



Engaging the next generation of plant geneticists through sustained research: an overview of a post-16 project

Jon Michael Hale ¹

Received: 19 January 2020 / Revised: 8 September 2020 / Accepted: 8 September 2020 / Published online: 17 September 2020
© The Genetics Society 2020

Abstract

Student career aspirations are directly linked to the careers that they are exposed to and the esteem that they are given in society. Where schools are located in areas with low visibility of scientific careers this will have an impact on student aspirations. This project is demonstrating that aspirations can be altered by engaging 16–18-year-old A level biologists in sustained research. A total of 20 students from schools across Jersey are attempting to sequence the chloroplast genomes from daffodils that they have collected from non-cultivated locations using Oxford Nanopore Technologies' MinION. Despite site closures due to COVID-19 control measures, the project has developed insight into different scientific careers through experience and ownership of the entire project pipeline. This project demonstrates an opportunity for schools and academics to collaborate to further science and potentially improve student outcomes.

Introduction

The landscape of biology education in England

Biology education in England has changed significantly since 2015 when both the GCSE and A level specifications were reformed. There is more of an emphasis on non-vertebrate biology, with topics such as plant diseases being introduced. However, with the increased content, and pressures on maximising student performance in examinations, it can be argued that the experience of being a scientist is being reduced to knowledge retention.

Unfortunately, educational research has driven this change. Implementation of “test-enhanced learning” (Roe-diger and Karpicke 2006) into UK classrooms has boomed in recent years as teachers become more comfortable with evidence-based decisions, thanks to the Education

Endowment Foundation (Holman and Yeomans 2018) and a wealth of other accessible resources (Busch and Watson 2019; Jones 2019). This shift towards repeated testing, where success is intrinsically rewarding, does little to improve visibility of different STEM careers. A short survey (Supplementary File S1) within our A level biology cohort in 2018 showed that 50% did not see how the topics they were learning about applied to real life, with 78% having no exposure to different careers in lessons ($n = 18$).

The impact of exposure to different careers has been researched through the ASPIRES project (Archer et al. 2020), which emphasises the role teachers have in shaping the aspirations of their students. Three main factors were identified that affect whether a young person wishes to pursue a career in STEM: science capital, which relates to the knowledge, experiences and social contacts an individual may have; educational and social representations of science, whereby it is often portrayed as a career for the gifted and educational practices where aspirations were seen to be undermined by the curriculum experience and careers delivery. The Gatsby Charitable Foundation commissioned research into how careers guidance should look like in schools, leading to eight benchmarks of good practice (Holman 2014). These include encounters of the workplace and of employers, which is particularly challenging for schools in remote locations.

Another aspect of the changing curricula that the Gatsby Charitable Foundation has researched is the role of practicals

Associate editor: Lindsey Compton

Supplementary information The online version of this article (<https://doi.org/10.1038/s41437-020-00370-0>) contains supplementary material, which is available to authorized users.

✉ Jon Michael Hale
halej@beaulieu.jersey.sch.uk

¹ Beaulieu Convent School, Wellington Road, St. Helier JE2 4RJ, Jersey

within science education (Holman 2017). The reforms at GCSE and A level have made some practicals “required” whereby every student should complete them during their studies. This was grounded in a 2011 survey of Russell Group universities, in which 97% of laboratory managers deemed their first year undergraduates to have inadequate practical skills (Grant 2011). Within the A level biology curriculum, students only have to complete 12 practicals. When this could include using Biuret’s in solution to test for protein, it is understandable that the desired effect of improving competency may not be achieved. Despite the backing of learned societies, many evidence-based teachers question the value of practical work within science, learning with educational research often not showing a positive effect on knowledge uptake (Abrahams and Millar 2008). Many teachers are therefore reluctant to commit time and money in delivering practical opportunities within the GCSE and A level cohorts for little, if any, gain in examination performance.

