Agrobacterium mediated CRISPR Cas9 genetic engineering of *Solanum tuberosum*

Thesis presented by
Dylan Barrett, BSc

For the degree of
Master of research

Student number: **114446538**

University College Cork
School of Biological, Earth and Environmental Science

Supervisors: Dr. Barbara Doyle Prestwich and Dr. Eoin Lettice

2020

Contents

"*This is to certify that the work I am submitting is my own and has not been submitted for another degree, either at University College Cork or elsewhere. All external references and sources are clearly acknowledged and identified within the contents. I have read and understood the regulations of University College Cork concerning plagiarism."*

1. Abstract

Global populations are continuing to rise and so too does the strain on the food production and agriculture industry. We are constantly in pursuit of new sustainable ways to feed the ever-increasing population. Genetic engineering of crop plants may prove to be a crucial tool in providing healthy sustainable food sources. The main aim of the project was to create an INDEL in The Rhamnose:beta-solanine/beta-chaconine rhamnosyltransferase (*SGT3*) gene region in *Solanum tuberosum* using *Agrobacterium* mediated CRISPR Cas9 genetic engineering. Sterol glycoside transferase (*SGT*) genes are involved in the biosynthetic pathway of glycoalkaloids in potatoes. This is a food safety issue which can result in reduced yields. In addition, the goal was to investigate the role of pre-transformation treatments on transformation efficiency. To that end, the effects of different wavelengths of light produced by light-emitting diodes on potatoes were examined in order to design a pre-transformation treatment. Red light and combinations of red and blue light were shown to have a significant positive impact on plant growth rate. Eight bacterial isolates from the soil rhizosphere were successfully isolated, identified and the volatile organic compounds produced by these bacterial isolates were measured in order to design a pre-transformation treatment. Six of the isolates were shown to produce known growth promoters through gas-chromatography mass-spectrometry and four of the isolates had a significant positive impact on plant growth rate in co-cultivation experiments. The role of heat shock treatment in transformation efficiency was also investigated. LED treatment with red-blue light (ratio of 3:1) and heat shock treatments were shown to increase *Agrobacterium* transformation efficiency in gus histochemical staining experiments. Finally, this project aimed to design guide RNA to target the *SGT3* gene and transform *Agrobacterium* with the CRISPR vectors in order to successfully create a knockout in *Solanum tuberosum.* Three guide RNAs to target the *SGT3* DNA were successfully designed and validated using an *in vitro* cleavage of *SGT3* DNA by CRISPR/Cas9 ribonucleoprotein complex, however, it is unclear whether an *in vivo* knockout has been created due to the activity of the *Agrobacterium*-mediated CRISPR Cas9 transformation.

2. Introduction

Due to an ever-increasing world population, genetic engineering of crops is now more important than ever. With world populations estimated to reach 10 billion by 2050, food security will become a major hurdle to overcome. It is estimated that food production will need to see an increase of 60-100 % in order to feed a population of this size (FAOSTAT, 2016). The ever-growing population combined with an increased level of biotic and abiotic stresses due to climate change and a decreasing amount of arable land available will all have to be combated against by new and improved crop production techniques. Genetic engineering could prove to be a vital tool to combat these mounting issues.

2.1 CRISPR genome engineering

Genome engineering refers to the altering of a target genome to produce a desired phenotypic trait in the target organism. The Clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR associated protein (Cas9) system is at the forefront of genome editing today. CRISPR is a type of adaptive immune system found in some prokaryotes and was first discovered by Yoshizumi Ishino (Ishino et al., 1987). Its function was later studied by Francisco Mojica (Mojica & Rodriguez-Valera, 2016). However, it was Jennifer Doudna and Emmanuelle Charpentier who first recognised the function of the CRISPR/Cas9 system as a tool for genetic engineering (Doudna & Charpentier, 2014).

CRISPR acts as an adaptive immune system in Archaea and some Bacteria, as shown in figure 1 below, and was first shown to be essential for defence against pathogens in *Streptococcus thermophiles* (Barrangou et al., 2007). The CRISPR system defends the host organism in three stages. Firstly, the CRISPR defence system adapts to the virus by incorporating foreign viral DNA (or RNA) into the CRISPR array. This allows for the CRISPR system to obtain sequence memory to defend against a similar infection in the future. This stage may often require the recognition of a Protospacer adjacent motif (PAM) sequence. Secondly, the CRISPR array is transcribed producing a precursor transcript (pre-crRNA) that matures to form CRISPR RNA (crRNA). The crRNA maturation is activated by trans-activating (tracrRNA) and both the pre-crRNA

and tracrRNA are cleaved by endogenous RNAase III (Deltcheva et al., 2011). Finally, the crRNA, along with the Cas proteins, cleave the DNA of the invading pathogen in a site-specific manner (Makarova et al., 2015). Cas9 is a large multifunctional protein and is considered a type II effector system. Cas9 contains a HNH nuclease domain and a RuvC like domain and forms a tracrRNA:crRNA:Cas9 complex. The HNH domain will cleave a strand of DNA complementary to the crRNA and the RuvC domain will cleave the strand opposite the complementary strand forming site-specific double-stranded DNA breaks. These double-strand breaks will be repaired either by non-homologous end joining (NHEJ) or homology-directed repair (HDR). The homologous recombination process, part of the HDR system, can be taken advantage of to insert donor DNA into the break site, adding in a short sequence that did not previously exist at this site as shown in figure 2. The non-homologous end-joining repair system is an error-prone process that rapidly ligates DNA causing the loss or acquisition of base pairs. The tracrRNA:crRNA system was engineered to form single guide RNA (sgRNA) which remains an approximately 20nt sequence and can bind to the Cas9 protein. The sgRNA acts as a guide to bring the Cas9 to the target site (Doudna & Charpentier, 2014).

The CRISPR system is not only capable of cutting viral DNA, but the DNA of many living organisms and as such is a powerful tool in gene editing. CRISPR has several advantages over other genome-editing nucleases such as Zinc finger nucleases (ZFNs) and transcription activator-like effector nucleases (TALENs). ZFN and TALENS require large scale remodelling (500-1500bp) for any new target sequence, whereas CRISPR brings a much greater level of versatility, as any new genomic sequence can be targeted by altering the 20bp guide RNA sequence. The CRISPR vector can remain unchanged, with only the guide sequence being altered for each new target. The CRISPR system also has the advantage of the ability to use multiple guide sequences in a single vector, allowing the CRISPR system to target multiple sites or to increase the precision of the system at one site. However the Cas9 protein is larger than both the TALEN monomer and the ZFN monomer which may cause issues in terms of the delivery of the CRISPR system into the targeted cell and the medium of delivery must account for this (Gupta & Musunuru, 2014). Site specificity is the major obstacle in

the path of CRISPR at the moment, as although the guide RNA is target-specific, one mismatch and sometimes several mismatches are tolerated leading to off-target activity which can have major consequences for the target organism, depending on where these off-target activities occur (Mali et al., 2013, Gupta & Musunuru, 2014).

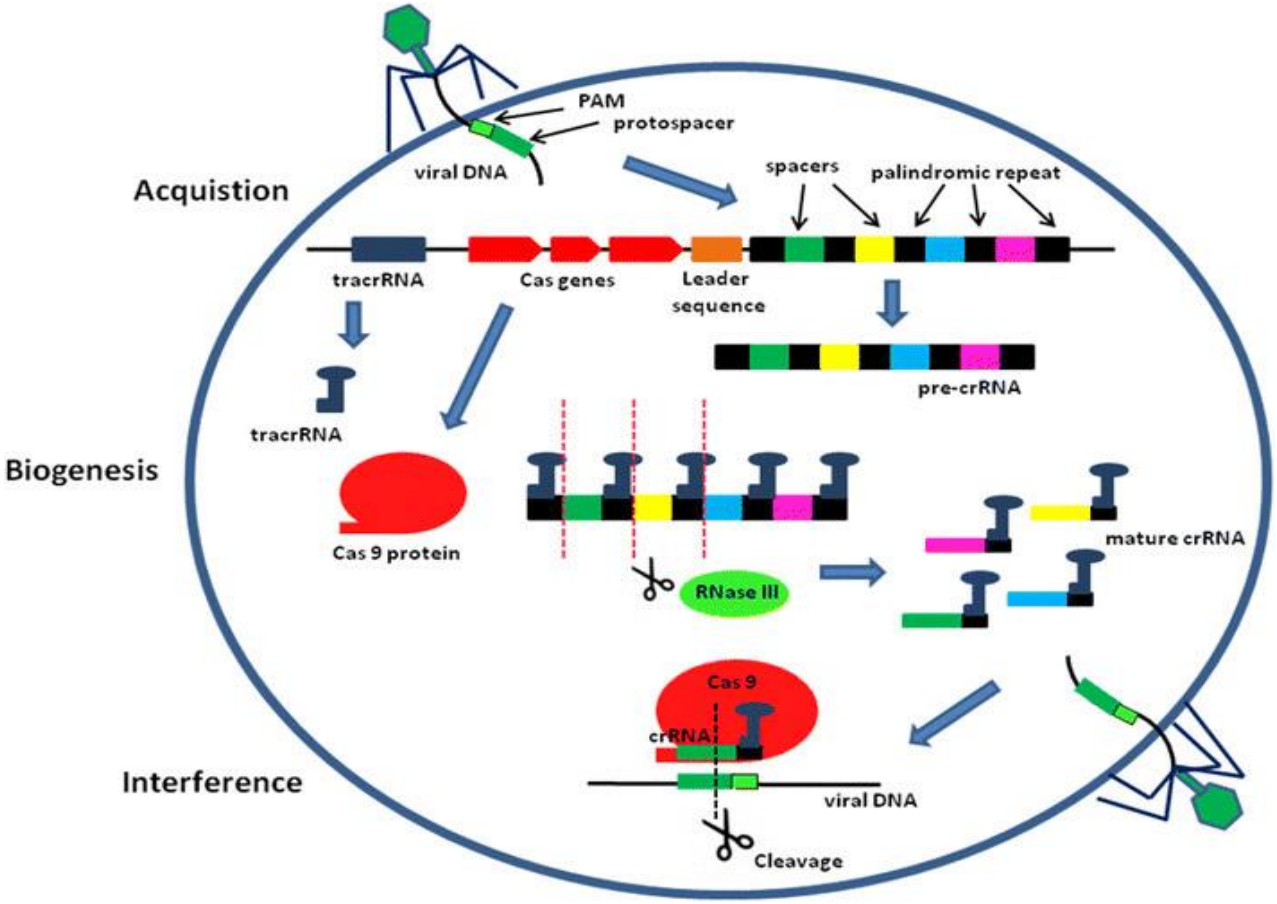

Fig 1. The three stages of the CRISPR/Cas bacterial adaptive immune system (Hryhorowicz et al., 2017).

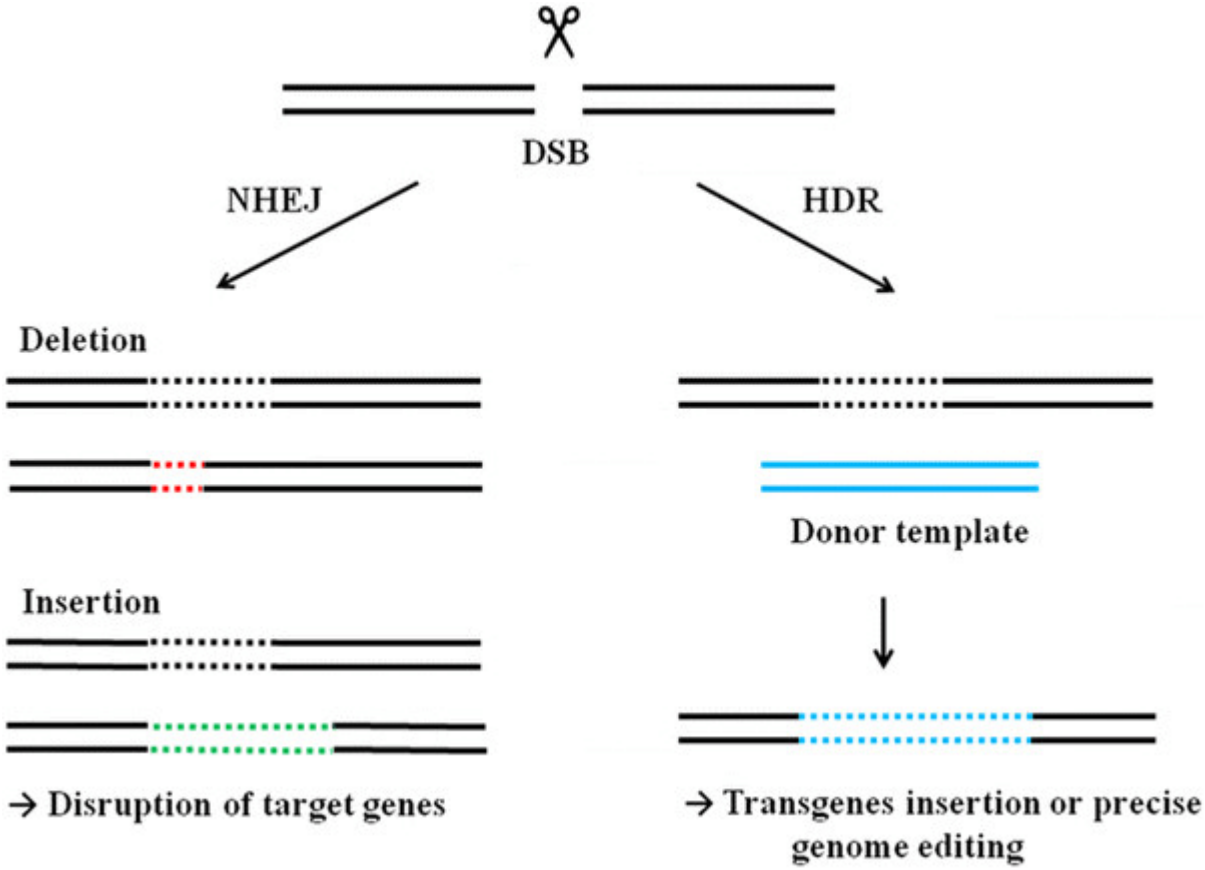

Fig 2. Engineered double-strand breaks will be repaired either by non-homologous end joining (NHEJ) or homology-directed repair (HDR) (Hryhorowicz et al., 2017).

2.2 *Solanum tuberosum* as a target crop for genome engineering

The total world potato (*Solanum tuberosum*) production is estimated at 388,191,000 tonnes in 2017 (FAOSTAT, 2019). Over a billion people consume potatoes every day. Potato cultivation has massively increased in developing countries, particularly in China, India and in Africa, with a third of all potatoes now harvested in China and India. Potatoes are now grown in over 100 countries worldwide (FAOSTAT, 2019). The potato is the fourth most widely consumed crop behind only rice (*Oryza sativa*), wheat (*Triticum*), and maize (*Zea mays*). The relatively recent explosion in popularity in developing countries may be a result of the potatoes' ability to provide a low-cost source of important nutrients, protein, and satiating carbohydrates. The potato is also a source of fibre and potassium, two nutrients of concern in the 2010 Dietary Guidelines for Americans (King & Slavin, 2013). The potato is a high-quality source of

important nutrients and relatively cost-effective making it an excellent choice for bio-fortification through genome engineering.

Potatoes produce biologically active secondary metabolites. One such secondary metabolite produced by potatoes and other members of the nightshade family are steroidal glycoalkaloids (SGA). The main SGAs found in potatoes are α-solanine and α-chaconine. They are found at the highest levels in the leaves of potatoes, however, they are also present in the tubers which are important in terms of human consumption. The amount of glycoalkaloids synthesised by the potato is genetically determined, however, environmental factors can also affect the concentration. Exposure to light and heat or damage to the tuber can increase SGA production. SGAs act as a defence mechanism against potential pests, which may be why exposure of a tuber to light increases production. Potatoes used for human consumption should not exceed 200 mg/kg concentration of SGAs. SGAs can also remain stable during the cooking process (Friedman, 2006). Symptoms of poisoning by eating potatoes with a glycoalkaloid level beyond this limit include nausea, vomiting, diarrhoea, abdominal pain, fever, and disorientation (Mcmillan & Thompson, 1978). It has also been shown that maternal consumption of potatoes with high glycoalkaloid levels during the periconceptional period may increase the risk of neural tube defects and orofacial clefts (Ni et al., 2018).

The Rhamnose:beta-solanine/beta-chaconine rhamnosyltransferase (*SGT3*) gene is located on chromosome 7. It has one isoform, one exon and a length of 4125bp (Mariot et al., 2016). *SGT3* codes for the enzyme β-steroidal glycoalkaloid rhamnosyltransferase which is necessary to convert β-solanine and β-chaconine to α-solanine and α-chaconine, in the terminal step of the potato glycoalkaloid biosynthetic pathway shown in figure 3 (McCue et al., 2007). The downregulation of the *SGT3* has been shown to reduce the levels of SGAs in potatoes. Unintended metabolic changes as a result of downregulating *SGT3* have also been accessed, and leads to an increase in fucosterol and β-sitosterol as cholesterol is a glycoalkaloid precursor (Shepherd et al., 2015). β-sitosterol has been labelled as an orphan phytosterol and clinical trials with β-sitosterol have shown beneficial effects in different diseases (Bin Sayeed et al., 2016).

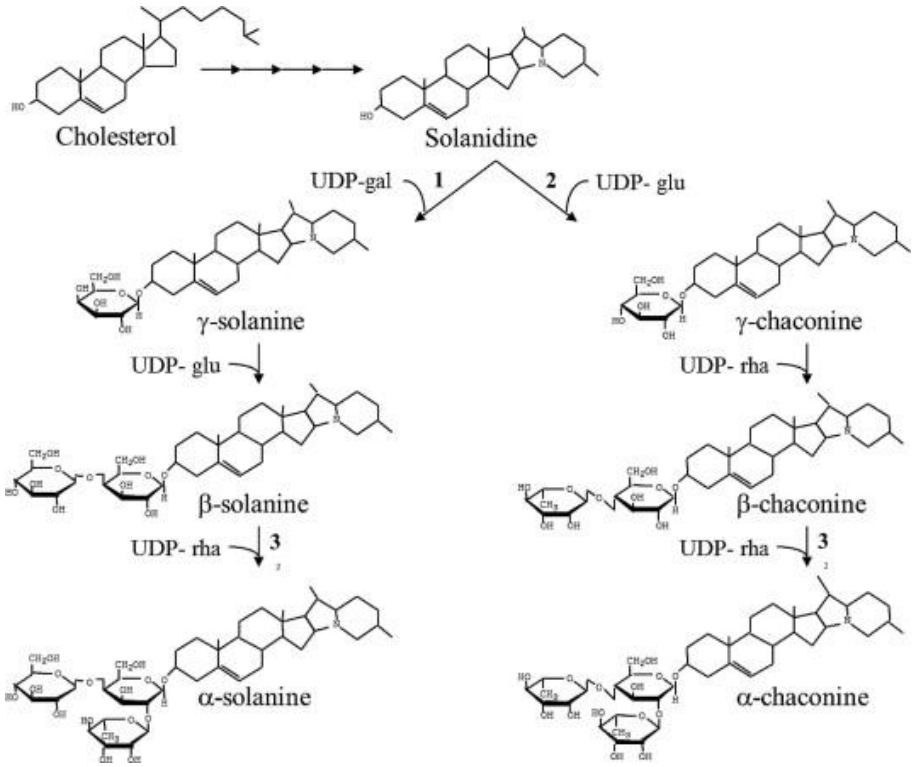

Fig 3. Glycoalkaloid biosynthetic pathway in potatoes. 3 *= SGT3 (UDP-rhamnose:β-steroidal glycoalkaloid rhamnosyltransferase); gal = galactose; glu = glucose; rha* = rhamnose (Shepherd et al., 2015).

