 1% per month. Initial work by Doudoroff [92] suggested that 7% NaCl is bacteriostatic.

I recalled discussing this project with CH Lee, JSH Oon, and KC Lee, and the response was positive. JSH Oon was interested to kick start this project and another graduating diploma student, WC Ng (JSH Oon's classmate), was interested to work on something while waiting for national service. At that time, the school hosted several interns from Assumption University, Thailand, and I was privileged to have 3 interns attached to me for about 10 weeks – G. Pairoh, S. Charoenlustavee, and N. Tongpradith. I decided to carry out 4 halophilization replicates in parallel, which proved to be instrumental in one of our findings. Together, we started this project. This project was eventually undertaken by JZR Lim, JA How, and DJW Goh; whom were cohort mates of ZQ Aw, SXZ Low, and BZL Loo of the second adaptation project. At the point in time when this project was started, DJW Goh was doing his internship in the same laboratory under another lecturer; hence, witnessed the initiation of this halophilization project. JZR Lim recounted their thought process and their choice to consider this project for their FYP:

"*I only had one consideration approaching FYP which may come across as single minded, but I had the strong belief that the project never really mattered so much as the supervisor I was working with. This could be attributed in a small way towards my disinterest*

towards sciences or inability to differentiate the bigger picture of the topics given, but for me the reality was this was a long-term project that would push and pull individuals within a group, thus the strongest bond we have should not be the interest towards the given topic but the relationship with the project supervisor. (Interest in a subject would not in any way insulate from poor working relationships with group members or supervisors). The best project handled by the most inept fool would amount to nothing, was what I believe I said to them. I viewed the project in terms of concentric circles of relationships, with the group within the core circle of trust and the supervisor as someone who should be brought into this circle to work with collaboratively.

The shortlist ultimately was between Mr Goh and Dr Maurice Ling. Jian Ann [JA How] / Desmond [DJW Goh] were in favour of Mr Goh but I was in favour of Dr Ling, and having worked with Dr Ling prior and having seen firsthand his efficiency and effectiveness in getting tasks done and how he entrusts his subordinates with the authority to get things done, along with his ability to convey clear, concise, precise instructions and meticulous attention to detail, it was clear to me Dr Ling was someone I wanted to work with – even though I had no clue what the project was about at that point. I did make a point to speak with him personally to find out more regarding the project, but the key decisional choices were based solely on potential working relationships and consistency / reliability / efficiency of the supervisor in consideration.

I successfully persuaded my teammates to work with Dr. Ling and choose him as our primary option in the ballot. When the close friends of Desmond and Jian Ann, Samuel [SXZ Low] / Bryan [BZL Loo] didn't manage to get their balloted option of Mr Goh, I saw it as an opportunity to bring everyone together by suggestion and vouching for Dr Ling to them, and I recall Desmond aiding me in trying to pull them in with us."

JZR Lim, JA How, and DJW Goh were coached by JSH Oon and WC Ng during the first 2 weeks of experiments. This project involved regular sub-culturing and generation time estimation; hence, we planned it to coincide with ZQ Aw, SXZ Low, and BZL Loo's project so that both teams can work together. It was also helpful that SXZ Low, BZL Loo, JA How, and DJW Goh were in the same class; hence, identical lesson schedule. This also meant that JZR Lim and ZQ Aw could be backup for experiments if required. The same arrangement happened with CH Lee, JSH Oon, and KC Lee as well, where CH Lee and JSH Oon were classmates and KC Lee acted as backup for experiments. As a result, project discussions were usually done together across both groups.

Over the course of project, we managed to culture *E. coli* to 8% NaCl, which was above bacteriostatic level suggested by Doudoroff [92]. Despite so, minimum inhibitory concentration assay demonstrated that 8% NaCl adapted *E. coli* could grow in concentrations as high as 11% NaCl. Comparatively, 11% NaCl is half the saturation level of NaCl in water and about 3 times the salinity of sea water (average NaCl concentration in sea water is 3.5%). Hence, we had reduced the water activity requirement of *E. coli*. We were concerned about *Staphylococcus aureus* contamination as *S. aureus* is commonly found on human skin and is known to be highly salt tolerant [93,94]. To ensure that *S. aureus* contamination was kept at bay, we cultured the cells on MacConkey agar before cryopreserved as glycerol stocks to ensure integrity

of the glycerol stocks. We performed Gram staining regularly and used DNA fingerprinting to assay for contaminations. I recalled that we had to revert a few passages from glycerol stocks due to contaminations. All four replicates showed an increasing rate of generation time as the concentration of NaCl increases, suggesting osmotic stress to the cells. Hence, there was a need to balance between culture survival and the rate of NaCl increase. As the cultures adapted to increasing NaCl concentrations, we wondered whether the cultures varied their salt optimality (increasingly able to grow in higher NaCl concentrations while reducing the ability to grow in lower NaCl concentrations) or broadening the range of salt tolerance (increasingly able to grow in higher NaCl concentrations while maintaining the ability to grow in lower NaCl concentrations). Interestingly, one of the cultures appeared to be broadening the range of salt tolerance while the other three cultures appeared to be varying their salt optimality. We published these findings [52,53].