However, excelling in subjects should mean more than attainment, and perhaps the soft skills of science need to be given greater emphasis. Bennett et al. (2016) completed a review of independent research projects in science including a number of international case studies. Through interviews they were able to identify the soft skills that are developed in the students, stating the development of self-esteem, independence and autonomy, self-regulation, tenacity, time management skills, a spirit of co-enquiry with teachers and a sense of scientific identity as additional benefits. With the number of students not continuing in higher education after their first year increasing in recent years (HESA 2018), perhaps development of these skills needs to be given higher priority.

Chloroplasts, their genomes and solving the sequences

Chloroplasts are often seen as the archetypal plant organelle. Instantly recognisable by secondary school-aged students, these organelles contain their own DNA genome called a plastome. The plastome is thought to occur in either circular or linear forms (Bendich 2014) containing a number of genes responsible for photosynthesis and their own housekeeping functions, such as protein synthesis, as shown with the plastome map of the daffodil *Narcissus poeticus* in Supplementary Fig. S1 (Könyves et al. 2018). With incredibly large numbers of chloroplasts within photosynthetic cells, this plastome offers a valuable resource for phylogenetic studies as shown by the sheer number of angiosperm plastome sequences now available in the NCBI GenBank. The high number of copies of the chloroplast within each cell potentially offers biotechnological importance exceeding nuclear DNA manipulation, from enhancing stress tolerance to producing vaccines (Daniell et al. 2016).

The first plastome sequence was published in 1986 coming from *Nicotiana tabacum* (Shinozaki et al. 1986). However,

the first daffodil plastome was only published in 2018 (Könyves et al. 2018). This complete plastome from *Narcissus poeticus* (accession MH706763) was completed using Illumina HiSeq 4000 at the Oxford Genomics Centre, showing that the total plastome sequence was 160,099 base pairs (bp) with two inverted repeat (IR) regions, each of 28,610 bp length. It is these repeat regions that potentially play to the strength of Oxford Nanopore Technologies’ (ONT) MinION in obtaining long reads. The near identical nature of these repeat regions makes synthesis of contigs difficult and therefore assembly challenging. If contigs entirely span the IR regions, then assembly will be much easier. This relies on obtaining relatively high molecular weight DNA during the library preparation. Bethune et al. (2019) were able to obtain plastome sequence reads of up to 25,828 bp plastome sequence using ONT MinION creating contigs of up to 125,727 bp showing that it can be achieved. The quality of ONT MinION has been validated by sequencing the complete *Arabidopsis thaliana* genome on a single flow cell (Michael et al. 2018). However, it does remain that the raw error rate is relatively high with ONT MinION, but it has been shown that these errors are effectively masked during phylogenetic analysis (Scheunert et al. 2019).

With classification within botany often relying on the observable characteristics, and therefore the skill and analysis experience of the taxonomist, the objective plastome sequence offers a different insight into the evolutionary relationship of the *Narcissus* genus (Könyves et al. 2019). Although currently only two complete plastome sequences exist within the subfamily of *Amaryllidoideae*, the concept of using these sequences for phylogenetic analysis has been demonstrated in different clades (Melton et al. 2015; Daniell et al. 2016; Manzanilla et al. 2018). The very concept of evolutionary trees to show phylogeny is taken as the norm in communicating just how related different species are. Although they have been used since before Charles Darwin sketched his first in 1837, they nevertheless continue to present many misconceptions (Gregory 2008) based upon the reading order of clades, interpreting divergence events or identifying shared phylogeny.

The use of ONT MinION within teaching and learning is continuing to grow, from the in situ use for informing biodiversity studies in Peru (Watsa et al. 2020) to more classical bioinformatics classes within university settings (Salazar et al. 2020). Fundamentally, this project is aiming to take genetics out of the textbook and give students a thorough understanding of the science behind the research.

Project background

The aim of this open-ended project is to improve the engagement of post-16 students with science through long-term scientific research sequencing the chloroplast genomes of *Narcissus* specimens.

Beaulieu Convent School is located in Jersey, Channel Islands. It is one of the five relatively small schools on the Island that offers 16–18 education. In the past, there have been very few students that have left the school to study a traditional science subject at university, though many have chosen to study healthcare or medicine. Students are rarely innately passionate about plant biology when they start A level biology. By collaborating with Dr Könyves at the Royal Horticultural Society, Wisley, the school is attempting to address these issues by raising the visibility of science within the community.