2.3 *Agrobacterium tumefaciens* mediated transformation

Agrobacterium is a genus of gram-negative soil-borne phytopathogenic bacteria. *Agrobacterium tumefaciens* is the causative agent of crown gall disease in a wide variety of plants, mainly woody and herbaceous dicots. It is unique in that it employs a trans-kingdom DNA transfer method. This combined with its broad range of susceptible targets make it a natural genetic engineer which can be taken advantage of to introduce foreign DNA into a target plant. *Agrobacterium* can detect signal molecules, such as acetosyringone, and sugars released by wounded plants in the rhizosphere. *Agrobacterium* infects the host plant and inserts T-DNA which is integrated into the plant genome. This is enhanced by the host plants' DNA repair enzymes (Păcurar et al., 2011). The mechanism of the transformation of a host plant cell by *Agrobacterium* is displayed in figure 4.

Although other transformation methods, such as particle bombardment, have the advantage of being genotype independent, *Agrobacterium* is the preferred transformation method as it does not cause mechanical damage to the plant and particle bombardment may cause uncontrolled multiple copy inserts (Travella et al., 2005). *Agrobacterium* also offers a higher transformation frequency and efficiency in potatoes and is the most widely used ahead of particle bombardment, direct DNA uptake by microinjection, protoplast treatment with polyethene glycol (PEG) or electroporation (Chakravarty et al., 2007).

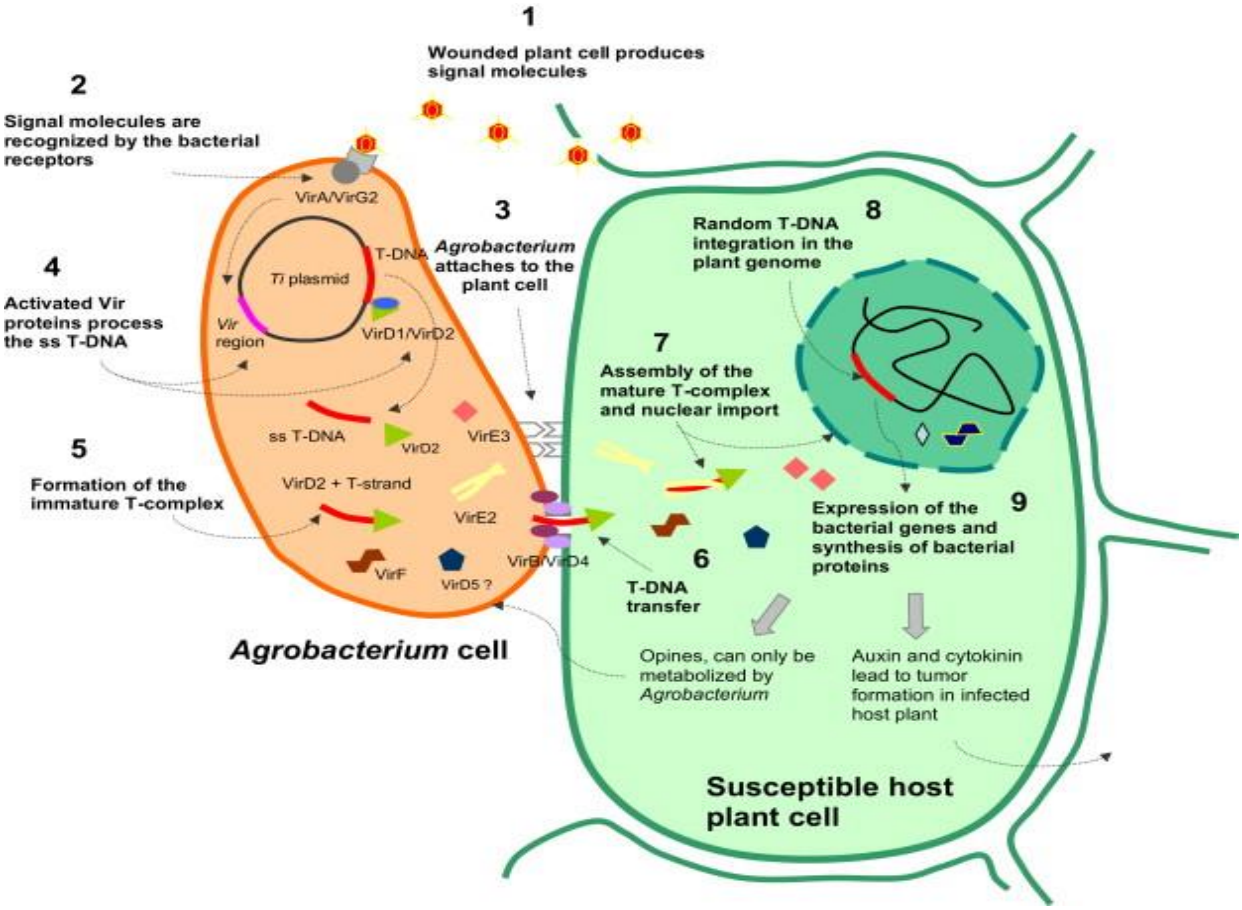

Fig 4. The steps involved in the transformation of a host plant cell by *Agrobacterium* (Păcurar et al., 2011).

2.4 Pre-transformation treatments to increase transformation efficiency

2.4.1 The role heat shock proteins in increasing transformation efficiency

During infection of a target plant with *Agrobacterium tumefaciens*, *Agrobacterium* transfers DNA along with effector virulence (Vir) proteins in a type IV secretion system. A small heat shock protein (HspL) aids in efficient DNA transfer by the type IV secretion system and overexpression of HspL can increase transformation efficiency (Hwang et al., 2015). Previous studies have shown the transformation efficiency of plants with *Agrobacterium* can be increased by subjecting the plants to a heat shock treatment prior to transformation. Rice (*Oryza sativa*) and maize (*Zea mays*) embryos showed increased transformation efficiency when heat-shocked prior to *Agrobacterium* infection (Hiei et al., 2006). Heat treatment has also been shown to increase transformation efficiency in sorghum (*Sorghum bicolour*) (Gurel et al., 2008). Heat treatment during Agrobacterium infection has also been shown to improve transformation efficiency in perennial ryegrass (*Lolium perenne*) and Rice (*Oryza sativa*). Heat treatment during infection instead of prior to may have the increased benefit of causing both bacterial cells and plant cells to release HspL simultaneously (Patel et al., 2013).

2.4.2 The effect of Light-emitting diodes (LED) on transformation efficiency

Light-emitting diodes (LED) offer several advantages over traditional lighting systems when growing plants prior to transformation. They are more energy-efficient and offer much more precise control over the type of light and intensity conditions used to grow the plants. They can be used to cause a direct physiological response, or as a more efficient source of energy for the plant, allowing for faster-growing plants (Pattison et al., 2018). LEDs also offer a mix of different light wavelengths, for example, in potatoes, a mixture of 75 % Red light (600–700 nm) and 25 % Blue light (400– 500 nm) have been shown to offer a higher increase in growth and desired physiological parameters. Wavelength combinations can be tailored to induce specific desirable physiological traits (Chen et al., 2018). Treatment with LED light has also been shown to increase transformation efficiency, transgene inheritance and decrease chimerism in *Agrobacterium* transformed chickpeas (Das Bhowmik et al., 2019). Extending the photoperiod can also result in increased growth rates while

maintaining normal physiological development and accelerated gene transformation pipelines through faster generation times (Ghosh et al., 2018).

2.4.3 The effect of exposure to Rhizospheric bacteria on transformation efficiency

The soil rhizosphere refers to a narrow dynamic area of soil where plant roots interact with the present microorganisms. Plant roots and microorganisms present in the soil interact with each other in harmful, beneficial and neutral ways. There may also be indirect effects between the plants and the microbiota present (Kennedy & de Luna, 2005). The bacteria that colonize this section of soil are known as rhizospheric bacteria. These bacteria can produce volatile organic compounds which impact the growth rates of the plants, as well as having an inhibitory effect on disease-causing microorganisms which infect the plants. These are highly adaptable species of bacteria found in a large range of soil environments. The rhizobacteria which have a positive impact on plant growth are termed plant growth-promoting rhizobacteria (PGPR). PGPR promote plant growth through nutrient solubilisation, nitrogen fixation, by producing plant growth promoters or by acting antagonistically towards phytopathogenic microorganisms (Bhattacharyya & Jha, 2011). PGPR can be isolated from soil samples and identified using techniques such as 16s rRNA sequencing and the volatile organic compounds being produced by the bacteria can also be detected and identified using mass spectrometry. These compounds can then be tested to determine their effect either as a plant growth promoter or as a microbial inhibitor (Ghyselinck et al., 2013). PGPR have previously been shown to reduce disease and increase yield in potatoes (Wang et al., 2019). PGPR treatment may be a useful tool in improving transformation efficiency by making a target plant more amenable to the stressful events of transformation through *Agrobacterium*. PGPR may also reduce ethylene levels in plants through ACC deaminase. Suppression of ethylene can lead to increased gene transfer efficiency (Nonaka & Ezura, 2014).

2.5 The future of plant genome engineering

One major obstacle preventing genetically engineered crops from widespread use throughout the world are the regulations placed upon them. Figure 5 shows the

different methods of genetic modification and the regulations they face. In the United States of America, the Federal Government developed a Coordinated Framework for the Regulation of Biotechnology in 1987 in order to regulate crops derived through genetic modifications. Genetically modified organisms (GMO), plants modified using methods such as *Agrobacterium*, which involved the insertion of foreign *Agrobacterium* DNA into the plant, were all regulated. However, with the advent of the new and more precise genetically engineered (GE) crops, these regulations are becoming increasingly obsolete. Genetic engineering of crops, using methods such as CRISPR are more precise. The United States Department of Agriculture now plans to deregulate many future GE crops in a hopeful step forward for the future of plant biotechnology. This will be the first major reform of these regulations and hopes to "provide a clear, predictable, and efficient regulatory pathway for innovators, facilitating the development of new and novel genetically engineered organisms" (United States Department of Agriculture, 2019). The percentages of GE crops planted in the United States in 2019 of all crops planted for *Zea mays* (corn), *Gossypium* (cotton), and *Glycine max* (soybean) were 92 %, 98 % and 94 % respectively (United States Department of Agriculture, 2019).

The European Union and its member states adopt a "process-based" regulation system whereby the technological process defines the sanctions applied rather than the crop itself ("product-based") as is used in the United States. The EU law has been limiting on GMO crops cultivation and sale and the Court of Justice of the European Union ruled that CRISPR edited crops would be subject to these same 2001 directives despite other countries granting exemptions. This means CRISPR edited crops in the European Union will still face a lengthy approval process and all products will be labelled as GM (Callaway, 2018).

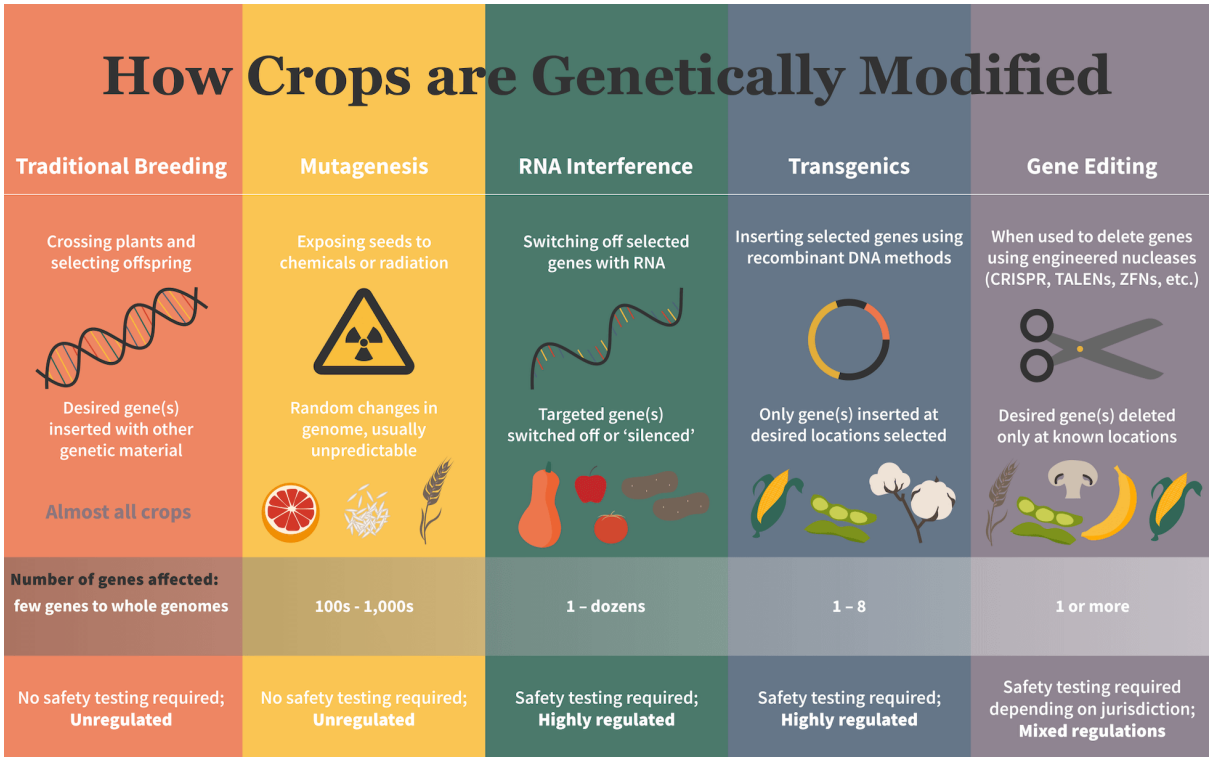

Fig 5. Methods of genetically modifying crops and the level of regulation they face (Schreiber, 2019).

2.6 Aims of the project

The aims of this research were as follows:

Investigate the effect of LEDs on plants prior to transformation and determine the optimal wavelength for each cultivar.

To isolate and identify soil bacteria from a commercial potato farm and determine which volatile organic compounds these bacteria produce and their impact on transformation efficiency.

To investigate the effect of heat treatment on transformation efficiency.

To design specific guide RNA for a CRISPR Cas9 system to target the *SGT3* gene in potatoes.

Ultimately to improve transformation efficiency of *Solanum tuberosum* and to create INDELs in the *SGT3* gene to disrupt the production of glycoalkaloids in these cultivars.

3. Materials and methods

3.1 Media preparation for plant tissue culture

All plant material necessary for this project were cultured on Murashige and Skoog basal salt medium (Murashige & Skoog, 1962). The media was made up to the specifications of table 1, and the pH was adjusted to 5.8. Materials were supplied by Sigma Aldrich (St. Louis, Missouri, United States). Following this, the media was autoclaved and poured into either food pots or Micro boxes and allowed to solidify. The boxes or pots were stored for a period of a week and routinely checked for any signs of contamination. Once the boxes or pots had passed a week without showing signs of contamination, they were used for tissue culture.

Table 1. Murashige and Skoog tissue culture medium composition

Agar	6 g
Sucrose	15 g
Murashige and Skoog powder	2.2 g
De Ionised water	950 mL
Total volume	1 L

3.2 Nodal tissue culture

Two cultivars of *Solanum tuberosum* were selected; Maris Piper and Golden Wonder. These micro plants were supplied by Teagasc (Oak Park Rd, Pollerton Little, Carlow, Ireland). Nodal plant tissue culture was carried out as per Tovar and Dodds Tissue culture propagation methodology (Tovar & Dodds, 1986). Internodal cuttings of about 2-3 mm in length were taken from the donor plant and placed in either food pots or micro boxes. All utensils were sterilized in a glass bead sterilizer before and

after cutting. The explants were placed in food pots at a rate of four per pot, or 16 per micro box, evenly spaced out. All tissue culture was carried out under aseptic conditions and in a laminar flow hood. The plants were stored in a growth room at approximately 20-22 °C with a long photoperiod of 16 hours light 8 hours dark supplied by fluorescent tubes. Plants were routinely checked for signs of contamination and any contaminated specimens were removed and autoclaved (Tovar & Dodds, 1986). This procedure was replicated in order to provide the necessary amount of plant material.

3.3 LED treatment

Freshly cultured food pots of each cultivar were placed under Heliospectra LX601C 630W LED Grow Light (Heliospectra AB, Fiskhamnsgatan, Göteborg, Sweden) units set to various wavelengths for a period of 4 weeks. 2 food pots containing 4 internodal cuttings of each cultivar were placed under the LED units for each of the different wavelengths, with each cutting acting as a replicate. The wavelengths chosen were; white (5700 K), red (660 nm), far-red (735 nm), blue (450 nm) and combinations of red and blue (75 % red: 25 % blue, 50 % red: 50 % blue and 25 % red: 75 % blue) (Chen et al., 2018). A longer photoperiod of 22 hours was employed to increase the growth rate of the plants (Ghosh et al., 2018). The LED growth room was kept at 20-22 °C. After 4 weeks, the plants grown under these conditions were removed and leaf number, stem length, and fresh weight were measured. The samples were placed in brown paper bags and dried to a constant weight in an oven at 60 °C overnight. The next day dry weight of the samples were measured. Averages for each cultivar and each treatment were calculated.

3.4 Soil sampling

Soil samples were taken from a commercial potato farm located in Fermoy, Cork, Ireland. The potato fields were 25 acres in size. Large soil samples were taken using

an auger from both the rhizosphere and the headland. Soil samples from each area were bagged and pooled.

3.5 Rhizospheric bacteria isolation

For bacteria isolation, 1 g of soil was added to 5 mL phosphate-buffered saline with 10 ~ 6 mm glass beads and vortexed for two minutes. Serial dilutions up to 10^{-6} were prepared. 100 µL of dilutions 10^{-3} and 10^{-4} was spread plated onto four-fold diluted tryptic soy agar (TSA) supplemented with 0.03 % cycloheximide (to inhibit fungal growth) supplied by Sigma Aldrich (St. Louis, Missouri, United States). The plated isolates were incubated at 20 ℃ and 30 ℃. Plates were observed over two weeks and colonies with differing morphologies were isolated to pure cultures on TSA plates resulting in 12 different isolates (Ghyselinck et al., 2013).

3.6 Sanger sequencing of isolates

12 falcon tubes containing Tryptic soy broth (TSB) were inoculated with bacteria from each isolate along with three controls and placed on a shaking incubator at 28 °C and 190 RPM. DNA was extracted using a Qiagen™ DNeasy Blood and Tissue kit (QIAGEN, Hilden, Germany). A region of approximately 1480bp of the 16s rRNA gene was amplified from each isolate using polymerase chain reaction (PCR). All reactions were conducted in 20 µL volume consisting of 10 µL 2X PCR master mix (TopBio™), 1 µM each of forward primer 8F (AGAGTTTGATCCTGGCTCAG) and reverse primer 1492R (GGTTACCTTGTTACGACTT) with 1 µL of DNA as template. Cycling conditions were as follows: 15 minutes at 95 °C, followed by 32 cycles of 30 seconds at 94 °C, 60 seconds at 54 °C and 90 seconds at 72 °C. A final extension stage for 7 minutes at 72 °C was also performed. PCR products were run on a 1 % agarose gel stained with SYBR Safe (Invitrogen, Carlsbad, California, United States). PCR bands were excised and placed in a safe lock tube (Eppendorf, Hamburg, Germany) with fibreglass wool and centrifuged for 3 minutes at 13000 RPM. For sequencing reactions, a mix of 0.5 µL Big dye mix, 1.75 µL sequencing buffer, 0.32 µL primer (8F, 100 µM), 6.43 µL ddH_2O, and 1 µL of purified DNA was made up for each sample. Samples were then placed on a thermocycler for 1 minute at 96 °C, followed by 25 cycles of 10 seconds at 96 °C, 5 seconds at 50 °C, and 4 minutes at 60 °C and stored at 4 °C until purification step. Each 10 µL sequencing reaction was added to 2.5 µL 125 mM EDTA solution (pH 8.0)

and 30 µL absolute ethanol. Samples were incubated at room temperature for 20 minutes followed by centrifuging at 1870 xg at 4 °C for 45 minutes to pellet the sequencing reaction. All liquid was then carefully removed and 30 µL of 70 % ethanol was added and samples were centrifuged for a further 15 minutes at 1870 xg at 4 °C. The supernatant was again pipetted off and samples were placed on a warm plate for 10 minutes at 30 °C to ensure pellets were fully dried. Following this 10 µL of Hi-Di formamide was added to each sample to re-suspend them. Sequences were run on a 3500 XL Genetic Analyzer (Thermo Fisher Scientific, Waltham, Massachusetts, United States).