Evolution by digital organism

I left the Singapore Polytechnic in December 2010 and one of the issues was the custodian of the glycerol stocks that I had left behind. Although I managed to hand-over the custodial duties to the school (approved by the director of the school), I was left wondering how much enthusiasm I will have if I am asked to restart this experimental evolution study. Unless I am certain of extended employment or if there was sufficient resource for third-party custodian, I would find it difficult to restart knowing the possibility of not being able to bring these glycerol stocks with me. This was the point where I wondered and started to examine alternative platforms for my evolutionary studies. One of the possibilities was to study evolution *in situ*, which was to observe natural evolution, much like Charles Darwin and Alfred Wallace. In this case, glycerol stocking is not logistically possible as there is no possible way to freeze an island for the purpose of preservation. Another possibility was to run computer simulations of evolution using artificial life (ALife) organisms, also known as digital organisms (DOs) [95]. DOs or ALife are based on cellular automata, described by Christopher Langton [96], and had been used to study many aspects of evolution [97-103]. Studies using DOs had been supported by evolutionary biologist; such as, Thomas Ray, whom wrote one of the earliest artificial life simulators [97]; and even well-known experimental evolutionary biology; such as, Richard Lenski; had used DOs in his studies [101,104]. In fact, Richard Lenski co-authored Avida, a well-known artificial life simulators, with Charles Ofria [101]. The premise of studying evolution using DOs is intriguing to me compared to observing natural evolution. Essentially, working with DOs amount to creating one or more populations to test evolutionary hypotheses or events, which can be difficult to perform experimentally [105]. At the same time, it fits my interests in bio-inspired algorithms; such as, genetic algorithms [106]; and resolves the issue with glycerol stocks as "glycerol stocks" in DO context are computer files, which can be duplicated and stored at various locations at ease. I decided to write a DO library for Python, Digital Organism Simulation Environment (DOSE) [107, 108] during my leisure time in South Dakota, USA.

As the Vice-President of Python User Group (Singapore), I was involved in the organization of PyCon Asia-Pacific (APAC)

Conference since its inception in 2010. During the conference days of the Third PyCon APAC, I was introduced to Clarence Castillo, whom was a student volunteer from Republic Polytechnic, at the conference. Clarence was pursuing a course in Diploma in Information Technology then. While chatting with the student volunteers, I found that Clarence showed strong enthusiasm and interest in problem solving; hence, I was toying the possibility of taking Clarence as my protégé. I was surprised that Clarence has an underlying interest in biology and saw the uniqueness of using computer simulations to study evolution.

My original implementation of DOSE [107,108] was rather rudimentary on hindsight and needed code refactoring, as well as feature enhancements. Clarence happily took up this challenge and painstaking read each line of my code to understand its operations. The goal was to simplify the use of DOSE. The main enhancement was to add a 3-dimensional "world structure" onto the DOs. This enabled movements of DOs within the world and could be used to generate multiple habitats or homing ranges, which was not possible in a "world-less" population. After a few months of work, DOSE was finally ready for use.

We tested our product using a case study to examine the effects of migration on population genetic diversity. Three migration schemes were tested – in case one, the DOs remained in their own habitat (no migration); in case two, DOs can migrate to adjacent habitats (adjacent migration); and in case three, DOs can migrate across habitats (long migration). No migration can represent isolated tribes or villages. Adjacent migration can represent nomadic movements while long migration resembles transcontinental migrations observed today. The social aspect to this work would be that in the present day where transcontinental mating and resettlement are commonplace, are we turning into a global village genetically? Our simulation results suggest that adjacent migration increased the overall inter-habitat diversity but long distance migration reduced inter-habitat diversity, suggesting that we are indeed turning into a global village from the genetic point of view. This concurred with findings from both hoverfly [109] and human studies [110]. We published the enhanced version of DOSE [54] and I took the pleasure in informing Dr. Liew Beng Keat, President of Python User Group (Singapore) and Assistant Director of School of IT in Republic Polytechnic, where Clarence was pursuing a diploma under his charge – "*As an educator, you will be pleased to know that Clarence had "sacrificed" much of online gaming time for this work and he is gearing for the second paper*."

In the next project, we focused on using DOSE to study evolution and we chose to study antibiotics resistance, which had been a theme in my work since the food additive project by CH Lee et al. Antibiotics resistance is without doubt one of the major concerns in the medical industry and had reached crisis proportions [111]. As chemical resistance (such as antibiotics resistance) is a result of selective pressure in the presence of the chemical, it can be expected that the prevalence of resistance will decrease following withdrawal and disuse of the chemical. A study shows that reduced antibiotics use correlates with reduced of microbial resistance [112] but others reported contradictory results [113]. For example, Enne et al. [114] report the persistence of resistant *E. coli* 8 years after reduced of antibiotics use. DO is the only means