Daffodils are ubiquitous in Jersey throughout spring. However, they are incredibly diverse in their characteristics within metres of each other. Through discussion with Dr Könyves it was proposed that students could carry out the necessary steps to collect daffodil samples and determine their plastome sequences, with the ambition of identifying phylogenetic relationships. Beyond the scientific rationale of obtaining plastome sequences to study the phylogeny of an understudied genus, this offered a curriculum link to photosynthesis as well as to gene technologies.

It is a commonly held misconception that the only role of a chloroplast is in photosynthesis. Only in year 13 do students acknowledge that this process requires more than chlorophyll to occur, but little is explored beyond the light-dependent stage and the Calvin cycle, with the plastome given a cursory mention of its existence. Assembling the plastome sequence also offers the opportunity to engage with the practical aspects of bioinformatics, which is another aspect of the A level biology specification that is mentioned without application. Providing assistance with this aspect of the project has been Dr Harkess who has developed chloroplast genome assembly procedures using ONT MinION data for his postgraduate classes (Harkess 2019) and is able to draw upon his experiences to negate issues with ONT MinION data (Li and Harkess 2018).

In summary, this student-led research project blends fieldwork with lab work to ensure that each participant has equal access to every discipline with the study in order to maximise their research experience.

Methods and materials

Participant recruitment

Schools offering post-16 biology and/or chemistry ($n = 5$) in Jersey were invited to participate in the project with Beaulieu Convent School hosting all the after-school sessions (~75 min per week). Hautlieu School provided the most external participants ($n = 4$), with no students turned away from the project ($n = 20$). Although it was not the intention, all of the students were female within this year's

cohort. Although the project is open to students in years 12 and 13, collaborating schools only nominated year 12 students. Each of the students was able to complete fieldwork and DNA extraction techniques, however only one student was able to complete the library preparation and flow cell loading.

Literature review

Students were introduced to scientific journals and how to find open access papers, and approach authors professionally. Students then worked collaboratively to research relevant reviews and methods, discussing their findings and summaries on a monthly basis.

Sample collection

Students and staff collected two samples from non-cultivated land, such as roadside verges, recording the coordinates of the samples. Each sample consisted of a flower, stem and a leaf. A voucher was produced for each sample, recording the flower type (cup, trumpet or rosette), colour of the flower components and leaves. Students selected leaf samples from which to extract the DNA. Students decided which specimens to use through debate rather than objective criteria.

DNA extraction and sequencing

Students extracted the DNA using the Qiagen DNAeasy kit (Qiagen, Manchester, UK). To date only one daffodil (Supplementary Fig. S2) has been sequenced using the Rapid DNA Sequencing Kit SQK-RAD004 (ONT, Oxford, UK) and the Flow Cell (R9.4.1, FLO-MIN106D) on the MinION device (ONT, Oxford, UK).

Sequence data analysis

Basecalling was undertaken using Guppy v3.6.0 using default parameters. To sample the quality of reads, students used NCBI BLASTn to identify the approximate percentage of plastome reads within the sample and probable error rate by comparing plastome reads against the *N. poeticus* (MH706763) alignment. Once basecalling is complete, students will use wtdbg2 (Ruan and Li 2019) to assemble contigs and the plastome.

Project collaborators

Dr Kálmán Könyves (Royal Horticultural Society, Wisley) was recruited via the Wellcome Genome Campus's Education Lead, Francesca Gale. Dr Könyves has been pivotal in the initial project plan, acting as the designated STEM

partner for the Royal Society Partnership Grant application. Since then, he has shared interesting papers to discuss with the students and sent consumables for the DNA extraction such as pestles and silica sand. Dr Alex Harkess (Hudson Alpha, Alabama) was recruited via social media. He has been incredibly supportive in building the necessary bioinformatics pipelines, providing walkthroughs via weekly video calls.