3.7 Identification of isolates

The resulting chromatograms were used to form a consensus sequence for each of the isolates. The Ribosomal database project (RDP) classifier tool was used to identify the genus of each isolate (Wang et al., 2007). The 16s rRNA sequences from each of the isolates were compared to the National Centre for Biotechnology Information (NCBI) GenBank database as well as the RDP database (National Center for Biotechnology Information, U.S. National Library of Medicine, Bethesda, Maryland, United States). BLAST search parameters were altered to tailor to the NCBI 16s ribosomal RNA database and similarity searches were also performed using the RDP tools. The species with the highest similarity scores from each database were matched to each isolate.

3.8 Identification of Rhizobacterial volatiles using gas chromatography-mass spectrometry

The Volatile organic compounds (VOCs) produced by the soil isolates were measured across three different media types; Tryptic soy broth (TSB), liquid Murashige and Skoog (MS) and Methyl Red and Voges-Proskauer broth (MRVP). The isolates were sub cultured onto fresh TSA plates and left to grow for 24 hours. The chosen medium was inoculated with a colony from the fresh TSA plates in a glass vial sealed with

Teflon faced butyl Septa and an aluminium seal (Supelco, Bellefonte, Pennsylvania, United States) and placed on a shaking incubator at 170 RPM at 28 °C for 24 hours. Solid-phase microextraction (SPME) fibres from Supelco (Bellefonte, Pennsylvania, United States) were used to absorb volatiles being produced by the samples in a 50 °C incubator for 40 minutes. Gas chromatography-Mass spectrometry was used to identify the volatiles being produced (Shimadzu, Kyoto, Kyoto Prefecture, Japan). Volatiles were accepted at a 90 % identity rate. The isolates were co-cultivated with the plant material in order to examine the effect on plant growth.

3.9 CRISPR guide RNA design

The guide RNA was designed to target the Rhamnose:beta-solanine/beta-chaconine rhamnosyltransferase (*SGT3*) gene which codes for an enzyme as part of the solanine production pathway. A 2500 base pair partial sequence of the *SGT3* gene in the potato genome was sourced on GenBank (accession number KC331037.1). The *SGT3* region for our potato cultivars, Maris Piper and Golden Wonder, were amplified and sequenced using PCR and Sanger sequencing and a multiple sequence alignment between our cultivars and the online sequence was performed to ensure the region was highly conserved in the area where gRNAs were designed. The partial sequence was inputted into CRISPRdirect to find suitable gRNA sequences and the specificity check was changed to *Solanum tuberosum* to ensure that the guides were specific to the target region and did not match anywhere else in the potato genome (Naito et al., 2015). The search results were filtered to only include highly specific targets. This resulted in four suitable sequences that occur adjacent to the appropriate PAM sequence (in the case of this vector that sequence is NGG). The Guide RNA must be adjacent to the protospacer adjacent motif for CRISPR to effectively target the selected region. The Cas9 is specific to an NGG PAM sequence. From this process four guide sequences were selected; 1- CCATACGCCATGACGAGTCATAT, 2- CCATGACGAGTCATATAACTCCA, 3- GCTAGACTCTTCGCCCTCCATGG, 4- CCTCAAAGTTACTATCATTGCCC. A guanine nucleotide was added to the first

two sequences as a U6 promoter is used in the CRISPR vector. A GATT linker was added to the forward strand of each and an AAAC linker was added to the reverse strand to allow for these oligonucleotides to be ligated to the CRISPR vector at the *BsaI* site. The oligonucleotides were supplied by Integrated DNA Technologies (Coralville, Iowa, United States). These oligonucleotides were made up to a concentration of 100 µM by adding ddH_2O and were stored at -20 °C.

Table 2. Guide RNA sequences. Yellow indicates linker. Green indicates added guanine.

Target	Oligo	Sequence (5'-3')
1	Forward	GATTGATATGACTCGTCATGGCGTA
1	Reverse	AAACTACGCCATGACGAGTCATATC
2	Forward	GATTGTGGAGTTATATGACTCGTCA
2	Reverse	AAACTGACGAGTCATATAACTCCAC
3	Forward	GATTGCTAGACTCTTCGCCCTCCA
3	Reverse	AAACTGGAGGGCGAAGAGTCTAGC
4	Forward	GATTGGGCAATGATAGTAACTTTG
4	Reverse	AAACCAAAGTTACTATCATTGCCC

3.10 Validation of guide RNA through *in vitro* digestion

Synthetic guide RNA was ordered from Synthego (Synthego, California, United States) in order to test the effectiveness of the guides *in vitro* (Mehravar et al., 2019). DNA was extracted from plants (leaves) using Edwards solution (Edwards et al., 1991). The leaves and internodes were ground up in safe lock tubes with a pestle and the samples were submerged in 500 µL Edwards solution (200 mM Tris-HCl (pH 7.5), 250 mM NaCl, 25 mM EDTA, and 0.5 % SDS). The samples were further crushed with the pestle and vortexed. Samples were centrifuged at 13000 RPM for 5 minutes. The supernatants were removed from each sample and placed in a new safe lock tube containing 300 µL of isopropanol. Samples were mixed well by inverting and left to sit for 5 minutes. Samples were centrifuged at 13000 RPM for 10 minutes. The supernatant was discarded and 500 µL of 70 % ethanol was used to wash the pellet. Samples were centrifuged at 13000 RPM for 5 minutes. Ethanol was removed without disturbing the pellet. Samples were left to air dry for 20 minutes and re-

suspended in nuclease-free water. Samples were stored at 4 °C for 20 minutes. Samples were vortexed to re-suspend and centrifuged at 13000 RPM for 3 minutes and the supernatant used directly for PCR. All reactions were conducted in 20 µL volume containing PCR buffer with 10 µL of 2X TopBio PCR master mix, 1 µM of each of primer and with 1 µL DNA as template. The primers were designed using the Primer3 tool (Untergasser et al., 2012). The primers spudfor 288 (CCCACTGACATGAAATTTTGGC) and spudrev 648 (GGGCTTGCGATGAAGTTTC) were selected and ordered from IDT (Integrated DNA Technologies, Iowa, United States). Cycling conditions consisted of an initial denaturation step of 3 minutes at 95 °C, followed by 35 cycles of 30 seconds at 95 °C, 30 seconds at 52 °C and 30 seconds at 72 °C. A final extension period of 5 minutes at 72 °C completed the reaction. The PCR products were run on a 1 % agarose gel stained with SYBR Safe (Invitrogen, Carlsbad, California, United States). The bands were excised and purified using a QIAquick gel extraction kit (QIAGEN, Hilden, Germany). The extracted DNA was nanodropped in order to determine the ng/µL concentration. The *in vitro* digestion reaction was assembled in a nuclease-free PCR tube in the following order as described in table 3.

Table 3. *In vitro* Digestion of *SGT3* DNA by guide RNA and Cas9 complex reaction.

Component	Volume
Nuclease-free water	21 µL
10X Cas9 Nuclease Reaction Buffer	3 µL
1 µM sgRNA	1 µL (~30 nM final)
1 µM Cas9 Nuclease, S. pyogenes	1 µL (~30 nM final)
20 nM substrate DNA	4 µL (~3 nM final)
Total reaction volume	30 µL

The reaction was mixed thoroughly and spun down briefly. The reaction was incubated at 37 °C for 1 hour, then heated to 65 °C for 10 minutes to deactivate Cas9 nuclease. The fragments were analysed by gel electrophoresis on a 1.2 % agarose gel.

3.11 Transforming *Escherichia coli* with CRISPR vectors

Ultracompetent Cells (XL10-Gold) were obtained from Agilent Technologies (Santa Clara, California, United States). The CRISPR vector (pGNK-LeCas9-AtUbp-gRNA) supplied by colleagues from the James Hutton Institute at the University of Dundee was measured for DNA concentration using a nanodrop. Three additional vector, p63(dicot)U6-gRNA:CMV-Cas9-beta-glucuronidase (*Agrobacterium* plasmid), were ordered with the custom guides pre-inserted from Sigma-Aldrich (St. Louis, Missouri, United States). The p63(dicot)U6-gRNA:CMV-Cas9-beta-glucuronidase vectors were at a concentration of 20 ng/µL. Five 15 mL falcon tubes were pre-chilled on ice. The NZY broth (see table 4 below) was preheated to 42 °C in a water bath. Five 100 µL vials of ultracompetent cells were thawed on ice. An aliquot of 100 µL of cells was pipetted into each of the pre-chilled falcon tubes. 4 µL of β-Mercaptoethanol mix was added to each falcon tube. Cells were incubated on ice for 10 minutes, swirling every 2 minutes. To insert vector DNA, 2 µL (50 ng) of the CRISPR vector (pGNK-LeCas9-AtUbp-gRNA) was added to one falcon tube and 2.5 µL (50 ng) of each of the CRISPR vectors p63(dicot)U6-gRNA:CMV-Cas9-beta-glucuronidase were added to three falcon tube, the final falcon tube acted as a negative control. The falcon tubes were swirled gently and incubated on ice for 30 minutes. The cells were heat pulsed at 42 °C in a water bath for 30 seconds. Following the heat pulse, the cells were incubated on ice for 2 minutes. Following this, 0.9 mL of preheated NZY broth was added to each of the falcon tubes and the tubes were incubated at 37 °C at 225 RPM for 1 hour. 200 µL of the transformation mixture was plated onto LB agar plates (see table 5 below) containing the appropriate antibiotics (50 mg/L Kanamycin and 10 mg/L tetracycline). Plates were incubated at 37 °C overnight to allow colonies to grow. Single colonies were selected from the plates and used to inoculate LB broth containing the appropriate antibiotics. The inoculated broths were placed on a shaking incubator at 37 °C at 170 RPM overnight. 4 mL of broth were used for plasmid extraction using a QIAprep spin miniprep kit (QIAGEN, Hilden, Germany). The DNA concentration in ng/µL was measured for each sample using a nanodrop nd-1000 spectrophotometer (Thermo Fischer Scientific, Waltham, Massachusetts, United States).

Validation of successfully extracted pGNK-LeCas9-AtUbp-gRNA and p63(dicot)U6-gRNA:CMV-Cas9-beta-glucuronidase vectors was performed by running the vectors on a 2 % agarose gel next to a 100bp ladder.

Table 4. NZY Broth outgrowth medium composition

NZY Broth (per Litre)
10 g NZ amine (casein hydrolysate)
5 g yeast extract
5 g NaCl
Add deionized H_2O to a final volume of a litre
Adjust pH to 7.5 and autoclave
Add the following filter sterilised prior to use:
12.5 mL 1 M $MgCl_2$
12.5 mL 1 M $MgSO_4$
20 mL 20 % (w/v) glucose

Table 5. LB agar medium composition

Luria Broth (LB) agar (per litre)
10 g NaCl
10 g tryptone
5 g yeast extract
20 g agar
Add deionized H_2O to a final volume of a litre
Adjust pH to 7.0 and autoclave
Filter sterilise and add appropriate antibiotics

3.12 Vector digestion

The pGNK-LeCas9-AtUbp-gRNA vector was cut at the *BsaI* sites to facilitate the ligation of the guide RNA to the CRISPR vector. A digestion reaction was made up of 1 µg or 6.5 µL of CRISPR vector DNA, 5 µL of 10xCutSmart buffer (New England Biolabs, Ipswich, Massachusetts, United States), 1 µL *BsaI* restriction enzyme (New England Biolabs, Ipswich, Massachusetts, United States) and 37.5 µL nuclease-free water to bring the final volume up to 50 µL. The reaction was incubated at 37 °C for one hour, followed by 65 °C for 20 minutes. The vector was purified by running it on a 1 % agarose gel, excising the band and extracting the vector using a QIAquick gel extraction kit (QIAGEN, Hilden, Germany).

3.13 Ligation of guide RNA to CRISPR vector

Duplexes of the guide RNAs were formed by annealing the oligonucleotides by heating to 95 °C for 6 minutes and allowing it to cool to room temperature. The guide RNA was ligated to the cut CRISPR vector using a T4 DNA ligase kit (Invitrogen, Carlsbad, California, United States). A 20 µL ligation reaction was made up to the specifications of table 6 below.

Table 6. Ligation of guides to CRISPR vector reaction

Component	Volume
5x ligase buffer	4 µL
guide RNA duplex (at two different concentrations, 1 µM and 100 µM)	2 µL
digested CRISPR vector (approx. 100 ng of DNA)	5 µL
T4 ligase enzyme	0.1 µL
ddH_2O	8.9 µL

The reaction was incubated at 27 °C for one hour. PCR was used to validate a successful ligation. Primers specific to the pGNK-LeCas9-AtUbp vector were designed on either side of the *BsaI* cutting site using the Primer3 tool (Untergasser et al., 2012). Each 10 µL PCR reaction consisted of 5 µL 2X PCR master mix, 1 µM CRISPR F1 forward primer (AACTCCAGAAACCCGGTACC) and 1 µM CRISPR R1 reverse primer

(TCTTCAAAAGTCCCACATCGC), and 1 µL of vector DNA to be amplified. Cycling was performed under the following conditions; 3 minutes at 95 °C, followed by 32 cycles of 30 seconds at 95 °C, 30 seconds at 54 °C and 30 seconds at 72 °C, 5 minutes at 72 °C and stored at 4 °C until further use. PCR products were then run on a 2.5% agarose gel stained with SYBR Safe (Invitrogen, Carlsbad, California, United States). 8 ligated vectors (4 guides with two different concentrations each) were amplified. The bands were excised, the DNA was extracted using a QIAquick gel extraction kit and sequencing was performed as described previously

3.14 Maintenance of *Agrobacterium* culture

Luria broth (LB) supplemented with rifampicin (25 mg/L) was prepared and inoculated with *Agrobacterium* strain AGL1. The broth was incubated on a shaking incubator at 28 °C and 190 RPM overnight. An aliquot of 100 µL of the *Agrobacterium* culture was transferred to LB agar plates containing 25 mg/L rifampicin and spread using a glass spreader. The cultures were incubated at 28 °C.

3.15 Transformation of *Agrobacteria*

In order to insert the CRISPR vector into the *Agrobacterium* strains a freeze/thaw shock transformation was carried out. A single colony of *Agrobacterium* AGL1 was used to inoculate 3 mL LB broth containing 25 mg/L rifampicin in a falcon tube. The broth was incubated on a shaking incubator at 28 °C and 190 RPM overnight. 50 mL of LB broth in a 250 mL flask was inoculated with 0.5 mL of the overnight culture and allowed to grow to mid-log phase (indicated by an appropriate OD reading) on a shaking incubator at 28 °C and 190 RPM. A spectrophotometer (Shimadzu, Kyoto, Kyoto Prefecture, Japan) was blanked with LB broth and used to take an OD_{600} of each of the three cultures. The cultures were diluted as necessary to achieve the desired optimal OD of between 0.6 and 0.9. The samples were chilled on ice for 5 minutes and placed in a refrigerated centrifuge (Sigma, St. Louis, Missouri, United States) and spun at 3000 RPM at 4 °C for 5 minutes in order to form a cell pellet. The supernatant was discarded, and the cell pellet was re-suspended in 1 mL of ice-cold 20 mM calcium chloride. Four safe lock tubes were chilled and 100 µL of each solution was

transferred into each safe lock tubes. 1 µg of plasmid DNA was added to each safe lock tube and mixed by tapping. The safe lock tubes were flash frozen in liquid nitrogen the thawed for 5 minutes at 37 °C. Once thawed, 1 mL of LB was added to each sample and the contents were transferred to 15 mL falcon tubes. The falcon tubes were incubated on a shaking incubator at 30 °C for 2 hours and the contents were transferred to 1.5 mL safe lock tubes and spun on a centrifuge for 5 minutes at 4000 RPM to pellet the cells. The supernatant was discarded, and the pelleted cells were re-suspended in 100 µL of LB. The suspension culture was plated onto LB using a glass spreader on plates containing 25 mg/L rifampicin (*Agrobacterium* resistance) and 50 mg/L Kanamycin (CRISPR vector resistance) and incubated for 2 days at 30 °C. Transformed colonies appeared after 2 days. The antibiotics present in the medium selected for transformed *Agrobacterium* only.

3.16 Pre-transformation treatments of *Solanum tuberosum*

Prior to transformation, the plant samples were treated with pre-transformation treatments based on the results from the bacterial isolation experiments and the LED experiments. The optimal wavelengths to induce plant growth from the LED experiments for each cultivar were used to design a pre-transformation 4-week growth period under the LED units as shown in figure 6 below. The bacterial isolates were co-cultivated with the plant samples for 10 days prior to transformation. Both Maris Piper and Golden Wonder plants were cultured in liquid MS medium using magentas with a foam square to allow the plant roots to access the medium. 4 foam magentas containing liquid MS (two for each cultivar) were placed in a micro box and allowed to grow for one week. Following one week of growth, a petri dish of the desired medium to induce the production of growth-promoting volatile organic compounds and one of the identified bacterial isolates streaked on the plate was placed in the micro box. This was repeated for each of the isolates and a control of a petri dish containing the medium but no isolate. The lids of the magentas were removed to allow the volatile organic compounds being produced to access the plants as shown in figure 7 below. The micro boxes were placed in the growth room under a long photoperiod of 16 hours on 8 hours off at room temperature under

fluorescent lights. This experimental method was devised by Darren Heenan Daly (PhD, Butler Building, University College Cork, Ireland).

Fig 6. LED treatments prior to transformation

Fig 7. Co-cultivation of *Solanum tuberosum* and bacterial isolates prior to transformation

3.17 Co-cultivation transformation

LB broth was inoculated with each of the transformed strains of *Agrobacterium* containing the CRISPR vector. The inoculated broths were placed on a shaking incubator at 28 °C and 190 RPM overnight to allow the cultures to grow. The next day a spectrophotometer (Shimadzu, Kyoto, Kyoto Prefecture, Japan) was blanked with

LB broth and used to take an OD_{600} of each of the three cultures. The cultures were grown as long as necessary to achieve the desired optimal OD of over 0.6. Internodal cuttings and leaf samples were taken from each cultivar, both treated and un-treated. The internodes were sliced down the middle using a sterile blade and placed in a safe lock tube containing *Agrobacterium* inoculated broth at the desired OD. Approximately 6 internodes were placed in each safe lock tube. The leaves were cut into squares and placed in a safe lock tube containing inoculated broth. All samples were placed on a shaking incubator at 28 °C for 10 minutes. In order to heat shock samples, several of the samples were incubated at 42 °C for 3 minutes prior to the 10 minute incubation period. The samples were removed from the safe lock tubes after 10 minutes and dried on Whatman filter paper (Whatman plc, Maidstone, United Kingdom). The dried samples were placed on MS plates wound side down and placed in the growth room for 2 days. A final set of co-cultivation transformations were carried out with *Agrobacterium* strain AGL1 which contained a gus gene but no CRISPR vector to test the effect of the different pre-transformation treatments without relying on the presence of a knockout.