to study this phenomenon as it is neither responsible nor ethical to induce antibiotics resistance. In the first instance, we assumed no fitness cost in maintaining resistance traits in the absence of selective pressure. Our results show that selective pressure is likely to result in maximum resistance to the selective pressure. During de-selection, mimicking disuse of the specific antibiotics, a large initial loss and prolonged stabilization of resistance are observed but resistance is not lost to the stage of pre-selection. This suggests that a pool of partial resistant organisms persist long after withdrawal of selective pressure at a relatively constant proportion. Hence, it is likely that contradictory results regarding the prevalence of resistance following withdrawal and disuse of the specific antibiotics may be a statistical variation about constant proportion. In addition, subsequent re-introduction of the same selective pressure results in rapid re-gain of maximal resistance. Therefore, complete elimination of specific antibiotics resistance is unlikely after the disuse of antibiotics, once a resistant pool of micro-organism has been established. This was published [55]. In the second instance, we repeated the original study [55] but considered fitness cost in the maintenance of resistance traits in the absence of selective pressure. However, the conclusions are consistent with that of no fitness cost, suggesting that fitness cost may have limited role to revert from resistance to susceptibility and other factors had to be considered in future studies [115]. This sequel to the first instance was also published [56].

My Reflections

As I was narrating my stories, I had a chance to re-examine the circumstances of each project and its outcomes. More importantly, gave me a chance to ask myself several important questions – (1) why I decided to be a project mentor, especially mentoring several projects that are not mandatory within my paid job scope (in fact, only the 3 FYPs are mandatory within my job scope); (2) what did I learn and would I have done it differently; and (3) given a second chance, would I have gone down this path again?

Looking back, I am hard-pressed to pinpoint a moment when I decided to be a mentor. However, I can list three factors that I consider to be instrumental. Firstly, I was the eldest in my generation (which means that I do not have any elder siblings and I did not have any elder cousins, both maternal and paternal) until I was 21 years old. What happened when I was 21 is a family matter, which is out of the scope of this article. Being the eldest in my generation predisposed myself to be a natural *de facto* mentor (or child leader) to several of my younger cousins; thereby, entrusting me with mentorship opportunities – more like, I was never given the opportunity to reject. Several studies [116,117] suggested that previous mentoring experiences are indicative of the person's willingness to mentor. Moreover, willingness to mentor is correlated to the number of mentoring experiences [117]. Secondly, I believe that my past experiences as a protégé played a significant role. The most extensive mentorship experience was during my postgraduate studies where I had the great privilege to be working with good thesis advisers, Drs. Kevin Nicholas and Christophe Lefevre, during my doctoral days. Both are old enough to be my father and they allowed me extensive latitude to explore my work while gently guiding my research. I could still remember the day when Kevin, who is also my honours project supervisor, told me that I was to start my PhD as a student

but to end as a colleague; and if I had never argued with him during my postgraduate, it would be his failure as an adviser. Both Allen et al. [116] and Ragins & Scandura [117] pointed out that positive experiences as protégé can be indicative of willingness to mentor. Perhaps it is also due to my birth order, I may be more willing to mentor from an intuitive point of view than logical reasoning of cost-benefits [118]. If I were to be pressed for a reason, I might only be able to say that "it just feels like it is a right and worthy thing to do, and I am willing to do it."

As I was writing these narratives, and re-reading them, a scene from *Sex and the City* (Season 5 Episode 7 – The Big Journey) popped into my mind – where Carrie Bradshaw met up with Mr. Big after her book reading and Mr. Big told Carrie, "*I had no idea I hurt you so much. I got to say it was tough seeing it in print …… Not all in one sitting, one reading, one right after another, bam… bam… bam…*" The same happened to my head – these events were scattered across more than a thousand pages of my diary and I had to simplify and put them into a structure as I put the events in writing; thereby, enabling me to see themes across various stories; which had been therapeutic for me [119-122].

I realized that all the students in these narratives did not have an idea what they wanted to do in terms of the nature of the project in the beginning but rather, they wanted or had to do something. I attributed this to the general lack of experience on the part of the students; hence, it is the mentor's duty to moot a project of sufficient content to interest yet not overwhelming to the students. It is also the mentor's duty to contextualize the project in the bigger realm of science. My students showed interests to want to do something; such as, taking up my project offer out of their course requirements or coming to me with the intention to do something; or willing to do something extra and on top of what is needed for them; such as, reading up more or drafting a manuscript. In hindsight, this is about all they can offer and it can be a challenge for them to stick their heads out for a non-mandatory research project as there is a perceived expectation of the project on the part of the students even though the mentor may not see it that way [123]. Despite so, there is a perceived burden on the students to meet the supposed expectations of the mentor as failure to meet such expectations can be distressing [124,125]. It then becomes the mentor's task to have to bridge this expectation gap and be extremely patient for students will ask the same stuffs multiple times. In a way, I did felt that the amount of effort to teach a semester and the amount of effort to mentor a project group is the same despite the number of students being significantly different – it is concentrating the attention needed by a lecture room of students onto just a team of 3 project students and it will not be just imparting of skills and knowledge. At the very initial stage, the role of the mentor is the advocate of the field where he should build interest and context of the work. Only then, can there be a stable path for knowledge and skills transfer. If mentoring is a graduate course, then knowledge/skills transfer is a hurdle requirement. The mentor should be a friend, a counsellor, and likely, even a surrogate parent. I was told by a former colleague that there was a student once that got into trouble and was detained by the police, and the student called up his project mentor instead of his parents. In a way, mentoring relationship is trust building between the mentor and the protégé

[126-128]. Once the trust is built, it can be fulfilling and Ragins & Scandura [117] had suggested that the willingness of mentors can be associated to previous positive experiences as a mentor.