Project evaluation

The impact of the project on students will be assessed using pre- and post-project questionnaires (Supplementary Files S1 and S2). Changes in attitudes towards STEM careers, student aspirations and how they perceive any changes to their soft skills will be identified. Any improvement in knowledge will be indirectly assessed by comparing trends in A level examination results between cohorts prior to 2018 and the project cohorts. Unfortunately it may not be possible to share examination performances of students outside of Beaulieu Convent School.

Project outreach

Four students (years 12 and 13, all from Beaulieu Convent School) were invited to Carlton House, London, to present their project plan to other students, members of the public and Fellows of the Royal Society in November 2019 as part of a student poster conference (Supplementary Fig. S3). One other student (year 12, Beaulieu Convent School) presented a similar poster at the Babraham Institute's Sixth Form Conference in December 2019. Six students (two from Beaulieu Convent School, four from Hautlieu School, all year 12) were invited to present at the Royal Society's Summer Science Exhibition as part of the inaugural Young Researchers Zone over the first 2 days of the exhibition. Unfortunately this event was cancelled due to the COVID-19 control measures.

Results

Developing the personal skills of students

Despite the project being heavily affected by the COVID-19 site closures in schools, early signs of success are now clear. At least two students are now considering applications to higher education to study plant biology. From initial demonstration of micropipetting, students are competent in this skill and therefore able to follow the protocols for DNA extraction and library preparation independently and with ease. As students' confidence has grown with the techniques of the project, there has been a growth of peer-to-peer

coaching ensuring that each student develops competency across all practical aspects.

During the Royal Society Student Conference, students had to defend the methodology of the project to staff and Fellows of the Royal Society. Removed from any immediate support from teachers, the students were treated as independent and professional scientists. Although they found the experience daunting at first, their collective confidence increased, with students increasingly identifying themselves as scientists carrying out research into the unknown. By the time the doors opened to the public, each and every member of the presentation team was conveying the science of nanopore sequencing, their expected results and most importantly, their excitement for science. Upon their return, they inspired a fellow student to present a poster at the Babraham Sixth Form Conference evening event, at which posters and presentations were judged by a panel representing the institute. Despite only having 3 months of A level Biology knowledge after studying a combined science GCSE, the student was awarded second place behind a year 13 student from a highly performing sixth form college in Cambridge. Upon her return, the student was adamant that she wanted a career in science as a result of this experience and having mixed with PhD students and researchers from the institute.

Narcissus sequence data analysis

A single daffodil sample (Supplementary Fig. S2) has been sequenced on the MinION device. Students have analysed a sample of 20,000 reads from a total of 800,000 basecalled reads against the *N. poeticus* reference sequence reporting an alignment rate of a little over 2%. Dr Harkess has independently confirmed this proportion, which he has also observed within his own labs. Along with relevant genomic data there are sequences that are likely to come from fungal and bacterial contamination at the time of sample collection. Of the samples that aligned with the reference sequence, there was a >95% similarity, with differences that could result from the error rate in basecalling.

Discussion

This project is already showing that aspirations can be raised by allowing students to engage with sustained research. As students have developed competency with the practical tasks, such as voucher preparation for the herbarium, they have been informally discussing their aspirations with each other, giving us a relatively accurate view of the changes observed so far. By analysing these observations and informal discussions, it is clear that collectively the students not only identify as scientists, but also through

interacting with researchers they have changed their pre-conceptions of scientists, with students describing them as “friendly”, “generous”, and “young”. A greater proportion of students from this project are considering careers in science compared to previous cohorts. Those students who go on to study plant genetics will already have an appreciation of this rapidly progressing field, which should make the transition to university easier. It may well prove that students have a better understanding of genetics, bioinformatics and phylogeny, which will be quantified through question level analysis of the examinations in 2021. However, when considering the financial and time investment, few schools would consider this to be cost-effective for any improvement in student examination outcomes.

This experience gives the students a flavour of research work, not only the excitement of collecting data that is entirely new to science (rather than simply reproducing trends from the textbook), but also the experience of the more “day-to-day” repetitive tasks, emphasising the need for skills such as time management and self-motivation.