3.18 Gus assay

Following co-cultivation transformation, a Gus assay was carried out. The Gus assay solution was made up of; 25 mL Phosphate buffer solution at pH 7 (PO_4), 5 mL triton X solution ($C_{14}H_{22}O(C_2H_4O)_n(_{n=9-10})$), 1 mL of 1 mM Potassium ferricyanide solution ($K_3[Fe(CN)_6]$), 5 mL of methanol (CH_3OH), 50 mg of x-gluc ($C_{14}H_{13}BrClNO_7$) dissolved in Dimethyl sulfoxide (C_2H_6OS), and 0.735 mL of 100 mM Sodium Citrate buffer ($Na_3C_6H_5O_7$), deionised water was added to bring the total volume to 50 mL. The transformed samples were placed in the Gus assay solution and incubated overnight at 37 °C. The next day the samples were bleached overnight in ethanol to remove the chlorophyll. Samples were examined for the presence of a blue colouration indicating that the sample had been successfully transformed.

3.19 Polymerase chain reaction and Sanger sequencing to detect the presence of INDELs

DNA was extracted from treated plants and control plants (internodes and leaves) using Edwards solution (Edwards et al., 1991) as previously described. The *SGT3* region was amplified, excised, purified and sequenced as previously described. The primers used were spudfor 288 (CCCACTGACATGAAATTTTGGC) and spudrev 648 (GGGCTTGCGATGAAGTTTC). The resulting chromatograms were viewed on Chromas (Technelysium Pty Ltd, South Brisbane, Australia) and used to form a consensus sequence. The sequences were compared to the control sequences using Clustal-Omega (Conway Institute, University College Dublin, Ireland) to perform a multiple sequence alignment in order to detect any deletions which may have occurred.

3.20 Statistical analysis

Analysis of variance (ANOVA) was carried out using Microsoft Excel to determine whether different treatments had a statistically significant effect. All graphs were produced on Microsoft excel (Microsoft Corporation, Albuquerque, New Mexico, United States). Independent sample T-tests were also performed on SPSS statistics (International Business Machines, New York, United States) to determine whether the difference between the different treatments and the controls were statistically significant (P-value < 0.05).

4. Results

4.1 Identification of bacterial isolates

Once each isolate was sequenced, the sequences were put through BLAST against the NCBI 16s ribosomal RNA database. The closest related strains were identified by the results with the highest percentage identity as can be seen in Table 7 below. In incidents where more than one strain had the same percentage identity, all strains

were included. A cut off point of 97 % identity was implemented, however, all isolates were above this range. Of the 12 bacterial isolates isolated from the rhizosphere and headland of a commercial potato farm, eight isolates were successfully sequenced. These sequences were used to classify the bacteria and to identify based on similarity to other sequences on the NCBI database using the BLAST tool. Four were successfully identified down to strain level, two were narrowed down to species level and two had more than one species with the same percentage identity. All identities were above 99 %.

Table 7. BLAST results of each isolate.

Isolate	Origin	Media	Incubator temperature	Classification	Closest related strain(s)	Percentage similarity
1	Rhizosphere	TSA	20 °C	*Bacillus*	*Bacillus mycoides strain NBRC 101228, strain ATCC 6462, strain DSM 11821, strain 273*	100 %
2	Rhizosphere	TSA	20 °C	*Bacillus*	*Bacillus mycoides strain NBRC 101228, strain DSM 11821, strain 273, strain ATCC 6462,*	100 %
4	Rhizosphere	TSA	20 °C	*Bacillus*	*Bacillus pumilus strain NBRC 12092*	99.72 %
5	Headland	TSA	20 °C	*Lysinibacillus*	*Lysinibacillus mangiferihumi strain M-GX18*	99.51 %

9	Headland	TSA	30 °C	*Bacillus*	*Bacillus aerius strain 24K*	100 %
10	Rhizosphere	TSA	30 °C	*Bacillus*	*Bacillus stratosphericus strain 41KF2a, Bacillus aerius strain 24K, Bacillus altitudinis 41KF2b*	99.93 %
11	Rhizosphere	TSA	30 °C	*Bacillus*	*Bacillus aerius strain 24K*	100 %
12	Rhizosphere	TSA	30 °C	*Bacillus*	*Bacillus altitudinis 41KF2b, Bacillus aerius strain 24K*	99.93 %

4.2 Phylogenetic tree of isolates

The sequence of each isolate was put into Clustal-Omega (Conway Institute, University College Dublin, Ireland) to perform a multiple sequence alignment. A phylogenetic tree was generated from the results of the multiple sequence alignment in order to show the relation of the isolates to each other as shown in figure 8 below. The values shown in the tree represent the "length" of the branch. These numbers represent the amount of genetic change. Generally, the larger the number, the larger the amount of genetic change. The phylogenetic tree analysis showed that the bacteria grown at the same temperature were more closely related to each other.

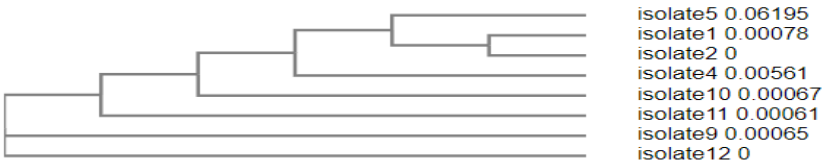

Fig 8. Phylogenetic tree of isolate sequences.

4.3 Volatile organic compounds detected by Gas Chromatography-Mass Spectrometry

Solid-phase microextraction (SPME) Gas chromatography-Mass spectrometry was used to identify the volatile organic compounds being produced by these isolates across a variety of different media types. Six of the isolates (isolate 1, isolate 4, isolate 5, isolate 9, isolate 10, isolate 11) produced known growth promoters such as 2,3-Butanediol and 3-hydroxy- 2-butanone, however, this varied between the different media types. Isolate 1, isolate 2, isolate 5 and isolate 9 also produced known fungal inhibitors. Tryptic soy broth and Methyl Red and Voges-Proskauer broth induced the production of known growth promoters in isolates 5 and 9. Methyl Red and Voges-Proskauer broth induced known growth promoters in isolates 1, 10 and 11 but tryptic soy broth did not. Tryptic soy broth induced known growth promoters in isolate 4 but Methyl Red and Voges-Proskauer broth did not. Liquid Murashige and Skoog medium did not induce the production of any growth promoters. Isolate 2 and isolate 12 did not produce any growth promoters across all media types. In some cases, volatiles were produced, but not at a high enough identity rate (90 %) to be included in the results. Following gas chromatography-mass spectrometry analysis, the volatile organic compound profile of each isolate under three different media types was formed shown below in table 8. The compounds that were known growth promoters as well as the compounds that were known fungal inhibitors were highlighted.

Table 8. Volatile organic compounds produced by each isolate under different media types. Volatiles that were also produced by controls (media with no isolates) were removed from the table. Red indicates a known growth promoter. Blue indicates a fungal inhibitor.

ISOLATE/MEDIA	MS	MRVP	TSB
ISOLATE 1	Propioloc acid C3H2O2 2-methyoxyamphetamine C10H15NO 1-Octanamine, N-methyl- C9H21N Carbamic acid, mono ammonium salt CH6N2O2	Nitrous oxide N2O Carbon dioxide CO2 L-alanine, ethylester C5H11NO2 Hydroxyurea CH4N2O2 Carbamic acid, monoamonium salt CH6N2O2 2-butanone, 3-hydroxy- C4H8O2 d-Alanine C3H7NO2 N-Hexylmethylamine C7H17 Propiolic acid C3H2O2	Benzenemethanol C9H13NO Cylclopropyl carbinol C4H8O Amphetamine-3-methyl Dodecane C12H26 Hexadecane C16H34 Pyrazine, 2,5-dimethyl. C6H8N2
ISOLATE 2	L-alanine ethylester C5H11NO2	Carbamic acid CH6N2O2 Ethylene oxide C2H4O Hydroxyurea CH4N2O2	Carbamic acid, monoamonium salt CH6N2O2 Ethylene oxide C2H4O 1-propanol, 2-amino-,(5)- C3H9NO Pyrazine 2,5-dimethyl. C6H8N2 Benzamethanop C9H13NO Hydroxyurea CH4N2O2
ISOLATE 4	Propiolic acid C3H2O2 2-propanamine, 1-methoxy C4H11NO L-alanine ethylester C5H11NO2 Hydroxyurea CH4N2O2 2-anamine C5H13N	Nitrous oxide N2O Carbon dioxide CO2 Carbamic acid, monoamonium salt CH6N2O2 L-alanine, ethylester C5H11NO2 R-(-)-1-Cyclohexylethylamine C8H17N Ethyne, fluoro- C2HF Propiolic acid C3H2O2 Silane, methyl CH6Si Ethylene oxide C2H4O Cyclopropapyl carbinol C4H8O Dextroamphetamine C9H13N	Carbamic acid, mono ammonium salt CH6N2O2 2-butanone, 3-hydroxy- C4H8O2 2,3-Butanediol C4H10O2 Cyclobutanol C4H8O Pyrazine 2,5-dimethyl- C6H8N2 Cyclopropyl carbinol C4H8O
ISOLATE 5		Carbamic acid, monoamonium salt CH6N2O2	Carbamic acid, mono ammonium salt CH6N2O2 Acetaldehyde C2H4O

		Ethylene oxide C2H4O Hydroxyurea CH4N2O2 1-propanol, 2-amino- C3H9NO Cyclopropapyl carbinol C4H8O Acetamide, 2-fluoro- C2H4FNO 2-butanone, 3-hydroxy- C4H8O2 2,3-Butanediol C4H10O2 2-propanamine 1-methods C4H11NO 1,3-Dioxolane C3H6O2	Acetamide, 2-fluoro C2H4FNO 2-butanone, 3-hydroxy- C4H8O2 Cyclobutanol C4H8O Pyrazine, 2,5-dimethyl- C6H8N2
ISOLATE 9	Carbon dioxide CO2 Carbamic acid, monoamonium salt CH6N2O2 L-alanine, ethylester C5H11NO2 Propiolic acid C3H2O2 Acetic acid, oxo- C2H2O3 Hydroxyurea CH4N2O2	Carbamic acid, monoamonium salt CH6N2O2 Propiolic acid C3H2O2 L-alanine, ethylester C5H11NO2 Ethyne, fluoro- C2HF Carbon dioxide CO2 1-propanol, 2-amino- C3H9NO Acetamide, 2-fluoro- C2H4FNO Silane, methyl- CH6Si Hydroxyurea CH4N2O2 Cyclopropyl carbinol C4H8O 2,3 butanediol C4H10O2	Carbamic acid, mono ammonium salt CH6N2O2 Ethylene oxide C2H4O Cyclopropyl carbinol C4H8O 2-butanone, 3-hydroxy- C4H8O2 Cyclobutanol C4H8O Pyrazine, 2,5-dimethyl- C6H8N2
ISOLATE 10	Nitrous oxide N2O	Carbamic acid, monoamonium salt CH6N2O2 Carbon dioxide CO2 L-alanine, ethylester C5H11NO2 Nitrous oxide N2O Ethyne, fluoro- C2HF 1-propanol, 2-amino- C3H9NO 2,3-Butanediol C4H10O2 (5)-(+)-1-Cyclohexylethylamine C8H17N	Nitrous oxide N2O

		d-Alanine C3H7NO2 2-propanamine, 1-methyoxy- C4H11NO Propiolic acid C3H2O2	
ISOLATE 11	Silane, tetramethyl- C4H12Si Nitrous oxide N2O Carbamic acid, mono ammonium salt CH6N2O2 Ethylene oxide C2H4O	Carbamic acid, monoamonium salt CH6N2O2 Ethylene oxide C2H4O Hydroxyurea CH4N2O2 2-butanone, 3-hydroxy- C4H8O2 Silane, methyl- CH6Si 2,3-butanediol C4H10O2 1,3-Dioxolane C3H6O2	Pyrazine 2,6-dimethyl C6H8N2
ISOLATE 12	Acetaldehyde C2H4O Acetamide, 2-fluoro C2H4FNO	Carbamic acid, monoamonium salt CH6N2O2 Hydroxyurea CH4N2O2 L-alanine, ethylester C5H11NO2 (R)-(+)-2-Amino-1-propanol C3H9NO 1-propanol, 2-amino- C3H9NO	L-Alanine, ethylester C5H11NO2 Carbamic acid, mono ammonium salt CH6N2O2 (s)-(+)-1-Cyclohexylethylamime C8H17N

4.4 LED treatment effect on plant growth

Following a 4-week treatment period under different wavelengths, measurements of stem length, leaf number, fresh weight and dry weight for each cultivar were taken.

In terms of leaf number, for Golden Wonder, the ANOVA showed a P-value of 5E-06 showing that wavelength had a statistically significant effect on average leaf number. Blue-red light at a ratio of 3:1 induced the largest number of leaves and outperformed all other wavelengths as can be seen in figure 9 below. This was a statistically significant positive result when compared to the control using an independent sample T-test with a P-value of 0.002. Red-blue light at a 1:1 ratio also had a statistically significant positive effect on leaf number (P-value: 0.047).

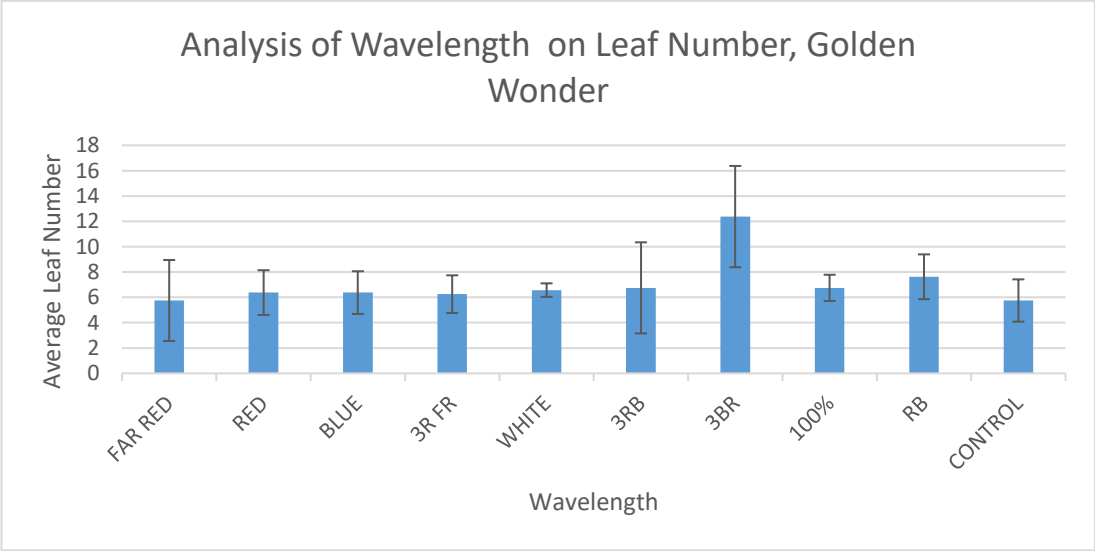

Fig 9. The bars show the average leaf number of 8 cuttings +/- 1 standard deviation under different wavelength treatments. FR = far-red. 3RFR= far-red-red (ratio 3:1). 3RB = red-blue (ratio 3:1). 3BR = blue-red (ratio 3:1). RB = red-blue (ratio 1:1). 100%= blue-red-far-red-white (ratio 1:1:1:1).

In Maris Piper, the ANOVA showed a P-value of 2.7E-10 indicating that wavelength had a statistically significant effect on average leaf number. Blue-red light at a ratio of 3:1 also performed the best in terms of leaf number (P-value: 0.022) but red-blue (3:1) and red-blue (1:1) also performed well as shown in figure 10 below. These were all statistically significant positive results when compared to the control using an independent sample T-test. White, blue and far-red wavelengths also had statistically significant positive effects.

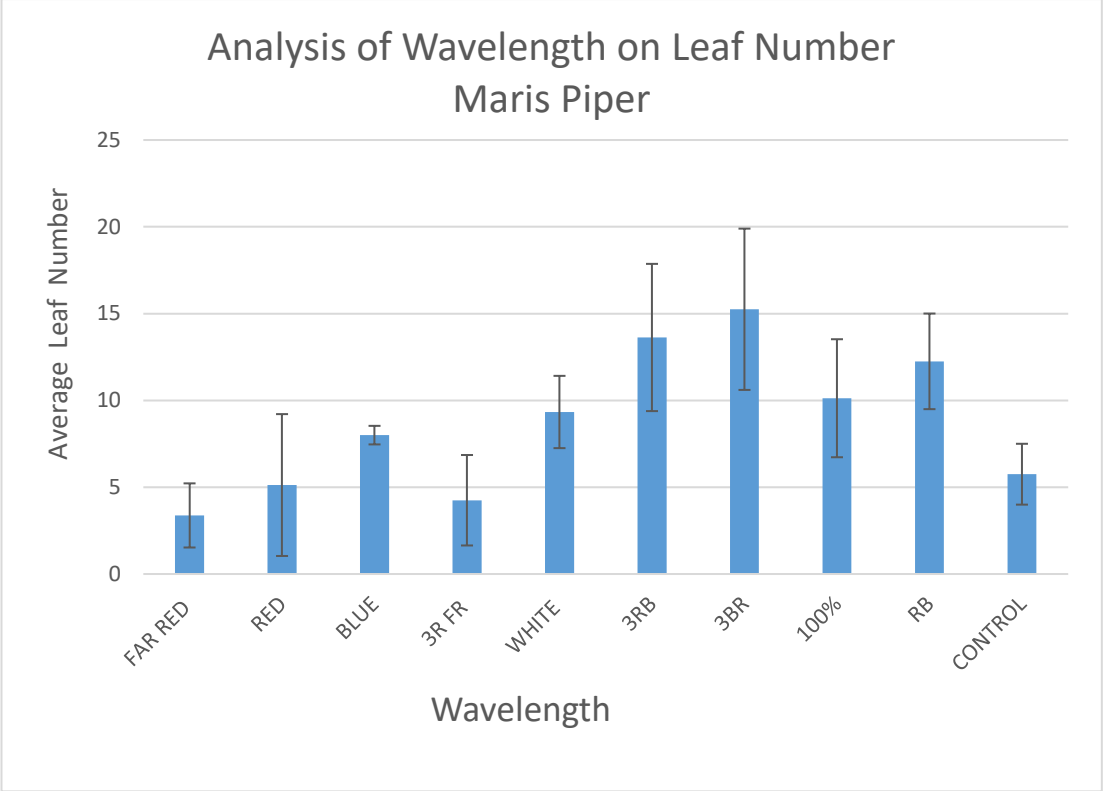

Fig 10. The bars show the average leaf number of 8 cuttings +/- 1 standard deviation under different wavelength treatments. FR = far-red. 3RFR= far-red-red (ratio 3:1). 3RB = red-blue (ratio 3:1). 3BR = blue-red (ratio 3:1). RB = red-blue (ratio 1:1). 100%= blue-red-far-red-white (ratio 1:1:1:1).

For stem length in Golden Wonder the ANOVA showed a P-value of 5.72E-13 indicating that wavelength had a statistically significant effect on stem length. Red

light induced the longest stems, with a 50/50 combination of far-red and red along with 3:1 red-blue having the second and third biggest impact on stem length respectively. When compared to the control, red light and far-red:red were statistically significant results (P-Values: 0.000 and 0.002), however, 3:1 red-blue was not statistically significant (P-value: 0.055). Blue-red (3:1) and red-blue (1:1) also had statically significant positive increase on stem length. Blue light and far-red light had statistically significant decrease on stem length when compared to the control using an independent sample T-test.

