Imaging green fluorescent protein in transgenic plants.

Jim Haseloff

MRC Laboratory of Molecular Biology, Hills Road, Cambridge. CB2 2QH. UK.

Marker genes have proved extremely useful for reporting gene expression in transformed plants. The ß-glucuronidase (GUS) gene has been used extensively [1]. Transformed tissues or patterns of gene expression can be identified histochemically, but this is generally a destructive test and is not suitable for assaying primary transformants, nor for following the time course of gene expression in living plants, nor as a means of rapidly screening segregating populations of seedlings. The green fluorescent protein (GFP) from the cnidarian jellyfish *Aequorea victoria* shares none of these problems, and there has been much interest in using the protein as a genetic marker in transgenic *Arabidopsis thaliana*.

Aequorea victoria are brightly luminescent, with glowing points around the margin of the jellyfish umbrella. Light arises from yellow tissue masses that each consist of about 6000-7000 photogenic cells. The cytoplasm of these cells is densely packed with fine granules that contain the components necessary for bioluminescence [2, 3]. In other bioluminescent coelenterates these have been characterised as 0.2 micron diameter particles enclosed by a unit membrane, and have been termed lumisomes [4]. The components required for bioluminescence include a Ca^{++} activated photoprotein, aequorin, that emits blue-green light, and an accessory green fluorescent protein (GFP), which accepts energy from aequorin and re-emits it as green light [5]. GFP is an extremely stable protein of 238 amino acids [6]. The fluorescent properties of the protein are unaffected by prolonged treatment with 6M guanidine HCl, 8M urea or 1% SDS, and two day treatment with various proteases such as trypsin, chymotrypsin, papain, subtilisin, thermolysin and pancreatin at concentrations up to 1 mg/ml fail to alter the intensity of GFP fluorescence [7]. GFP is stable in neutral buffers up to 65°C, and displays a broad range of pH stability from 5.5 to 12. The protein is intensely fluorescent, with a quantum efficiency of approximately 80% and molar extinction coefficient of 2.2 $x10^4$ cm^{-1} M^{-1} [5] (after correction for the known molecular weight). GFP fluoresces maximally when excited at 400 nm with a lesser peak at 475 nm, and fluorescence emission peaks at 509nm .

The intrinsic fluorescence of the protein is due to a unique covalently-attached chromophore which is formed post-translationally within the protein upon cyclisation and oxidation of residues 65-67, Ser-Tyr-Gly [6, 8, 9]. Several genomic and cDNA clones of *gfp* have been obtained from a population of *A. victoria* [6]. The *gfp* gene contains at least three introns, and the coding sequence derived from one of the cDNA clones, pGFP10.1 has been used for protein expression, first in *Escherichia coli*, *Caenorhabditis elegans* [9-11] and *Drosophila melanogaster* [12]. Fluorescent protein has now been produced in a number of heterologous cell types and there appears to be little requirement for specific additional factors for post-translational modification of the protein, which may be autocatalytic or require ubiquitous factors.

GFP expression in plants

GFP has been successfully expressed at high levels in tobacco plants using the cytoplasmic RNA viruses potato virus X [13] and tobacco mosaic virus [14]. In these experiments, the gene was directly expressed as a viral mRNA in infected cells, and very high levels of GFP fluorescence were seen.

In contrast to the efficient RNA virus-mediated expression of GFP, variable results have been obtained with transformed cells and plants. Although green fluorescence has been seen in *gfp* transformed protoplasts of citrus [15] and maize [16, 17], we and others have seen no fluorescence in *Arabidopsis* [18, 19]. Hu and Cheng [17] have reported that no signal was seen in *gfp* transformed *Arabidopsis thaliana* protoplasts. Reichel and colleagues also failed to dectect fluorescence in *gfp* transformed *Arabidopsis*, tobacco or barley protoplasts [20]. Sheen and colleagues also saw no expression of a CAB2-driven *gfp* gene in transgenic *Arabidopsis* plants [16], and Pang *et al.* [21] saw little or no expression in *gfp* transformed wheat, corn, tobacco and *Arabidopsis* plants. There appeared to be a need for substantial improvement of expression the wild-type *gfp* gene in plants.

The production of GFP fluorescence in plants requires that: (i) the GFP apoprotein be produced in suitable amounts within plant cells, and (ii) the non-fluorescent apoprotein undergoes efficient post-translational modification to produce the mature GFP. The high levels of GFP fluorescence seen in plants infected with suitable RNA virus vectors [13, 14] demonstrate that the protein can undergo efficient post-translation maturation in plants. It has now been shown that expression of the *gfp* is curtailed by aberrant mRNA splicing in *Arabidopsis*, and alteration of the codon usage of *gfp* is

required to avoid recognition of a cryptic intron, and allow proper expression of the fluorescent protein in this and other plant species.