Given a second chance, would I have gone down this path again? My answer will be a resounding yes. At an autistic level, being a mentor had given me substantial satisfaction [129] that I had contributed significantly to someone's learning and the contributing to my own explorations in science. Being asked by students is a great way for me to improve and consolidate my understanding in the field [130]. By explaining to my students, I gained insights into how I learn and the inherent blind spots to my knowledge. Nothing beats the motivation to flip open a textbook or to search for another publication just to support your argument or to clear your protégés' doubts. At the external level, I had seen my protégés' growth in their intellect and reasoning skills. More concretely, there were multiple instances where my protégés rode on their published papers and gained entries into their desired course of studies and awarded competitive scholarships. I believe that their success is one of the best returns for my efforts [129,130]. Each of my mentors played an irreplaceable role in my achievements, now and in future. Therefore, why will I walk this path? Perhaps the simplest answer is – to repay the gratitude I have for my own mentors.

Conflict of Interest

The author declares no conflict of interest.

Acknowledgement

References

1. McNerney K, Hall N (2017) Developing a framework of scientific enquiry in early childhood: an action research project to support staff development and improve science teaching. Early Child Development and Care 187: 206-220.

2. Ling MHT (2017) Problem-based learning (PBL), an important paradigm for bioinformatics education. MOJ Proteomics and Bioinformatics 5(4): 00166.

3. Krasny ME (1999) Reflections on nine years of conducting high school research programs. Journal of Natural Resources and Life Sciences Education 28: 17-23.

4. Gordon C (1999) Students as Authentic Researchers: A new prescription for the high school research Assignment. School Library Media Research 2: 1-21.

5. Machluf Y, Gelbart H, Ben-Dor S, Yarden A (2017) Making authentic science accessible-the benefits and challenges of integrating bioinformatics into a high-school science curriculum. Brief Bioinform 18(1): 145-159.

6. Coker JS, Davies E (2002) Involvement of plant biologists in undergraduate and high school student research. Journal of Natural Resources and Life Sciences Education 31: 44-47.

7. Roberts LF, Wassersug RJ (2009) Does doing scientific research in high school correlate with students staying in science? A half-century retrospective study. Research in Science Education 39(2): 251-256.

8. Hathaway RS, Nagda BA, Gregerman SR (2002) The relationship of undergraduate research participation to graduate and professional education pursuit: an empirical study. Journal of College Student Development 43: 614.

9. Kappler U, Rowland SL, Pedwell RK (2017) A unique large-scale undergraduate research experience in molecular systems biology for non-mathematics majors. Biochem Mol Biol Educ 45(3): 235-248.

10. Kardash CM (2000) Evaluation of undergraduate research experience: Perceptions of undergraduate interns and their faculty mentors. Journal of Educational Psychology 92: 191.

11. Landrum RE, Nelsen LR (2002) The undergraduate research assistantship: An analysis of the benefits. Teaching of Psychology 29(1): 15-19.

12. Lopatto D (2007) Undergraduate Research Experiences Support Science Career Decisions and Active Learning. CBE-Life Sciences Education 6(4): 297-306.

13. Russell SH, Hancock MP, McCullough J (2007) Benefits of undergraduate research experiences. Science 316(5824): 548-549.

14. Smith JT, Harris JC, Lopez OJ, Valverde L, Borchert GM (2015) "On the job" learning: A bioinformatics course incorporating undergraduates in actual research projects and manuscript submissions. Biochem Mol Biol Educ 43(3): 154-161.

15. Via A, Blicher T, Bongcam-Rudloff E, Brazas MD, Brooksbank C, et al. (2013) Best practices in bioinformatics training for life scientists. Briefings in Bioinformatics 14(5): 528-537.

16. Wightman B, Hark AT (2012) Integration of bioinformatics into an undergraduate biology curriculum and the impact on development of mathematical skills. Biochem Mol Biol Educ 40(5): 310-319.

17. Gary W Houchens, Tom A Stewart, Sara Jennings (2017) Enhancing instructional leadership through collaborative coaching: a multi-case study. International Journal of Mentoring and Coaching in Education 6: 34-49.

18. Ma Q (2017) A multi-case study of university students' language-learning experience mediated by mobile technologies: a socio-cultural perspective. Computer Assisted Language Learning 30(3-4): 183-203.

19. Uiboleht K, Karm M, Postareff L (2016) How do university teachers combine different approaches to teaching in a specific course? A qualitative multi-case study. Teaching in Higher Education 21(7): 854-869.