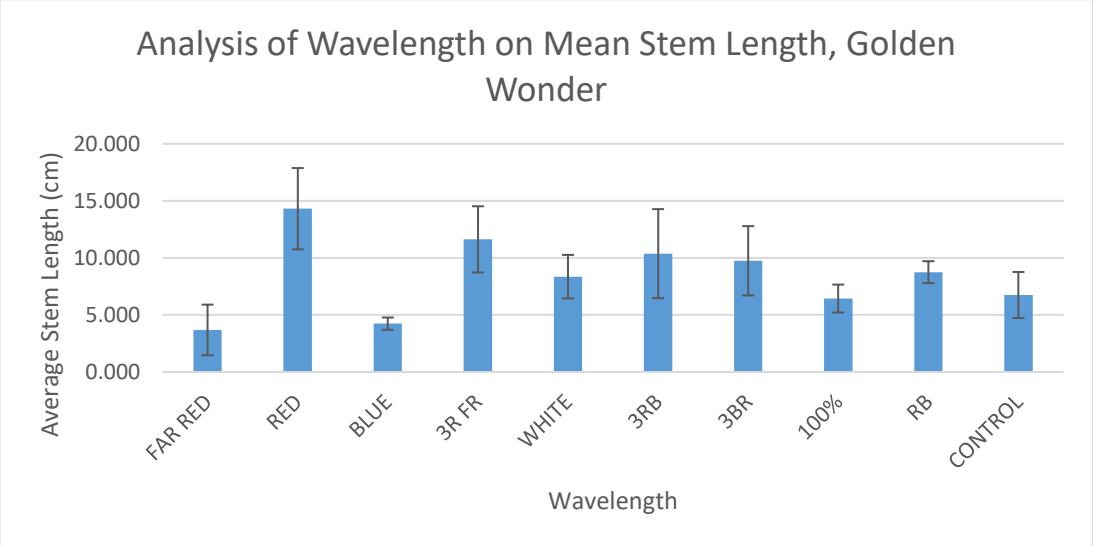

Fig 11. The bars show the average stem length of 8 cuttings +/- 1 standard deviation under different wavelength treatments. FR = far-red. 3RFR= far-red-red (3:1). 3RB = red-blue (ratio 3:1). 3BR = blue-red (ratio 3:1). RB = red-blue (ratio 1:1). 100%= blue-red-far-red-white (ratio 1:1:1:1).

With Maris Piper, the ANOVA gave a P-value of 0.003002 indicating that wavelength had a statistically significant effect on stem length. It was blue-red light (3:1) that saw the greatest impact on stem length, with 50/50 red-blue and 50/50 far-red-red also inducing length as shown in figure 12 below. However, when compared to the control using an independent sample T-test, blue-red (3:1) and red-blue (1:1) were statistically significant (P-values: 0.001 and 0.001), red-far-red (1:1) was not (P-value: 0.15). Red-blue (3:1) also had a statistically significant positive impact on stem length, while blue and far-red had a statistically significant negative impact.

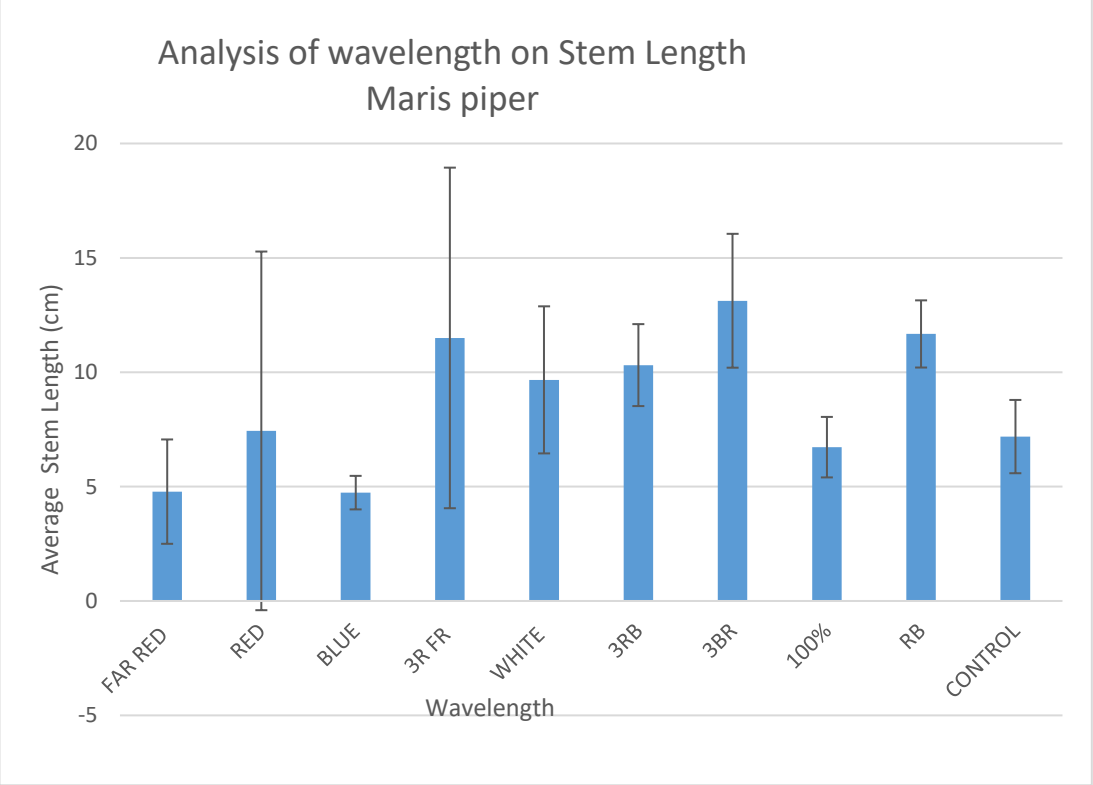

Fig 12. The bars show the average stem length of 8 cuttings +/- 1 standard deviation under different wavelength treatments. FR = far-red. 3RFR= far-red-red (ratio 3:1). 3RB = red-blue (ratio 3:1). 3BR = blue-red (ratio 3:1). RB = red-blue (ratio 1:1). 100%= blue-red-far-red-white (ratio 1:1:1:1).

In Golden Wonder, based on the ANOVA wavelength had a statistically significant effect on fresh weight and dry weight with P-values of 9.01E-09 and 1.68E-21 respectively. Red light had the greatest impact on biomass, while the combinations of red and blue as well as white light performed well shown in figures 13 and 14 below. For fresh weights, while red, white light, red-blue (1:1) and blue-red (3:1) were all statistically significant (P-values: 0.000, 0.001, 0.001, 0.001) positive increases on fresh weight when compared to the control using an independent sample T-test, red-blue (3:1) was not (P-value: 0.196). Combinations of all wavelengths (100%) and red-far-red also had significant positive effects on fresh weight. For dry weight, the only statistically significant effects when compared to the control were negative impacts, from far-red, red-far-red and white light (P-value: 0.000, 0.002, and 0.004).

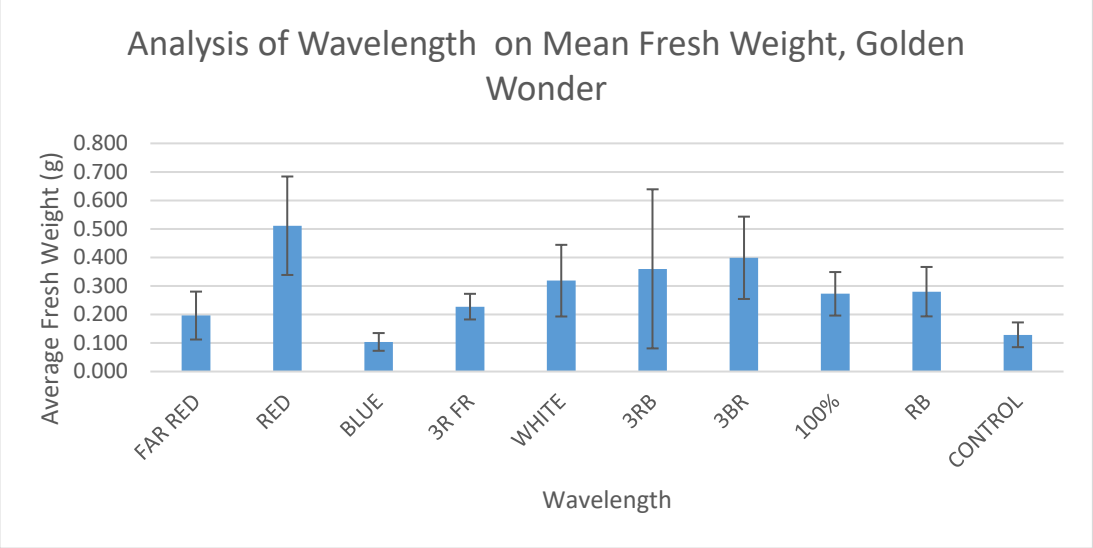

Fig 13. The bars show the average fresh weight of 8 cuttings +/- 1 standard deviation under different wavelength treatments. FR = far-red. 3RFR= far-red-red (ratio 3:1). 3RB = red-blue (ratio 3:1). 3BR = blue-red (ratio 3:1). RB = red-blue (ratio 1:1). 100%= blue-red-far-red-white (ratio 1:1:1:1).

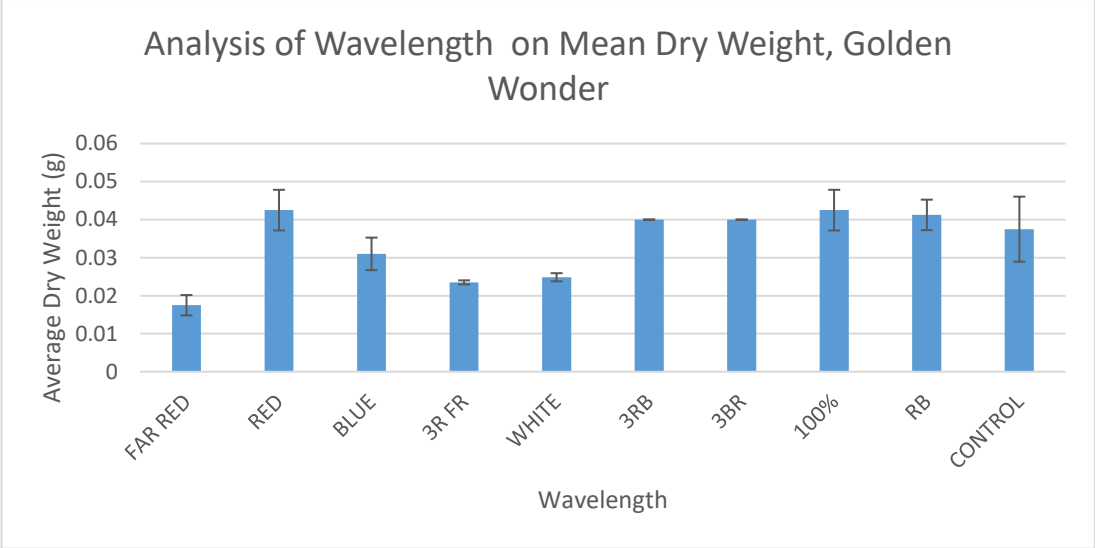

Fig 14. The bars show the average dry weight of 8 cuttings +/- 1 standard deviation under different wavelength treatments. FR = far-red. 3RFR= far-red-red (ratio 3:1). 3RB = red-blue (ratio 3:1). 3BR = blue-red (ratio 3:1). RB = red-blue (ratio 1:1). 100%= blue-red-far-red-white (ratio 1:1:1:1).

For Maris Piper, based on the ANOVA performed, wavelength had a statistically significant effects on fresh weight and dry weight with P-values of 1.92E-05 and 0.002582 respectively. Red-blue (3:1 ratio) performed the best in terms of fresh weight and dry weight as can be seen in figures 15 and 16 below. The combinations of red and blue, white light and the combination of all wavelengths (100%) also had a strong impact on biomass. Red-blue (3:1), blue-red (3:1), red-blue (1:1) and red were all statistically significant positive effects on fresh weight when compared to the control using an independent sample T-test (P-values: 0.000, 0.02, 0.045, 0.041), however white light was not statistically significant (P-value: 0.221). the combinations of red and blue light, red-blue (3:1), red-blue (1:1) and blue-red (3:1) all had a statistically significant positive impact on dry weight when compared to the control (P-value: 0.000, 0.009, 0.008). Far-red, blue and red-far-red light all had statistically significant negative impacts on dry weight when compared to the control (P-values: 0.000, 0.004, and 0.001).

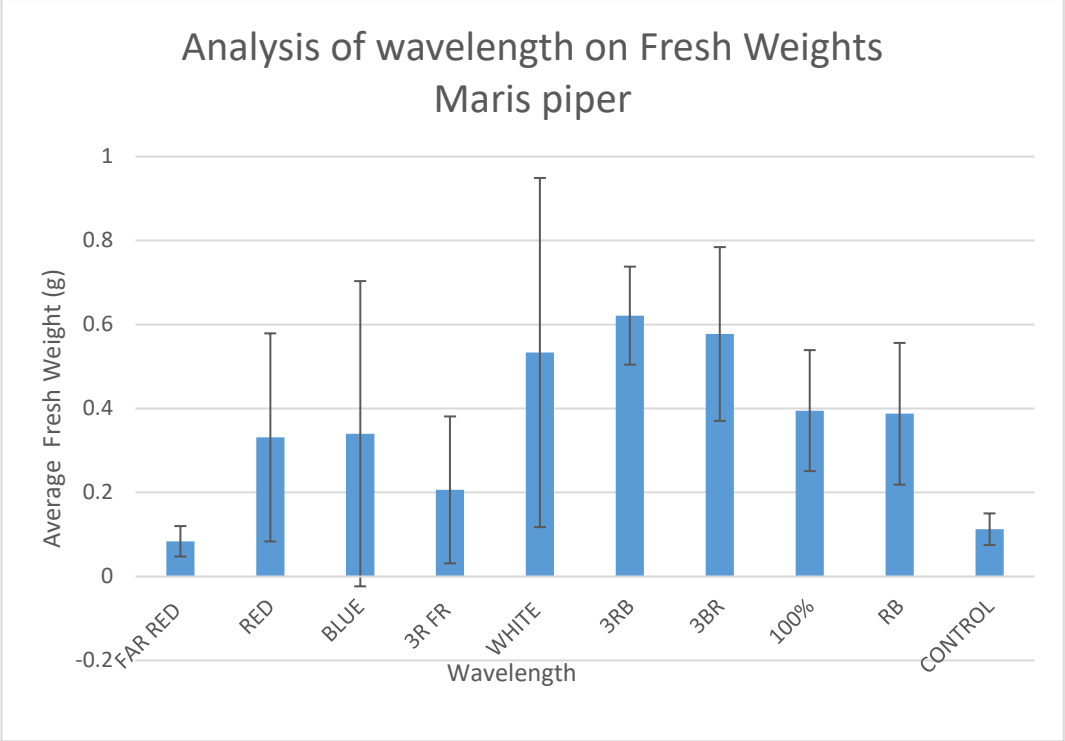

Fig 15. The bars show the average fresh weight of 8 cuttings +/- 1 standard deviation under different wavelength treatments. FR = far-red. 3RFR= far-red-red (ratio 3:1). 3RB = red-blue (ratio 3:1). 3BR = blue-red (ratio 3:1). RB = red-blue (ratio 1:1). 100%= blue-red-far-red-white (ratio 1:1:1:1).

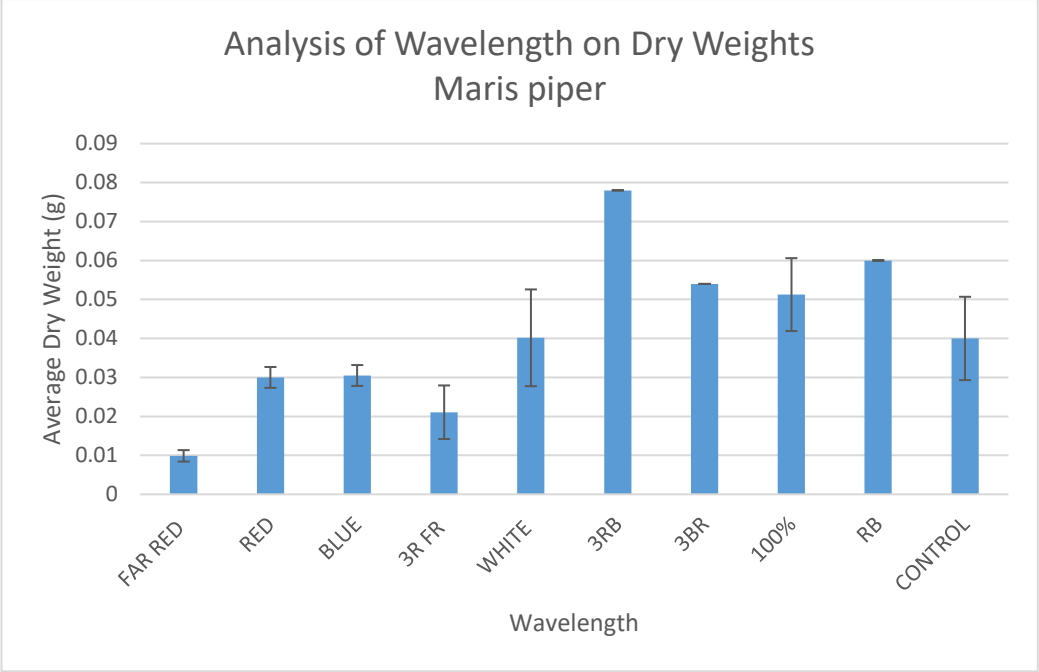

Fig 16. The bars show the average dry weight of 8 cuttings +/- 1 standard deviation under different wavelength treatments. FR = far-red. 3RFR= far-red-red (ratio 3:1). 3RB = red-blue (ratio 3:1). 3BR = blue-red (ratio 3:1). RB = red-blue (ratio 1:1). 100%= blue-red-far-red-white (ratio 1:1:1:1).

4.5 Effect of Bacterial isolates co-cultivation on plant growth

Following one week of growth and a further 10 days co-cultivation with each isolate, leaf number, stem length, fresh weight and dry weight were measured for each cultivar.

The ANOVA performed showed that the isolates had a statistically significant effect on leaf number with a P-value of 0.009958. Isolate 11 performed best in both cultivars. Isolate 4 also performed well in Maris Piper. However, isolate 11 was the only statistically significant result with P-values of 0.021 (GW) and 0.038 (MP) when compared to the control.