Allowing students to become competent in a range of skills over a longer period of time than would be achievable within the A level curriculum has shifted the status quo within the group. Students are much more independent, delegating voucher preparation to others while DNA extraction from a 100 mg sample of a leaf is completed, to ensure that the specimen has been fully processed within the after-school session. They have shown a great deal of empathy as well as negotiation skills to ensure that each student has been able to complete every step of the protocols in a calm and supportive environment.

Undoubtedly, the positive teamwork was rapidly accelerated by the Royal Society Student Conference. The conference took place only a couple of months into the project, before any daffodils had been identified. Although only four students attended the conference, every member of the team contributed to the poster and to the knowledge base of those attending through discussing their literature reviews and developing a method. For example, the misconception that chloroplasts only photosynthesise was dispelled as students discussed the role of genes within the plastome following analysis of the *N. poeticus* plastome using OGDRAW, leading to students gaining an appreciation of other plastids not mentioned in their course. The intense scrutiny from Fellows of the Royal Society received by the four attendees had a profound effect on them. They became less accepting of concepts in lessons as they assessed the experimental evidence rather than simply rote learning processes. This is expected to have a positive impact on their application skills across the entire A level specification as well as stretching those around them to understand the “why”, rather than the “what” of biological processes.

Once the dry lab section of the project has concluded the qualitative and quantitative impact on students’ soft skills

and aspirations will be explored in depth using the post-project survey. Ideally, this would be completed before students make applications for employment or university to enable greater reflection and targeted guidance.

With respect to the sequencing data, students will next attempt to complete a de novo assembly of the first daffodil sample (Supplementary Fig. S2). If successful, students will be able to compare this sequence to the reference *N. poeticus* plastome sequence, which may identify regions of variation and allow phylogenetic analysis. One benefit of the temporary pause in the project due to the COVID-19 control measures will be the opportunity for overlapping cohorts within the project allowing peer-to-peer teaching of procedures. With the space created by the departure of the year 13 students, the project will again be opened to all of the post-16 biology and chemistry students on the Island.

Conclusions

Despite the relatively low-cost of this project, it would still be out of reach of most UK schools without financial assistance. The loan of a MinION and Flongle adaptor alongside micropipettes would dramatically reduce the cost to schools, with an outlay of ~£150 per sequencing run. If this was to be administered by regional universities it may help support links with higher education for remote schools. With straightforward DNA extraction protocols, schools could explore flora that is locally significant to them and improve the science capital of each student. If schools intended on using a Flongle flow cell, it may be necessary to isolate chloroplasts before DNA extraction to mitigate the lower data output in comparison with a MinION flow cell.

Observing the impact on individual students, it is clear that project-based opportunities can have life-changing effects on the aspirations of young people. The informal learning environment has been particularly beneficial to those students with lower prior attainment levels and/or less confident students. It may well be the sense of co-enquiry that levels the playing field with their peers as each and every student’s contribution is equally valued. For this reason alone, we propose that larger and sustained collaborative investigations should be part of the next A level biology curriculum reform.

Acknowledgements This project has been funded by the Royal Society through a Partnership Grant in collaboration with Dr Kálmán Könyves of the Royal Horticultural Society, Wisley. We are indebted to the bioinformatics assistance provided by Dr Alex Harkess of Hudson Alpha, Huntsville, Alabama.

Compliance with ethical standards

Conflict of interest The author declares no conflict of interest.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