20. Liz F (2017) A Multi-Case Study of Student Perceptions of Instructor-Created Videos in Online Courses. International Journal for Scholarship of Technology Enhanced Learning 1(2): 67-90.

21. Dahlstrom MF (2014) Using narratives and storytelling to communicate science with nonexpert audiences. Proc Natl Acad Sci U S A 111(Suppl 4): 13614-13620.

22. Edwards SL (2015) The Personal Narrative of a Nurse. J Holist Nurs 34(2): 154-161.

23. Bhattacharya K (2015) The Vulnerable Academic. Qualitative Inquiry 22(5): 309-321.

24. Dush L (2016) Nonprofit Collections of Digital Personal Experience Narratives. Journal of Business and Technical Communication 31(2): 188-221.

25. Ingraham C (2016) The Scope and Autonomy of Personal Narrative. Written Communication 34(1): 54-74.

26. Hamm JA, Leonhardt BL (2016) The Role of Interpersonal Connection, Personal Narrative, and Metacognition in Integrative Psychotherapy for Schizophrenia: A Case Report. J Clin Psychol 72(2): 132-141.

27. Bennett D, Reid A, Rowley J (2017) Student musicians' experiences of reflexivity during internships: Personal narratives and complex modalities. International Journal of Music Education 35(3): 460-475.

28. Baum F, MacDougall C, Smith D (2006) Participatory action research. J Epidemiol Community Health 60(10): 854-857.

29. Apgar JM, Cohen PJ, Ratner BD, de Silva S, Buisson MC, et al. (2017) Identifying opportunities to improve governance of aquatic agricultural systems through participatory action research. Ecology and Society 22(1): 9.

30. Darby S (2017) Making space for co-produced research "impact": learning from a participatory action research case study. Area 49(2): 230-237.

31. Yin RK (2003) Case study research: design and methods. Sage Publications, London.

32. Gomm R, Hammersley M, Foster P (2002) Case study method - key issues, key texts. Sage Publications, London.

33. Towgood KJ, Meuwese JD, Gilbert SJ, Turner MS, Burgess PW (2009) Advantages of the multiple case series approach to the study of cognitive deficits in autism spectrum disorder. Neuropsychologia 47(13): 2981-2988.

34. Akhavan P, Pezeshkan A (2014) Knowledge management critical failure factors: a multi-case study. VINE: The Journal of Information and Knowledge Management Systems 44(1): 22-41.

35. Chua A, Lam W (2005) Why KM projects fail: a multi-case analysis. Journal of Knowledge Management 9(3): 6-17.

36. Eryılmaz R (2016) Self-identification of three bilingual language teachers: a multi-case study. Sustainable Multilingualism 9(3): 32-49.

37. Oliveira R, Sousa L, de Mello R, Valentim N, Lopes A, et al. (2017) Collaborative identification of code smells: a multi-case study. In Proceedings of the 39th International Conference on Software Engineering: Software Engineering in Practice Track, (IEEE Press), p. 33-42.

38. Ng YY, Ling MHT (2010) Electronic laboratory notebook on Web2Py framework. The Python Papers 5: 7.

39. Chia CY, Lim CWX, Leong WT, Ling MHT (2010) High expression stability of microtubule affinity regulating kinase 3 (MARK3) makes it a reliable reference gene. IUBMB Life 62(3): 200-203.

40. Too IHK, Heng SSJ, Chan OYW, Keng BMH, Chia CY, et al. (2013) Identification of Reference Genes by Meta-Microarray Analyses. In James V Rogers (Ed), Microarrays: Principles, Applications and Technologies. Nova Science Publishers, Inc, USA.

41. Heng SSJ, Chan OYW, Keng BMH, Ling MHT (2011) Glucan biosynthesis protein G (mdoG) is a suitable reference gene in Escherichia coli K-12. ISRN Microbiol 2011:469053.

42. Keng BMH, Chan OYW, Heng SSJ, Ling MHT (2013) Transcriptome analysis of Spermophilus lateralis and Spermophilus tridecemlineatus liver does not suggest the presence of spermophilus-liver-specific reference genes. ISRN Bioinformatics 2013: 8.

43. Chan OYW, Keng BMH, Ling MHT (2014) Correlation and variation based method for reference genes identification from large datasets. Electron Physician 6(1): 719-727.

44. Chan OYW, Keng BMH, Ling MHT (2014) Bactome III: OLIgonucleotide Variable Expression Ranker (OLIVER) 1.0, Tool for identifying suitable reference (invariant) genes from large microarray datasets. The Python Papers Source Codes 6: 2.

45. Keng BMH, Chan OYW, Ling MHT (2014) Codon usage bias is evolutionarily conserved. Asia Pacific Journal of Life Sciences 7(3): 233-242.

46. Lee CH, Lee KC, Oon JSH, Ling MHT (2010) Bactome, I: Python in DNA fingerprinting. The Python Papers 5: 6.

47. Lee CH, Oon JS, Lee KC, Ling MHT (2012) Escherichia coli ATCC 8739 adapts to the presence of sodium chloride, monosodium glutamate, and benzoic acid after extended culture. ISRN Microbiol 2012: 965356.