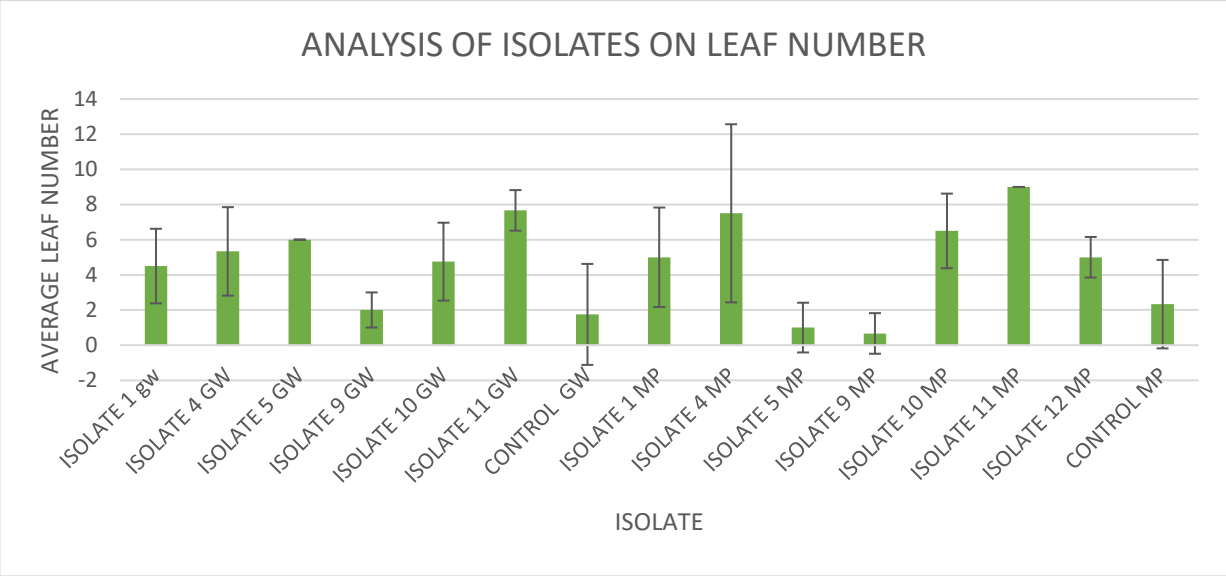

Fig 17. The bars show the average leaf number of 8 cuttings +/- 1 standard deviation under different isolate co-cultivation treatments. GW = Golden Wonder MP = Maris Piper

For stem length the ANOVA performed showed that the isolates did not have a statistically significant effect with a P-value of 0.173283. isolate 11 had the strongest impact on stem length. Isolate 4 also had a strong positive impact in Maris Piper, however only the Golden Wonder isolate 11 result was statistically significant with a P-value of 0.029 when compared to the control using an independent sample T-test. Both the Maris Piper isolate 11 and Maris Piper isolate 4 were not significant against the control (P-values: 0.748 and 0.944).

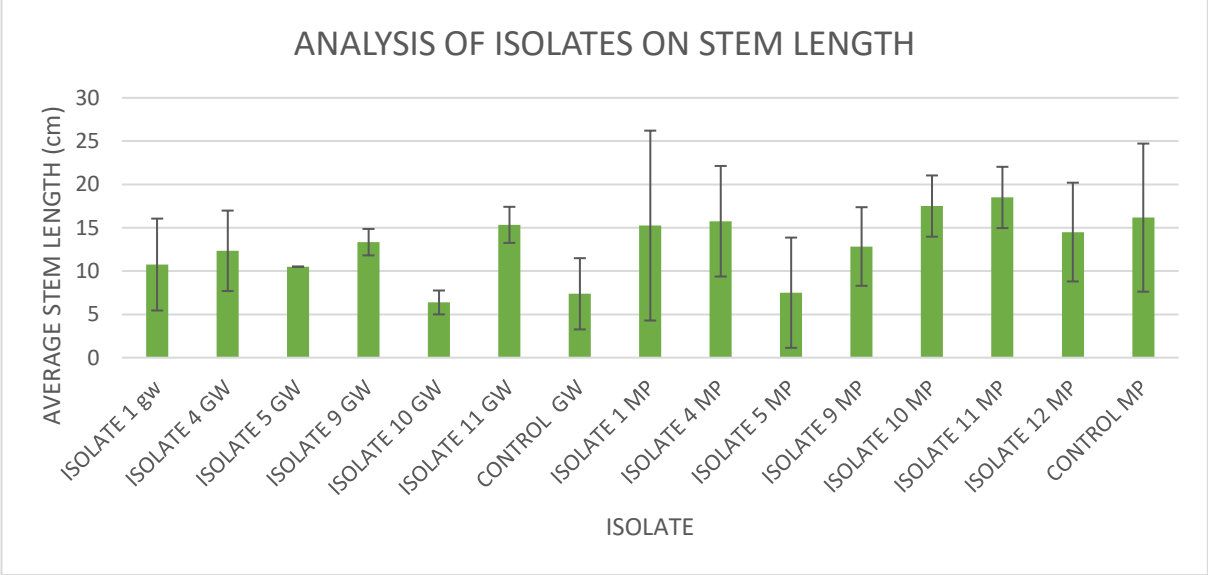

Fig 18. The bars show the average stem length of 8 cuttings +/- 1 standard deviation under different isolate co-cultivation treatments. GW = Golden Wonder MP = Maris Piper

In the fresh weight experiments the ANOVA performed showed that the isolates had a statistically significant effect on the fresh weights with a P-value of 0.023072. Once again isolate 11 had the greatest positive impact on fresh weight. In Golden Wonder, isolate 9 and isolate 5 also performed well. In Maris Piper, isolate 10 and isolate 4 were also strong performers. In Golden Wonder, although isolate 11 vastly outperformed all other isolates, it was not a statistically significant result when compared to the control using an independent sample T-test, with a P-value of 0.185. All other treatments were statistically significant against the control in terms of fresh weight in Golden Wonder. In Maris Piper, only isolate 11 had a statistically significant result with a P-value of 0.045.

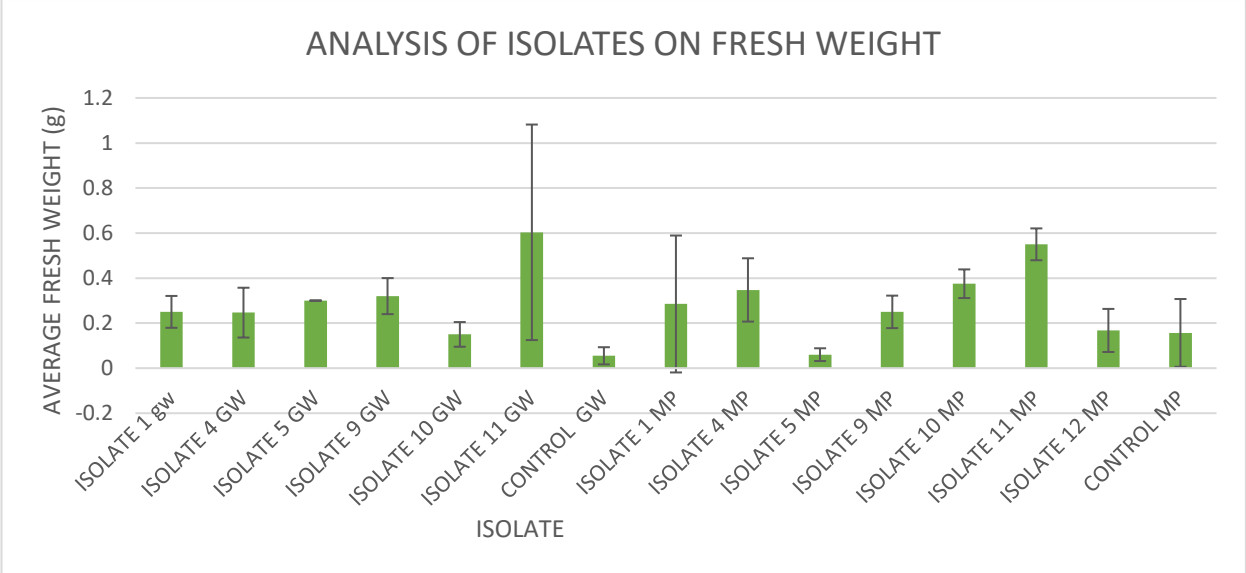

Fig 19. The bars show the average fresh weight of 8 cuttings +/- 1 standard deviation under different isolate co-cultivation treatments. GW = Golden Wonder MP = Maris Piper

In terms of dry weight, again isolate 11 outperformed all other treatments in both cultivars. Isolates 10 and 4 also performed well in Maris Piper. Due to the fact that

each plant from each treatment was pooled together to measure dry weights, independent sample T-tests and the ANOVA could not calculate P-values for the dry weight experiments as there was no variance between each sample since there was only one result for each treatment.

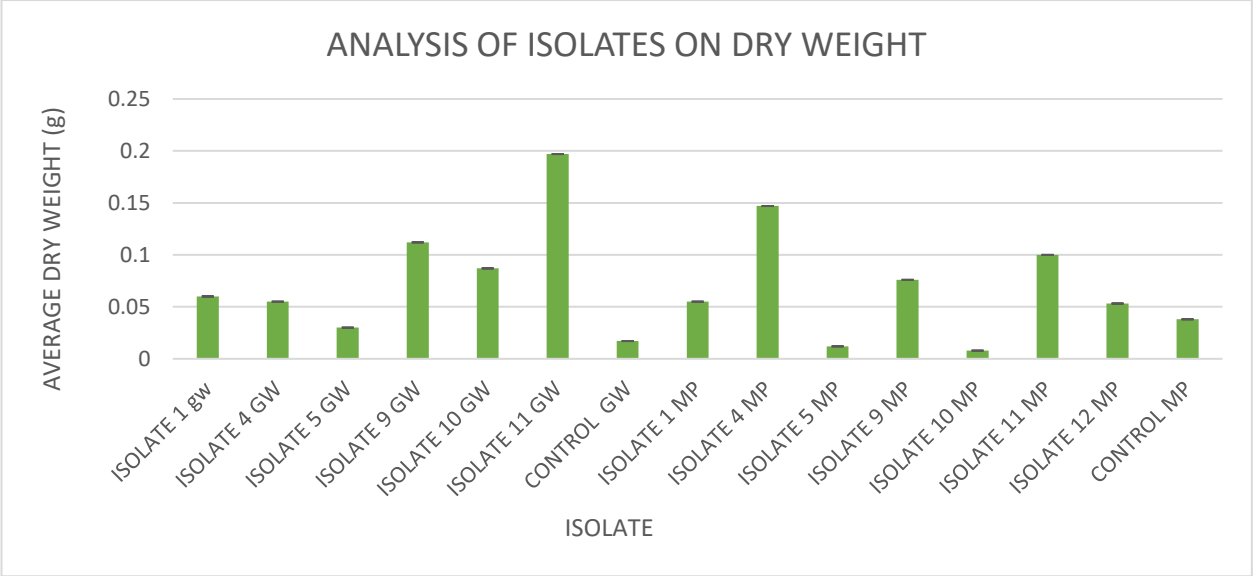

Fig 20. The bars show the average dry weight of 8 cuttings under different isolate co-cultivation treatments. GW = Golden Wonder MP = Maris Piper.

4.6 E.coli transformation

Post transformation, the *E.coli* containing the CRISPR vector and *E.coli* that had not been transformed (negative control) were streaked on LB agar plates containing 50 mg/L kanamycin and 10 mg/L tetracycline. Only *E.coli* which had been successfully

transformed could grow on these plates as they carried resistance to both kanamycin (from the CRISPR vector) and tetracycline (from the *E.coli*) as can be seen in figure 21.

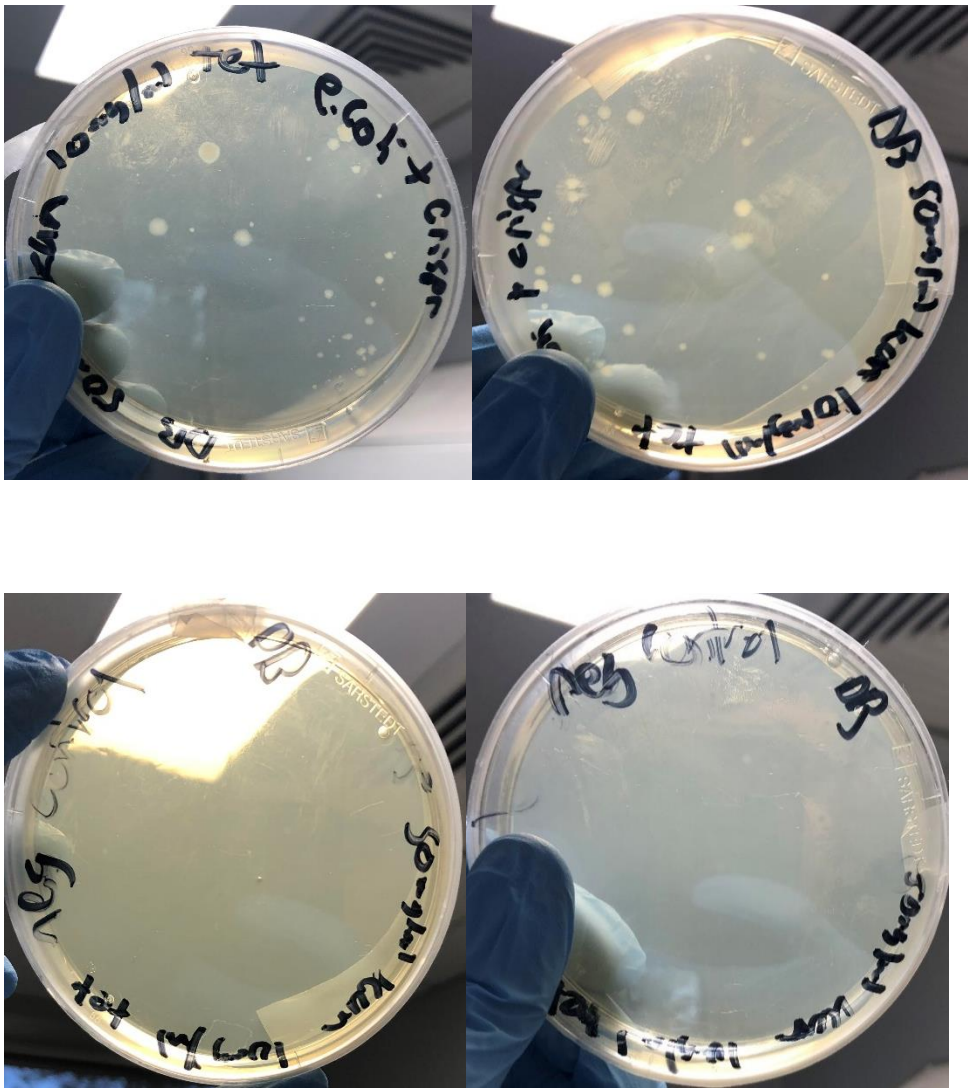

Fig 21. Transformed *E.coli* growing on LB plates containing 50 mg/L kanamycin and 10 mg/L tetracycline and non-transformed *E.coli* streaked on LB plates containing 50 mg/L kanamycin and 10 mg/L tetracycline showing no growth.

4.7 Plasmid extraction from transformed *E.coli*

The DNA concentrations in ng/µL are shown in table 9.

Table 9. DNA concentration post plasmid extraction

Sample I.D.	Plasmid mini prep DNA concentrations ng/µl
CRISPR 1 pGNK-LeCas9-AtUbp-gRNA	191.6
CRISPR 2 pGNK-LeCas9-AtUbp-gRNA	42.3
CRISPR 3 pGNK-LeCas9-AtUbp-gRNA	178.4
CRISPR 4 pGNK-LeCas9-AtUbp-gRNA	87.8
CRISPR 5 pGNK-LeCas9-AtUbp-gRNA	67.8
CRISPR guide 1 A p63(dicot)U6-gRNA:CMV-Cas9-beta-glucuronidase	173.4
CRISPR guide 1 B p63(dicot)U6-gRNA:CMV-Cas9-beta-glucuronidase	253.8
CRISPR guide 1 C p63(dicot)U6-gRNA:CMV-Cas9-beta-glucuronidase	150.7
CRISPR guide 1 D p63(dicot)U6-gRNA:CMV-Cas9-beta-glucuronidase	159.8
CRISPR guide 2 A p63(dicot)U6-gRNA:CMV-Cas9-beta-glucuronidase	198.7
CRISPR guide 2 B p63(dicot)U6-gRNA:CMV-Cas9-beta-glucuronidase	147.3
CRISPR guide 2 C p63(dicot)U6-gRNA:CMV-Cas9-beta-glucuronidase	150.7
CRISPR guide 2 D p63(dicot)U6-gRNA:CMV-Cas9-beta-glucuronidase	212.6
CRISPR guide 2 E p63(dicot)U6-gRNA:CMV-Cas9-beta-glucuronidase	198.8

CRISPR guide 3 A p63(dicot)U6-gRNA:CMV-Cas9-beta-glucuronidase	198.4
CRISPR guide 3 B p63(dicot)U6-gRNA:CMV-Cas9-beta-glucuronidase	144.9
CRISPR guide 3 C p63(dicot)U6-gRNA:CMV-Cas9-beta-glucuronidase	203.9
CRISPR guide 3 D p63(dicot)U6-gRNA:CMV-Cas9-beta-glucuronidase	233.8
CRISPR guide 3 E p63(dicot)U6-gRNA:CMV-Cas9-beta-glucuronidase	250.7
CRISPR Control	-0.6

Validation of successfully extracted pGNK-LeCas9-AtUbp-gRNA and p63(dicot)U6-gRNA:CMV-Cas9-beta-glucuronidase vectors was performed by running the vectors on a 2 % agarose gel next to a 100bp ladder to ensure the vector was the appropriate size. Gel electrophoresis results are shown in figures 22 and 23.

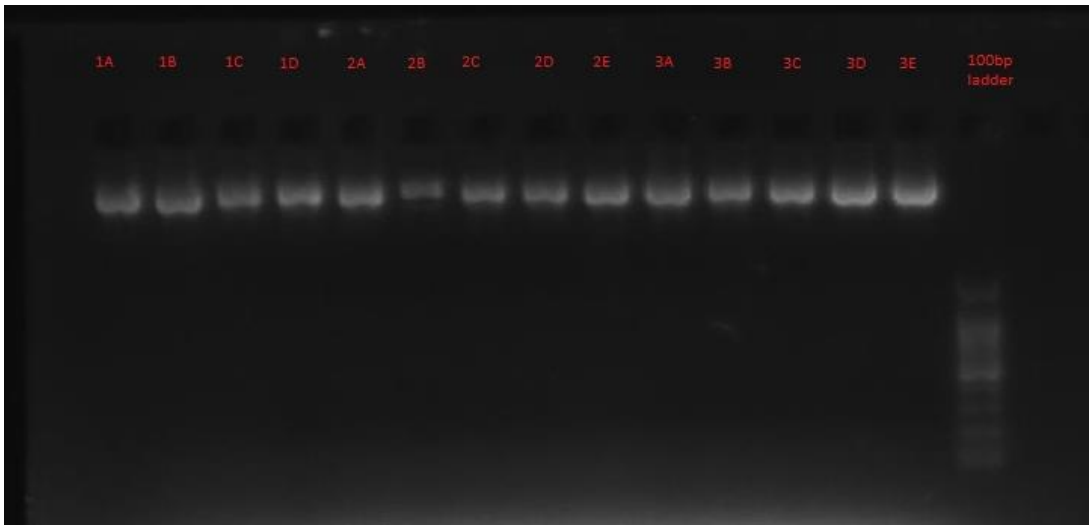

Fig 22. Gel electrophoresis of extracted p63 vectors 1A-D, 2A-E and 3A-E. The numbers refer to the guide present in the vector and the letters correspond to the concentrations of DNA post extraction in table 10.