- Abrahams I, Millar R (2008) Does practical work really work? A study of the effectiveness of practical work as a teaching and learning method in school science. *Int J Sci Educ* 30(14):1945–1969
- Archer L, Moote J, MacLeod E, Francis B, De Witt J (2020) ASPIRES 2: young people's science and career aspirations, age 10–19. UCL Institute of Education, London
- Bendich AJ (2014) Circular chloroplast chromosomes: the grand illusion. *Plant Cell* 16:1661–1666
- Bennett J, Dunlop L, Knox KJ, Reiss MJ, Jenkins RT (2016) A rapid evidence review of practical independent research projects in science. Wellcome Trust, London
- Bethune K, Mariac C, Couderc M, Scarcelli N, Santoni S, Ardisson M, Martin J-F et al. (2019) Long-fragment targeted capture for long-read sequencing of plastomes. *Appl Plant Sci* 5:e1243
- Busch B, Watson E (2019) *The science of learning: 77 studies that every teacher needs to know*. Routledge, Oxon
- Daniell H, Lin C, Yu M (2016) Chloroplast genomes: diversity, evolution, and applications in genetic engineering. *Genome Biol* 17:134
- Grant L (2011) *Practical skills of new undergraduates*. Russell Group survey. The Gatsby Charitable Foundation, London
- Gregory TR (2008) Understanding evolutionary trees. *Evolution* 1:121–137
- Harkess A (2019) Mini maize class. Github. https://github.com/alexharkess/mini_maize_class. Accessed 3 Jan 2020
- HESA (2018) Non-continuation summary: UK performance indicators 2016/17. HESA. <https://www.hesa.ac.uk/news/08-03-2018/non-continuation-summary>. Accessed 16 Jan 2020
- Holman J (2014) *Good career guidance—a handbook for secondary schools*. The Gatsby Charitable Foundation, London
- Holman J (2017) *Good practical science*. The Gatsby Charitable Foundation, London
- Holman J, Yeomans E (2018) *Improving secondary science—a guidance report*. The Education Endowment Foundation, London
- Jones K (2019) *Retrieval practice*. John Catt Educational, Woodbridge
- Könyves K, Bilsborrow J, David J, Culham A (2018) The complete chloroplast genome of *Narcissus poeticus* L. (Amaryllidaceae: Amaryllidoideae). *Mitochondrial DNA Part B* 3(2):1137–1138
- Könyves K, David J, Culham A (2019) Jumping through the hoops: the challenges of daffodil (*Narcissus*) classification. *Bot J Linn Soc* 190(4):389–404
- Li F-W, Harkess A (2018) A guide to sequence your favorite plant genomes. *Appl Plant Sci* 6(3):e1030
- Manzanilla V, Kool A, Nguyen Nhat L, Nong Van H, Le Thi Thu H, de Boer HJ (2018) Phylogenomics and barcoding of the *Panax*: towards the identification of the ginseng species. *BMC Evol Biol* 18:44
- Melton III JT, Leliaert F, Tronholm A, Lopez-Bautista (2015) The complete chloroplast and mitochondrial genomes of the green macroalga *Ulva* sp. UNA00071828 (*Ulvophyceae, Chlorophyta*). *PLoS ONE* 10(4):e0121020
- Michael TP, Jupe F, Bemm F, Motley ST, Sandoval JP, Lanz C et al. (2018) High contiguity *Arabidopsis thaliana* genome assembly with a single nanopore flow cell. *Nat Commun* 9:541
- Roediger HL, Karpicke JD (2006) Test-enhanced learning: taking memory tests improves long-term retention. *Psychol Sci* 17(3):249–255
- Ruan J, Li H (2019) Fast and accurate long-read assembly with wtdbg2. *Nat Methods*. <https://doi.org/10.1038/s41592-019-0669-3>
- Salazar AN, Nobrega FL, Anyansi C, Aparicio-Maldonado C, Costa AR, Haagsma AC et al. (2020) An educational guide for nanopore sequencing in the classroom. *PLoS Comput Biol* 16(1):e1007314
- Scheunert A, Dorfner M, Lingl T, Oberprieler C (2019) Nanopore data for plastome sequencing. <https://www.biorxiv.org/content/10.1101/855981v1>. Accessed 27 Nov 2019
- Shinozaki K, Ohme M, Tanaka M, Wakasugi T, Hayashida N, Matsumabayashi T, Zaita N (1986) The complete nucleotide sequence of the tobacco chloroplast genome: its gene organization and expression. *EMBO J* 5:2043–2049
- Watsa M, Erkenwick GA, Pomerantz A, Prost S (2020) Portable sequencing as a teaching tool in conservation and biodiversity research. *PLoS Biol* 18(4):e3000667