48. Chay ZE, Ling MHT (2010) COPADS, II: Chi-Square test, F-test and t-test routines from Gopal Kanji's 100 statistical tests. The Python Papers Source Codes 2: 3.

49. Chay ZE, Lee CH, Lee KC, Oon JSH, Ling MHT (2010) Russel and Rao coefficient is a suitable substitute for Dice coefficient in studying restriction mapped genetic distances of Escherichia coli. Concept Journal of Computational and Mathematical Biology 1: 1.

50. Loo BZL, Low SXZ, Aw ZQ, Lee KC, Oon JSH, et al. (2014) Escherichia coli ATCC 8739 adapts specifically to sodium chloride, monosodium glutamate, and benzoic acid after prolonged stress. Asia Pacific Journal of Life Sciences 7: 243-258.

51. Low SX, Aw ZQ, Loo BZ, Lee KC, Oon JS, et al. (2013) Viability of Escherichia coli ATCC 8739 in nutrient broth, Luria-Bertani broth and brain heart infusion over 11 weeks. Electron Physician 5(1): 576-581.

52. Goh DJW, How JA, Lim JZR, Ng WC, Oon JSH, et al. (2012) Gradual and step-wise halophilization enables Escherichia coli ATCC 8739 to adapt to 11% NaCl. Electronic Physician 4(3): 527-535.

53. How JA, Lim JZR, Goh DJW, Ng WC, Oon JSH, et al. (2013) Adaptation of Escherichia coli ATCC 8739 to 11% NaCl. Dataset Papers in Biology 2013: 219095.

54. Castillo CFG, Ling MHT (2014) Digital organism simulation environment (DOSE): A library for ecologically-based in silico experimental evolution. Advances in Computer Science: an International Journal 3(1): 44-50.

55. Castillo CFG, Ling MHT (2014) Resistant traits in digital organisms do not revert preselection status despite extended deselection: Implications to microbial antibiotics resistance. BioMed Research International 2014: 648389.

56. Castillo CFG, Chay ZE, Ling MHT (2015) Resistance maintained in digital organisms despite guanine/cytosine-based fitness cost and extended de-selection: Implications to microbial antibiotics resistance. MOJ Proteomics & Bioinformatics 2(2): 00039.

57. Ling MHT (2009) Understanding mouse lactogenesis by transcriptomics and literature analysis. Doctor of Philosophy. Department of Zoology, The University of Melbourne, Australia.

58. Myers JD (2003) Collaborative electronic notebooks as electronic records: Design issues for the secure electronic laboratory notebook (ELN). Simulation Series 35: 13-22.

59. Cock PJA, Antao T, Chang JT, Chapman BA, Cox CJ, et al. (2009) Biopython: freely available Python tools for computational molecular biology and bioinformatics. Bioinformatics 25(11): 1422-1423.

60. Hellegouarch S (2007) CherryPy essentials: Rapid python web application development. Packt Publishing Ltd, UK.

61. Di Pierro M (2009) Web2py: enterprise web framework. (2nd edn), Wiley Publishers, USA.

62. US FDA (1997) Electronic records; Electronic signatures (21 CFR Part 11). Federal Register 62(54): 13430-13466.

63. Dundas J, Ling M (2012) Reference genes for measuring mRNA expression. Theory in Biosciences 131(4): 1-9.

64. Kozera B, Rapacz M (2013) Reference genes in real-time PCR. J Appl Genet 54(4): 391-406.

65. Chapman JR, Waldenström J (2015) With Reference to Reference Genes: A Systematic Review of Endogenous Controls in Gene Expression Studies. PLoS One 10(11): e0141853.

66. Silver N, Best S, Jiang J, Thein SL (2006) Selection of housekeeping genes for gene expression studies in human reticulocytes using real-time PCR. BMC Mol Biol 7: 33.

67. Al-Bader MD, Al-Sarraf HA (2005) Housekeeping gene expression during fetal brain development in the rat—validation by semi-quantitative RT-PCR. Brain Res Dev Brain Res 156(1): 38-45.

68. Andersen CL, Jensen JL, Ørntoft TF (2004) Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. Cancer Res 64(15): 5245-5250.

69. Williams DR, Epperson LE, Li W, Hughes MA, Taylor R, et al. (2005) Seasonally hibernating phenotype assessed through transcript screening. Physiol Genomics 24(1): 13-22.

70. Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, et al. (2002) Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biol 3(7): RESEARCH0034.

71. Pfaffl MW, Tichopad A, Prgomet C, Neuvians TP (2004) Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper--Excel-based tool using pair-wise correlations. Biotechnol Lett 26(6): 509-515.

72. Klie M, Debener T (2011) Identification of superior reference genes for data normalisation of expression studies via quantitative PCR in hybrid roses (Rosa hybrida). BMC Res Notes 4: 518.

73. Ling MH, Ban Y, Wen H, Wang SM, Ge SX (2013) Conserved expression of natural antisense transcripts in mammals. BMC Genomics 14: 243.