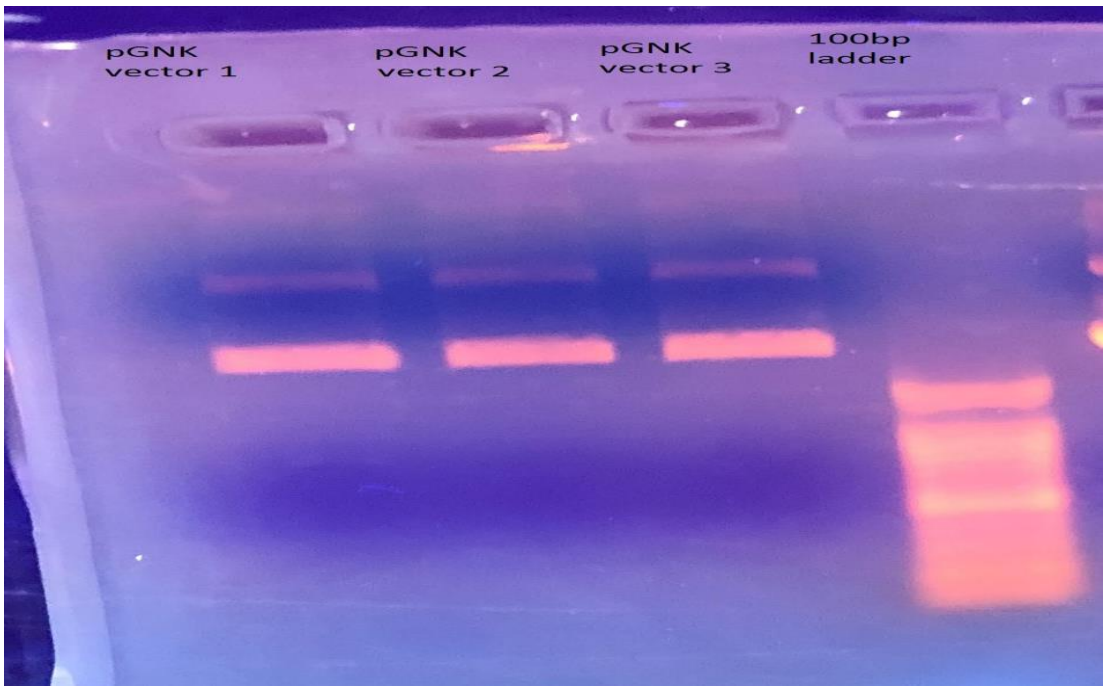

Fig 23. Gel electrophoresis of extracted vectors pGNK 1-3.

4.8 Validation of guide RNA through *in vitro* digestion

The extracted DNA from the *SGT3* region in potatoes was successfully digested by the synthetic cas9:sgRNA complex for all 3 guides. The gel electrophoresis results in figure 24 show that the DNA had been successfully digested when compared to the undigested control samples. This result validates the guides which were previously designed on CRISPRdirect.

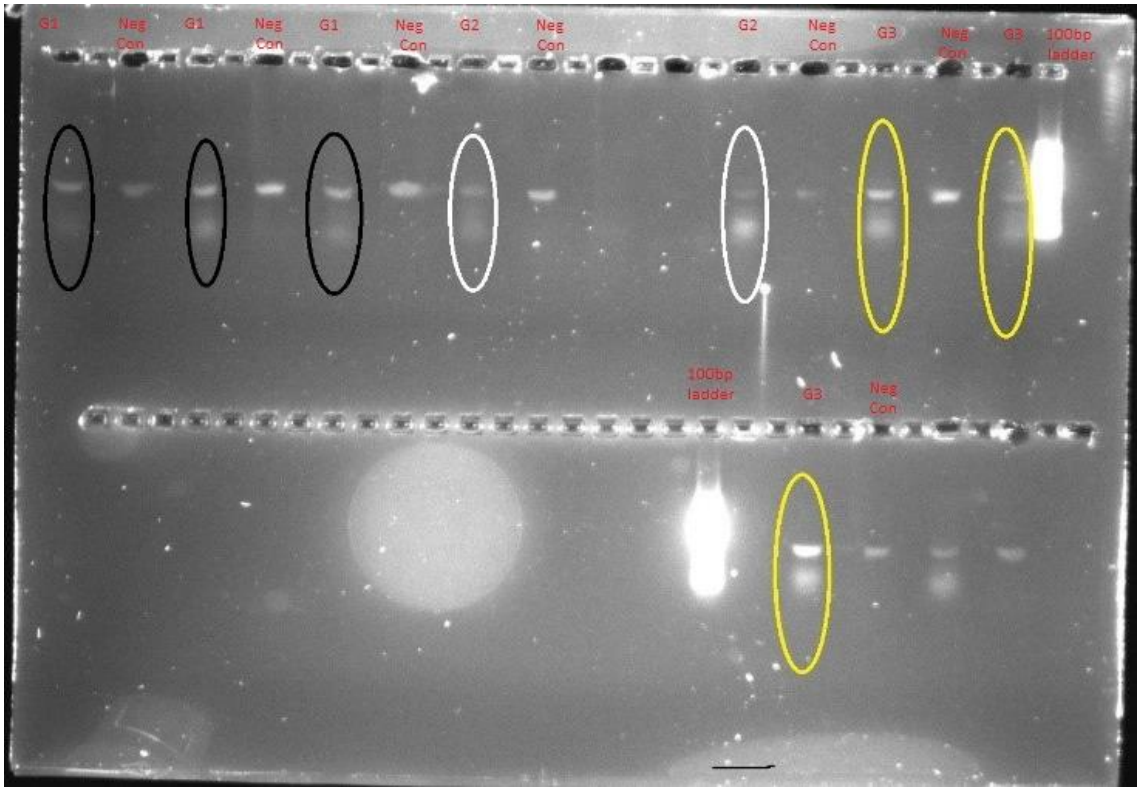

Fig 24. Gel electrophoresis of *in vitro* digestion showing digested DNA band next to undigested control. Colour indicates different guides (Black= digestion by Cas9 and guide 1, White= digestion by Cas9 and guide 2, Yellow= digestion by Cas9 and guide 3). G1 = guide 1, G2 = guide 2, G3 = guide 3, Neg con = negative control (DNA but no Cas9 or guide).

4.9 Multiple sequence alignment of the ligated vectors

Once the CRISPR vector (pGNK-LeCas9-AtUbp-gRNA) had been digested and purified, a ligation was carried out with the pre-validated guide sequences. Following ligation, the vectors were sequenced to detect the presence of the inserted guide RNA sequences. A multiple sequence alignment of the ligated vector and an uncut vector was performed using Clustal-Omega (Conway Institute, University College Dublin, Ireland). The multiple sequence alignment showed that the guide sequences had not been successfully inserted into the vector as can be seen in figure 25 below. Furthermore, the original sequence that should have been cut out of the vector by

the *BsaI* restriction enzyme was still present in the ligated vectors. This suggests that the vector was not digested properly, or that the vector was digested and not purified properly so that the sequence that was cut out was simply ligated back into the vector during the ligation process

```
CLUSTAL O(1.2.4) multiple sequence alignment


2           --TSTGGTAAAWTTTGCAAAAAARGCACCGACTCGGKGCCASTTTTKSAAGTTGATAACG 58
1           ----------TTTGCAAAAAAARGCACCGACTCGGTGCCACTTTTTSAAGTTGATAACG 49
4           ---ATCTGGAATTTTGCAAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACG 57
uncut       GGATCTGGAAATTTTGCAAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACG 60
3           ------GGAAATTTTGCAAAAAARGCACCGACTCGGTGCCACTTTTTSAAGTTGATAACG 54
                      **    ****** *********** **** ****  ***********

2           GASTAGCCTTAKTTTRACTTGCTATTTCTAGCTCKAAAACTGAGACCTTGGTCTCGAATC 118
1           GACTASCCTTATTTTAACTTGCTATTTCTAGCTCKAAAACTGAGACCTTGGTCTCRAATC 109
4           GACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACTGAGACCTTGGTCTCGAATC 117
uncut       GACTAGCCYTATTTTAACTTGCTATTWCTAGCTCTAAAACTGAGACCTTGGTCTCGAATC 120
3           GACTASCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACTGAGACCTTGGTCTCGAATC 114
            ** ** ** ** *** ********** ******* ****************** ****

2           ASTACTTSGACTCTAGCKGTATATAAACTCAGCTTCGTTTTCTTATCTAAGCGATGKGGG 178
1           ACTACTTCGACTCTAGCTGTATATAAACTCAGCTTCGTTTTCTTATCTAAGCGATGTGGG 169
4           ACTACTTCGACTCTAGCTGTATATAAACTCAGCTTCKTTTTCTTATCTAAGCGATGTGGG 177
uncut       ACTACTTCGACTCTAGCTGTATATAAACTCAGCTTCGTTTTCTTATCTAAGCGATGTGGG 180
3           ACTACTTCGACTCTAGCTGTATATAAACTCAGCTTCKTTTTCTTATCTAAGCGATGTGGG 174
            * ***** ********* ***************** ****************** ***

2           ACTTTTGAARAA 190
1           ACTTTTGAARAA 181
4           ACTTTTGAARAA 189
uncut       ACTTTTGAARAA 192
3           ACTTTTGAAGAA 186
            ********* **
```

Fig 25. Multiple sequence alignment of the ligated vectors with each guide and an undigested vector showing the desired insert is not present. 1 = guide 1, 2 = guide 2, 3 = guide 3, 4 = guide 4, uncut = undigested vector.

4.10 Gus assay

The Gus assay for the samples transformed with the p63 CRISPR vectors did not show any positive results. When the Gus assay was performed on the samples transformed with *Agrobacterium* that did not contain the CRISPR vector (115 in total), two of the treatments showed some positive results displayed in figure 26. Maris Piper heat-shocked samples showed a transformation efficiency of 16 % (3/18). Golden Wonder

LED (wavelength 3RB) showed a transformation efficiency of 50 % (3/6). The overall transformation efficiency across all treatments was 5.2 % (6/115).

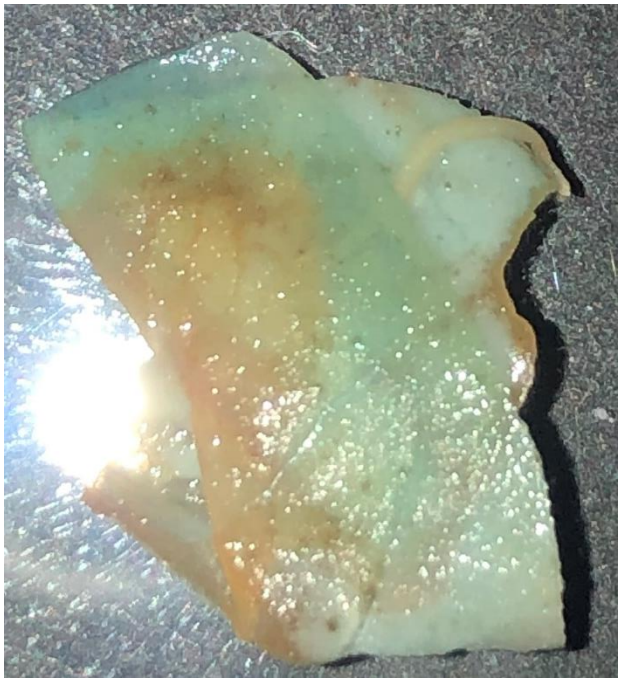

Fig 26. Sample successful transformed by AGL1 containing the gus gene post gus assay. Blue colour indicates successful transformed sample.

Fig 27. Control samples that had not been transformed by AGL1, incubated in gus solution overnight and bleached in ethanol for two days to remove chlorophyll.

4.11 Polymerase chain reaction to detect the presence of INDELs post co-cultivation transformation

In total, 62 plants were selected for sequencing (60 experimental plants, and two control plants). From these PCRs, 43 bands were successful excised, purified and used as templates DNA for Sanger sequencing (Figure 28).

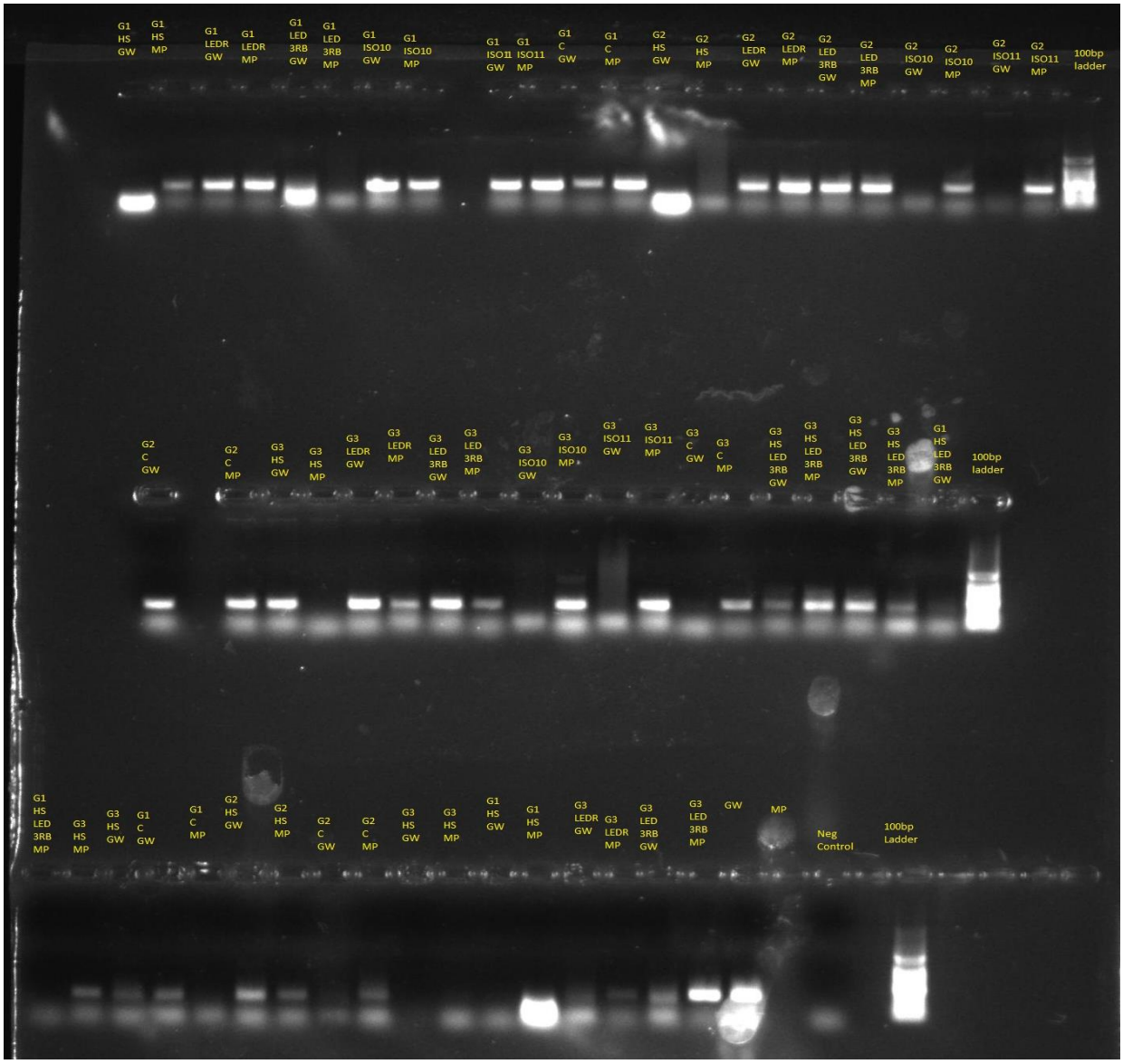

Fig 28. Gel electrophoresis of pcr product post co-cultivation transformation. G1 = guide 1, G2 = guide 2, G3 = guide 3, GW = Golden Wonder, MP = Maris Piper, HS = heat shock treatment, LEDR = LED red light treatment, LED3RB = LED red-blue (3:1) treatment, ISO10 = isolate 10 co-cultivation treatment, ISO11 = isolate 11 co-cultivation treatment, C = control (no treatment), Neg control = negative control (not transformed).

4.12 Sanger sequencing of purified DNA

Once the samples were sequenced, a multiple sequence alignment was performed comparing the sequence to a reference sequence of the *SGT3* region in potato using Clustal-Omega. The target region was then highlighted in order to see if this region had been affected.

As can be seen below in figure 29, most sequences appeared to be unaffected and no knockout was present.

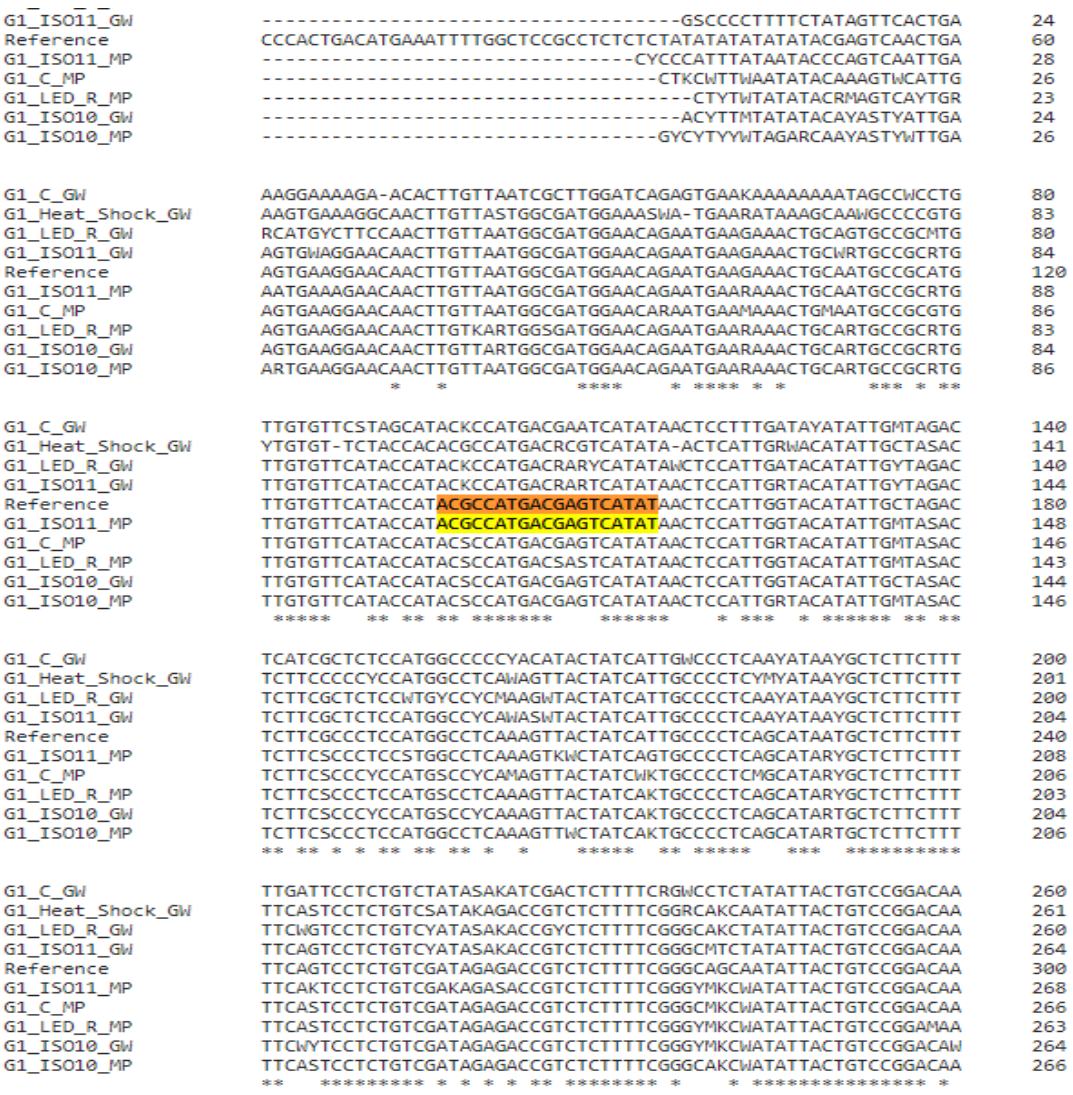

Fig 29. Multiple sequence alignment of samples post co-cultivation transformation. Guide 1 has been highlighted to show the target region.

Several of the sequences were cut short upstream of the targeted region. This may be due to sequencing error rather than any effect of the co-cultivation transformation.

Two of the samples showed significant gaps in their sequence. One of these gaps occurred in the target region as can be seen in figures 30 and 31 below. Several other gaps also appear elsewhere in the sequence. These sequences did not align well with the reference sequence.

Fig 30. Multiple sequence alignment of heat-shocked LED 3RB treated sample show some gaps in the sequence. Part of Guide 3 is highlighted to show the target region for this sample.