74. Comeron JM, Aguadé M (1998) An evaluation of measures of synonymous codon usage bias. J Mol Evol 47(3): 268-274.

75. Jordan IK, Mariño-Ramírez L, Koonin EV (2005) Evolutionary significance of gene expression divergence. Gene 345(1): 119-126.

76. Dobzhansky T (1973) Nothing in biology makes sense except in the light of evolution. The American Biology Teacher 35: 125-129.

77. Travisano M, Lenski RE (1996) Long-term experimental evolution in Escherichia coli. IV. Targets of selection and the specificity of adaptation. Genetics 143(1): 15-26.

78. Travisano M (1997) Long-term experimental evolution in Escherichia coli. VI. Environmental constraints on adaptation and divergence. Genetics 146(2): 471-479.

79. Cooper VS (2002) Long-term experimental evolution in Escherichia coli. X. Quantifying the fundamental and realized niche. BMC Evol Biol 2: 12.

80. Crozat E, Philippe N, Lenski RE, Geiselmann J, Schneider D (2005) Long-term experimental evolution in Escherichia coli. XII. DNA topology as a key target of selection. Genetics 169(2): 523-532.

81. Bibbal D, Dupouy V, Prère MF, Toutain PL, Bousquet-Mélou A (2009) Relatedness of Escherichia coli strains with different susceptibility phenotypes isolated from swine feces during ampicillin treatment. Appl Environ Microbiol 75(10): 2999-3006.

82. Bibbal D, Dupouy V, Ferré JP, Toutain PL, Fayet O, et al. (2007) Impact of three ampicillin dosage regimens on selection of ampicillin resistance in Enterobacteriaceae and excretion of blaTEM genes in swine feces. Appl Environ Microbiol 73(15): 4785-4790.

83. Furtula V, Farrell EG, Diarrassouba F, Rempel H, Pritchard J, et al. (2010) Veterinary pharmaceuticals and antibiotic resistance of Escherichia coli isolates in poultry litter from commercial farms and controlled feeding trials. Poult Sci 89(1): 180-188.

84. Cuny C, Lesbats M, Dukan S (2007) Induction of a global stress response during the first step of Escherichia coli plate growth. Appl Environ Microbiol 73(3): 885-889.

85. Gou N, Onnis-Hayden A, Gu AZ (2010) Mechanistic toxicity assessment of nanomaterials by whole-cell-array stress genes expression analysis. Environmental Science and Technology 44(15): 5964-5970.

86. Ling MHT (2010) COPADS, I: Distances measures between two lists or sets. The Python Papers Source Codes 2: 2.

87. Nei M, Li WH (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc Natl Acad Sci U S A 76(10): 5269-5273.

88. Dice LR (1945) Measures of the amount of ecologic association between species. Ecology 26: 6.

89. Fallaw WC (1979) A test of the Simpson coefficient and other binary coefficients of faunal similarity. Journal of Paleontology 53(4): 1029.

90. Russel PF, Rao TR (1940) On habitat and association of species of anopheline larvae in south-eastern Madras. Journal of the Malaria Institute of India 3: 154-178.

91. Sezonov G, Joseleau-Petit D, D'Ari R (2007) Escherichia coli physiology in Luria-Bertani broth. Journal of Bacteriology 189(23): 8746-8749.

92. Doudoroff M (1940) Experiments on the adaptation of Escherichia coli to sodium chloride. J Gen Physiol 23(5): 585.

93. Jay JMM (1992) E. coli gastroenterritis syndromes. Modern Food Microbiology 4: 570-575.

94. Ochiai T (1999) Salt-sensitive growth of Staphylococcus aureus: Stimulation of salt-induced autolysis by multiple environmental factors. Microbiol Immunol 43(7): 705-709.

95. Wilke CO, Adami C (2002) The biology of digital organisms. Trends in Ecology & Evolution 17(11): 528-532.

96. Langton CG (1986) Studying artificial life with cellular automata. Physica D: Nonlinear Phenomena 2291-3): 120-149.

97. Ray TS (1992) An approach to the synthesis of life. In Langton CG. et al. (Eds.), Proceedings of Artificial Life II, pp. 371-408.

98. Adami C (2006) Digital genetics: unravelling the genetic basis of evolution. Nature Review Genetics 7: 109-118.

99. Elena SF, Sanjuan R (2008) The effect of genetic robustness on evolvability in digital organisms. BMC Evol Biol 8: 284.

100. Goldsby HJ, Dornhaus A, Kerr B, Ofria C (2012) Task-switching costs promote the evolution of division of labor and shifts in individuality. Proceedings of the National Academy of Sciences of the United States of America 109(34): 13686-13691.

101. Lenski RE, Ofria C, Pennock RT, Adami C (2003) The evolutionary origin of complex features. Nature 423(6936): 139-144.

102. Nelson CW, Sanford JC (2011) The effects of low-impact mutations in digital organisms. Theoretical Biology & Medical Modelling 8: 9.

103. Ofria C, Huang W, Torng E (2008) On the gradual evolution of complexity and the sudden emergence of complex features. Artif Life 14(3): 255-263.