Fig 31. Multiple sequence alignment of LED 3RB treated sample show some gaps in the sequence. Part of Guide 3 is highlighted to show the target region for this sample.

5. Discussion

Genetic engineering of crops may prove to be an important part of global agricultures adaptation to climate change and help with issues such food security and increasing populations. This project aimed to use genetic engineering to create a knockout in the *SGT3* region of potato for crop enhancement and to investigate methods of improving this process.

5.1 LED treatment

After 4 weeks growth under each of the different experimental wavelengths, the recorded leaf number, stem length, fresh weight and dry weight measurements showed that there was a substantial difference between the effect of the different wavelengths and that the two cultivars responded differently to the different wavelengths. Interestingly far-red combinations performed well in stem length experiments but not in any other category. This is most likely due to the fact that far-red light mimics shade, encouraging the plant to grow taller to reach direct sunlight (Gelderen et al., 2018). To measure the effectiveness of the different wavelengths on the biomass on the plants' fresh weight and dry weight measurements were taken. In Golden Wonder, red light had the greatest impact on biomass. Clearly, in terms of Golden Wonder, red light was the key factor to any successful wavelength treatment. For Maris Piper, red-blue (3:1 ratio) performed the best in terms of fresh weight and dry weight. The combinations of red and blue and the combination of all wavelengths also had a strong impact on biomass. In both cultivars the combinations of red and blue light outperformed the control treatments of fluorescent lights in each category. In fact, the control treatment was one of the worst performing treatments in terms of leaf number, stem length, and fresh weight. In terms of dry weight, the control

treatment performed well in Golden Wonder. Overall the LED treatments, particularly the combinations of red and blue light, well outperformed the controls. This may be due to the longer photoperiod experienced by the LED treated plants of 22 hours on 2 hours off versus the 16 hours on 8 hours off in the fluorescent tube growth room. However the mix of red and blue light having the greatest positive impact on plant growth, particularly with regards to 3RB is in line with previously published studies (Chen et al., 2018). It is also worth noting that each cultivar had a slightly different reaction to the different wavelengths. Based on this analysis, red light and red-blue (3:1) were chosen as the wavelengths for the pre-transformation treatment. The impact of these wavelengths supports previously published studies showing the effect of different wavelengths on growth rates of potatoes and could help in designing future pre-transformation treatments (Chen et al., 2018). However, since the controls were grown in a different growth room under a different photoperiod, a direct comparison is limited. Light intensity is another factor which may have influenced the growth rate which was not looked at in this project. These factors should be considered in any future work.

5.2 Bacterial isolates

Of the 12 bacterial isolates isolated from the rhizosphere and headland of a commercial potato farm, eight isolates were successfully sequenced. All eight isolates were identified as *Bacillus* species of bacteria. This is an expected result as *Bacillus* species are well known soil rhizosphere bacteria, particularly in crops. *Bacillus* species have also previously been isolated and identified from potato rhizospheres (Calvo et al., 2010).

Solid-phase microextraction (SPME) Gas chromatography-Mass spectrometry was used to identify the volatile organic compounds being produced by these isolates across a variety of different media types. All but two of the isolates were shown to produce known plant growth promoters in at least one medium type, such as 2,3-butanediol (Ryu et al., 2003). Some isolates also produced known fungal inhibitors, such as isolate 1 which produced the hydrocarbons dodecane and hexadecane (Hughes et al., 2006).

Following the Gas chromatography-Mass spectrometry analysis, the isolates were co-cultivated with Maris Piper and Golden Wonder cultivars on the appropriate medium to induce the production of plant growth promoters. After 10 days of co-cultivation, growth was measured in terms of leaf number, stem length, fresh weight and dry weight. For each cultivar, isolate 11 outperformed all other isolates in all categories. This is likely due to the isolates ability to produce the growth promoters 3-hydroxy-2-butanone, and 2,3-butanediol (Ryu et al., 2003). In the Golden Wonder experiments, isolate 9 also performed well. This is to be expected as isolate 9 also produces known growth promoters and the identification experiments indicated that isolate 9 and isolate 11 are the same bacteria. For Maris Piper, isolates 4 and 10 also performed well in each category. Again, these isolates were shown to produce known plant growth promoters, so this result is to be expected (Ryu et al., 2003). Both cultivars did not grow when exposed to isolate 2 and Golden Wonder did not grow when exposed to isolate 12. Whether this was because of the exposure to the isolates is not clear, but these two isolates were also the only isolates that did not produce any growth promoters. However, several of the other isolates had less than all 4 sets of cultured plants grow, even though they did produce plant growth promoters. Contamination was also an issue for this experiment. The magentas had to be left open in order to expose the plant to the bacterial volatiles, however, this also made it easier for samples to be contaminated. The impact of plant growth promoting rhizospheric bacteria has been well characterised by previous studies and these experiments have reiterated the ability of rhizobacteria to produce growth promoters and to promote plant growth through co-cultivation (Ghyselinck et al., 2013).

5.3 *In vitro* digestion

The extracted DNA from the *SGT3* region in potatoes was successfully digested by the synthetic cas9:sgRNA complex. The gel electrophoresis results show that the DNA had been successfully digested when compared to the undigested control samples. All 3 guides showed successful digestion. This result validates the guides which were previously designed on CRISPRdirect as has been previously shown in other studies (Mehravar et al., 2019).

5.4 Vector and digestion

Once the CRISPR vector (pGNK-LeCas9-AtUbp-gRNA) had been digested and purified, a ligation was carried out with the pre-validated guide sequences. Following ligation, the vectors were sequenced to detect the presence of the inserted guide RNA sequences. A multiple sequence alignment of the ligated vector and an uncut vector was performed using Clustal-Omega (Conway Institute, University College Dublin, Ireland). The multiple sequence alignment showed that the guide sequences had not been successfully inserted into the vector. Because the ligation of gRNAs to the pGNK-LeCas9-AtUbp-gRNA vector failed, the digestion and ligation process was repeated several times with different parameters. However, the sequencing results continued to show that the digestion and ligation had been unsuccessful. It is likely that in order to successfully achieve this, a more complex approach to digestion and ligation would need to be taken which would prevent re-ligation of the vector to itself without uptake of the desired insert. Due to these difficulties, a new vector, p63(dicot)U6-gRNA:CMV-Cas9-beta-glucuronidase (*Agrobacterium* plasmid), was ordered from Sigma-Aldrich (St. Louis, Missouri, United States). This removed the need to digest the vector and ligate the guides to the vector as the custom guides were pre-inserted. This new vector had the additional benefit of containing a beta-glucuronidase gene allowing for a GUS assay to be performed in order to determine whether the *Agrobacterium* transformation had been successful.

5.5 Co-cultivation transformation

Following co-cultivation transformation of the samples with AGL1 containing each of the 3 guides, *SGT3* DNA was extracted, amplified, purified and sequenced. These sequences were compared to a reference sequence using a multiple sequence alignment in other to determine whether a knockout had occurred. For the majority of the samples, it was clear that no knockout had occurred. Several of the sequences appeared to be cut short upstream of the target region. This may be due to sequencing error rather than as a result of CRISPR activity. These sequences would need to be sequenced again using the reverse primer in order to see if this was a one-off error. Two of the samples had sequences which did not match the reference sequence well. The samples appeared to have gaps in their sequences. In one of the

samples, the gap occurs in the target region as shown in figure 30. In order to determine whether this is a result of CRISPR activity or sequencing error, the samples would need to be sequenced again with the reverse primer in order to form a consensus sequence. If the gaps in the sequence are a result of CRISPR deletions, then this would be an example of off-target activity and may have unknown effects of the *SGT3* gene. Off-target activity is a well-known limitation associated with CRISPR based transformations. (Zhang et al., 2015).

5.6 Pre-transformation treatments

As a pre-transformation treatment, the LED treatments did help increase plant growth, providing more plant material for transformation. It did not, however, appear to have a direct impact on transformation efficiency. Since there was no clear knockout, improving efficiency was determined by a Gus assay. The Golden Wonder sample treated with 3RB LED light showed the highest transformation efficiency of any of the treatments at 50 %. The overall efficiency was so low however, that it cannot be determined whether this is just by chance rather than a direct effect of the LED treatment as shown in previous studies (Das Bhowmik et al., 2019). The other set of successfully transformed samples came from heat shock treated plants. Again, this may or may not be a direct effect of the heat shock treatment, however heat shock treatment has previously been shown to improve transformation efficiency (Hwang et al., 2015). No bacterial isolate treated samples were successfully transformed with the gus gene.

6. Conclusion

This project set out to create a knockout in the *SGT3* gene region of *Solanum tuberosum* and to investigate the role of several pre-transformation treatments on transformation efficiency. Guide RNAs to target the *SGT3* region were successfully designed and validated through the *in vitro* digestion of *SGT3* DNA by the gRNA and Cas9 complex. The effect of several different LED wavelengths was examined and shown to have had a statistically significant impact on the leaf number, stem length,

fresh weight and dry weight of two different cultivars. Red light and combinations of red and blue light were shown to have a positive impact on plant growth rate. Eight soil rhizospheric isolates were successful isolated, sequenced, identified and volatile organic compound profiles for each isolate was determined. These isolates were then used to successfully determine the effect of these volatiles on leaf number, stem length, fresh weight and dry weight. Several of the isolates, mainly isolate 11, were shown to have a positive impact on plant growth rate. At this time, it is unclear whether a successful knockout has been created as a result of CRISPR activity and further experimentation is required to determine this.

7. Acknowledgments

I would like to thank my project supervisors Dr. Barbara Doyle-Prestwich and Dr. Eoin Lettice for their endless guidance, support and assistance. I would also like to thank Dr. Eileen Dillane for her technical support throughout the project. Finally, I would like to thank Darren Daly, Jack Daly and the Lab technicians in the Butler building for all their invaluable help.

Reference list

Barrangou, R. et al., 2007. CRISPR Provides Acquired Resistance Against Viruses in Prokaryotes. *Science,* 315(5819), pp. 1709-1712.

Bhattacharyya, P. N. & Jha, D. K., 2011. Plant growth-promoting rhizobacteria (PGPR): emergence in agriculture. *World Journal of Microbiology and Biotechnology,* 28(4), p. 1327–1350.

Bin Sayeed, M. S., Rezaul Karim, S. M., Sharmin, T. & Morshed, M. M., 2016. Critical Analysis on Characterization, Systemic Effect, and Therapeutic Potential of Beta-Sitosterol: A Plant-Derived Orphan Phytosterol. *Medicines,* 3(4), p. 29.

Callaway, E., 2018. CRISPR plants now subject to tough GM laws in European Union. *Nature ,* 560(16).

Calvo, P., Ormeño-Orrillo, E., Martínez-Romero, E. & Zúñiga, D., 2010. Characterization of Bacillus isolates of potato rhizosphere from andean soils of Peru and their potential PGPR characteristics. *Brazilian Journal of Microbiology,* 41(4), p. 899–906.

Chakravarty, B. et al., 2007. Genetic Transformation in Potato: Approaches and Strategies. *American Journal of Potato Research,* 84(4), pp. 301-311.

Chen, L. et al., 2018. Effects of red and blue LEDs on in vitro growth and microtuberization of potato single-node cuttings. *Frontiers of Agricultural science and Engineering,* 5(2), pp. 197-205.

Das Bhowmik, S. S. et al., 2019. Robust Genetic Transformation System to Obtain Non-chimeric Transgenic Chickpea. *Frontiers in Plant Science,* Volume 10, p. 524.

Deltcheva, E. et al., 2011. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature,* 471(7340), p. 602–607.

Doudna, J. A. & Charpentier, E., 2014. The new frontier of genome engineering with CRISPR-Cas9. *Science,* 346(6213).

Edwards, K., Johnstone, C. & Thompson, C., 1991. A simple and rapid method for the preparation of plant genomic DNA for PCR analysis.. *Nucleic Acids Research,* Issue 19, p. 1349.

FAOSTAT, 2016. *FAOSTAT Database.* [Online]
Available at: http://faostat3.fao.org/faostat-gateway/go/to/download/Q/QC/E

FAOSTAT, 2019. *Food and Agricultural Organization of the United Nations Statistical Database.* [Online]
Available at: http://www.fao.org/home/en

Friedman, M., 2006. Potato Glycoalkaloids and Metabolites: Roles in the Plant and in the Diet. *Journal of Agricultural and Food Chemistry,* 54(23), pp. 8655-8681.

Gelderen, K. v. et al., 2018. Far-Red Light Detection in the Shoot Regulates Lateral Root Development through the HY5 Transcription Factor. *The Plant Cell,* Volume 30, p. 101–116.

Ghosh, S. et al., 2018. Speed breeding in growth chambers and glasshouses for crop breeding and model plant research. *Nature Protocols,* Volume 13, p. 2944–2963.

Ghyselinck, J. et al., 2013. Bioprospecting in potato fields in the Central Andean Highlands: Screening of rhizobacteria for plant growth promoting properties. *Systematic and Applied Microbiology,* Volume 36, pp. 116-127.

Gupta, R. M. & Musunuru, K., 2014. Expanding the genetic editing tool kit: ZFNs, TALENs, and CRISPR-Cas9. *The Journal of Clinical Investigation,* 124(10), pp. 4154-4161.

Gurel, S. et al., 2008. Efficient, reproducible Agrobacterium-mediated transformation of sorghum using heat treatment of immature embryos. *Plant Cell Reports ,* 28(3), p. 429–444.

Hiei, Y., Ishida, Y., Kasaoka, K. & Komari, T., 2006. Improved frequency of transformation in rice and maize by treatment of immature embryos with centrifugation and heat prior to infection with Agrobacterium tumefaciens. *Journal of Plant Biotechnology,* 87(3), p. 233–243.

Hryhorowicz, M., Lipiński, D., Zeyland, J. & Słomski, R., 2017. CRISPR/Cas9 Immune System as a Tool for Genome Engineering. *Archivum Immunologiae et Therapiae Experimentalis,* 65(3), p. 233–240.

Hughes, K. A., Bridge, P. & Clark, M. S., 2006. Tolerance of Antarctic soil fungi to hydrocarbons. *Science of the Total Environment,* 372(2-3), pp. 539-548.

Hwang, H.-H.et al., 2015. Overexpression of the HspL Promotes Agrobacterium tumefaciens Virulence in Arabidopsis Under Heat Shock Conditions. *Phytopathology,* Volume 105, pp. 160-168.

Ishino, Y. et al., 1987. Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in Escherichia coli, and identification of the gene product.. *Journal of Bacteriology,* 169(12), p. 5429–5433.

Kennedy, A. C. & de Luna, L. Z., 2005. RHIZOSPHERE. *Encyclopedia of Soils in the Environment,* pp. 399-406.

King, J. C. & Slavin, J. L., 2013. White Potatoes, Human Health, and Dietary Guidance. *Advances in Nutrition,* 4(3), p. 393–401.

Makarova, K. S. et al., 2015. An updated evolutionary classification of CRISPR–Cas systems. *Nature Reviews Microbiology,* Volume 13, pp. 722-736.

Mali, P. et al., 2013. CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nature Biotechnology,* Volume 31, p. 833–838.

Mariot, R. F. et al., 2016. Characterization and Transcriptional Profile of Genes Involved in Glycoalkaloid Biosynthesis in New Varieties of Solanum tuberosum L.. *Journal of Agricultural and Food Chemistry,* 64(4), pp. 988-996.

McCue, K. F. et al., 2007. Potato glycosterol rhamnosyltransferase, the terminal step in triose side-chain biosynthesis. *Phytochemistry,* 68(3), pp. 327-334.

Mcmillan, M. & Thompson, J. C., 1978. An Outbreak of Suspected Solanine Poisoning in Schoolboys. *Quarterly Journal of Medicine,* 47(190), pp. 227-243.

Mehravar, M., Shirazi, A., Mehrazar, M. M. & Nazari, M., 2019. In Vitro Pre-validation of Gene Editing by CRISPR/Cas9 Ribonucleoprotein. *Avicenna Journal of Medical Biotechnology,* 11(3), p. 259–263.

Mojica, F. . J. M. & Rodriguez-Valera, . F., 2016. The discovery of CRISPR in archaea and bacteria. *The FEBS Journal,* 283(17), pp. 3162-3169.

Murashige, T. & Skoog, F., 1962. A Revised Medium for Rapid Growth and Bio Assays with Tobacco Tissue Cultures. *Physiologia Plantarum,* Volume 13.

Naito, Y., Hino, K., Bono, H. & Ui-Tei, K., 2015. CRISPRdirect: software for designing CRISPR/Cas guide RNA with reduced off-target sites. *Bioinformatics,* 31(7), p. 1120–1123.

Ni, W. et al., 2018. Maternal periconceptional consumption of sprouted potato and risks of neural tube defects and orofacial clefts. *Nutrition Journal ,* 17(112).

Nonaka, S. & Ezura, H., 2014. Plant–Agrobacterium interaction mediated by ethylene and super-Agrobacterium conferring efficient gene transfer. *Frontiers in Plant Science,* Volume 5, p. 681.

Păcurar, D. I. et al., 2011. Agrobacterium tumefaciens: From crown gall tumors to genetic transformation. *Physiological and Molecular Plant Pathology,* 76(2), pp. 76-81.

Patel, M., Dewey, R. E. & Qu, R., 2013. Enhancing Agrobacterium tumefaciens-mediated transformationefficiency of perennial ryegrass and rice using heat and highmaltose treatments during bacterial infection. *Journal of Plant Biotechnology,* Volume 114, p. 19–29.

Pattison, P. M., Tsao, J. Y., Brainard, G. C. & Bugbee, B., 2018. LEDs for photons, physiology and food. *Nature,* Volume 563, p. 493–500.

Ryu, C.-M.et al., 2003. Bacterial volatiles promote growth in Arabidopsis. *Plant Biology,* 100(8), pp. 4927-4932.

Schreiber, K., 2019. *Genetic Literacy Project.* [Online]
Available at: https://geneticliteracyproject.org/2019/02/19/infographic-are-genetically-engineered-crops-less-safe-than-classically-bred-food/

Shepherd, L. V. T. et al., 2015. Modifying glycoalkaloid content in transgenic potato – Metabolome impacts. *Food Chemistry ,* Volume 187, pp. 437-443.

Tovar, P. & Dodds, J. H., 1986. *Tissue Culture Propagation of Potato.* 5th ed. Ithaca, New York: International Potato Centre (CIP) .

Travella, S. et al., 2005. A comparison of transgenic barley lines produced by particle bombardment and Agrobacterium-mediated techniques. *Plant Cell Reports,* 23(12), p. 780–789.

United States Department of Agriculture, 2019. *Federal Register.* [Online]
Available at: https://www.federalregister.gov/documents/2019/06/06/2019-11704/movement-of-certain-genetically-engineered-organisms

United States Department of Agriculture, 2019. *United States Department of Agriculture Economic Research Service.* [Online]
Available at: https://www.ers.usda.gov/data-products/adoption-of-genetically-engineered-crops-in-the-us.aspx

Untergasser, A. et al., 2012. Primer3—new capabilities and interfaces. *Nucleic Acids Research,* 40(15), p. 115.

Wang, Q., Garrity, G. M., Tiedje, J. M. & Cole, J. R., 2007. Naïve Bayesian Classifier for Rapid Assignment of rRNA Sequences into the New Bacterial Taxonomy. *Applied and environmental microbiology,* 73(16), p. 5261–5267.

Wang, Z. et al., 2019. A Rhizosphere-Derived Consortium of Bacillus subtilis and Trichoderma harzianum Suppresses Common Scab of Potato and Increases Yield. *Computational and Structural Biotechnology Journal,* Volume 17, p. 645–653.

Zhang, X.-H.et al., 2015. Off-target Effects in CRISPR/Cas9-mediated Genome. *Molecular Therapy—Nucleic Acids,* Volume 4, p. 264.