104. Lenski R (2003) The ecology, genetics and evolution of bacteria in an experimental setting. Curr Biol 13(12): R466-467.

105. Ling MHT (2014) Applications of artificial life and digital organisms in the study of genetic evolution. Advances in Computer Science: an international journal 3(4): 107-112.

106. Hamblin S (2013) On the practical usage of genetic algorithms in ecology and evolution. Methods in Ecology and Evolution 4(2): 184-194.

107. Ling MHT (2012) An artificial life simulation library based on genetic algorithm, 3-character genetic code and biological hierarchy. The Python Papers 7: 5.

108. Ling MHT (2012) Ragaraja 1.0: The genome interpreter of Digital Organism Simulation Environment (DOSE). The Python Papers Source Codes 4: 2.

109. Raymond L, Plantegenest M, Vialatte A (2013) Migration and dispersal may drive to high genetic variation and significant genetic mixing: the case of two agriculturally important, continental hoverflies (Episyrphus balteatus and Sphaerophoria scripta). Molecular Ecology 22(21): 5329-5339.

110. Relethford JH (1988) Heterogeneity of long-distance migration in studies of genetic structure. Ann Hum Biol 15(1): 55-63.

111. Balkhair A (2017) Antibiotics resistance: When the armamentarium gets to the verge of being empty. Oman Med J 32(4): 267-268.

112. Guillemot D, Varon E, Bernede C, Weber P, Henriet L, et al. (2005) Reduction of antibiotic use in the community reduces the rate of colonization with penicillin G-nonsusceptible Streptococcus pneumonia. Clin Infect Dis 41(7): 930-938.

113. Johnsen PJ, Townsend JP, Bohn T, Simonsen GS, Sundsfjord A, et al. (2009) Factors affecting the reversal of antimicrobial-drug resistance. The Lancet Infect Dis 9(6): 357-364.

114. Enne VI, Livermore DM, Stephens P, Hall LM (2001) Persistence of sulphonamide resistance in Escherichia coli in the UK despite national prescribing restriction. Lancet 357(9265): 1325-1328.

115. Knight GM, Budd EL, Lindsay JA (2013) Large mobile genetic elements carrying resistance genes that do not confer a fitness burden in healthcare-associated meticillin-resistant Staphylococcus aureus. Microbiology 159: 1661-1672.

116. Allen TD, Poteet ML, Russell JE, Dobbins GH (1997) A field study of factors related to supervisors' willingness to mentor others. Journal of Vocational Behavior 50(1): 1-22.

117. Ragins BR, Scandura TA (1999) Burden or Blessing? Expected Costs and Benefits of Being a Mentor. Journal of Organizational Behavior 20(4): 493-509.

118. Janssen S, van Vuuren M, de Jong MDT (2014) Motives to mentor: Self-focused, protégé-focused, relationship-focused, organization-focused, and unfocused motives. Journal of Vocational Behavior 85(3): 266-275.

119. Fink L, Drake JE (2016) Mood and Flow. Empirical Studies of the Arts 34(2): 177-192.

120. Hibbin R (2016) The psychosocial benefits of oral storytelling in school: developing identity and empathy through narrative. Pastoral Care in Education 34(4): 218-231.

121. Pennebaker JW (2000) Telling stories: the health benefits of narrative. Lit Med 19(1): 3-18.

122. Pennebaker JW, Seagal JD (1999) Forming a story: the health benefits of narrative. J Clin Psychol 55(10): 1243-1254.

123. Morales DX, Grineski SE, Collins TW (2017) Faculty Motivation to Mentor Students Through Undergraduate Research Programs: A Study of Enabling and Constraining Factors. Research in Higher Education 58(5): 520-544.

124. Agliata AK, Renk K (2007) College Students' Adjustment: The Role of Parent-College Student Expectation Discrepancies and Communication Reciprocity. Journal of Youth and Adolescence 37: 967.

125. Agliata AK, Renk K (2008) College Students' Affective Distress: The Role of Expectation Discrepancies and Communication. Journal of Child and Family Studies 18: 396.

126. Erdem F, Aytemur JÖ (2008) Mentoring-A relationship based on trust: Qualitative research. Public Personnel Management 37(1): 55-65.

127. Fleig Palmer MM, Schoorman FD (2011) Trust as a moderator of the relationship between mentoring and knowledge transfer. Journal of Leadership & Organizational Studies 18(3): 334-343.

128. Kabdulovaa E, Öztürk EB (2017) The Relationship between Mentoring Functions Received and Psychological Empowerment: The Role of Trust and Political Skill. The Journal of Human and Work 4(1): 47-60.

129. Ghosh R, Reio TG (2013) Career benefits associated with mentoring for mentors: A meta-analysis. Journal of Vocational Behavior 83(1): 106-116.

130. Reddick R, Griffin K, Cherwitz R, Cérda Pražák A, Bunch N (2012) What you get when you give: how graduate students benefit from serving as mentors. The Journal of Faculty Development 26(1): 37-49.