Cryptic splicing of *gfp* mRNA in *Arabidopsis*.

In our first attempts to use the gene, we used *Agrobacterium* mediated transformation to produce transgenic *Arabidopsis* plantlets containing a 35S promoter-driven *gfp* cDNA. However, at no stage during the transformation procedure did we detect GFP-related fluorescence, using UV lamp illumination and epifluorescence microscopy. Therefore we used PCR-based methods to verify the correct insertion of the gene, and to check mRNA transcription and processing in these transformed plantlets. DNA and mRNA samples were separately extracted, and *gfp* sequences were amplified via PCR from the separate extracts and analysed. While the expected full-length *gfp* product was obtained after amplification of the integrated gene, RT-PCR of *gfp* mRNA sequences gave rise to a truncated product.

The shortened RT-PCR product was cloned and sequenced, and a deletion of 84 nucleotides was found between residues 380-463 of the GFP coding sequence (Fig. 1). The missing sequence bears close similarity to known plant introns and it is likely that expression of *gfp* in *Arabidopsis* is curtailed by aberrant mRNA splicing, with an 84 nucleotide sequence being recognised as a cryptic intron. This explanation would also account for the efficient expression of *gfp* from RNA virus vectors which replicate in the cytoplasm, and thus evade splicing. The nucleotide sequences bordering the deletion are shown in Figure 1, and demonstrate similarity to known plant introns. Matches were found for sequences that are conserved at the 5' and 3' splice sites of plant introns (reviewed in [22], and for conserved branchpoint nucleotides in plant introns [23, 24]. The excised *gfp* sequence contains a high AU content (68%) that has also been shown to be important for recognition of plant introns [25-28] . It is likely that this 84 nucleotide region of the jellyfish *gfp* cDNA sequence is efficiently recognised as an intron when transcribed in *Arabidopsis*, resulting in an in-frame deletion and the production of a defective protein product, which is predicted to be 28 amino acids shorter. It should be noted that the borders of this cryptic intron do not coincide with any of the natural spliced junctions found after processing of the *gfp* mRNA in *A. victoria* [6]. No full-length *gfp* mRNA is detectable by RT-PCR, and so misprocessing must be close to complete in transformed *Arabidopsis* plantlets.

Removal of the cryptic intron.

In order to destroy this cryptic intron we have altered the codon usage for GFP, deliberately mutating recognition sequences at the putative 5' splice site and branchpoint and decreasing the AU content of the intron. All of the sequence modifications affected only codon usage, and this modified gene, *mgfp4,* encodes a protein product which is identical to that of the jellyfish. When the *mgfp4* sequence was inserted behind the 35S promoter and introduced into *Arabidopsis* using the root transformation technique [29], bright green fluorescent plant cells were detected within 2-3 days of cocultivation. As cell proliferation continued, the brightest clumps of callus and developing shoot tissue were so intensely fluorescent that they were clearly visible by eye, using a 100 Watt long wavelength hand-held UV lamp (UV Products, B100AP). We have also adapted an inverted fluorescence microscope (Leitz DM-IL) to allow more sensitive, higher magnification observation of cells in sterile culture during transformation and regeneration. The microscope was fitted with a filter set (Leitz-D excitation BP355-425, dichroic 455, emission LP460) suitable for the main 395nm excitation and 509nm emission peaks of GFP, and we have used a 7mm threaded extension tube with a 4x objective (EF 4/0.12) to give a greater working distance above the microscope stage. This allows the convenient direct observation of transformed tissues and plantlets within sealed inverted petri dishes.

The ease with which fluorescent proteins can be monitored in living tissues allows new approaches for improving transformation and regeneration of intractable or slow-growing plant species. During our own regeneration experiments, we observed a wide range of GFP fluorescence intensities in 35S-*mgfp4* transformed plantlets, which we expect arose from position-dependent modulation of gene expression in different transformants. It proved difficult to regenerate fertile plants from the brightest transformants, with cells remaining as a highly fluorescent callus or mass of shoots after several months of culture. It is possible that high levels of GFP expression were mildly toxic or interfered with differentiation. This is of special concern with a fluorescent molecule such as GFP, which would be expected to generate free radicals upon excitation, and which undergoes oxidative modification and could possess catalytic properties. The conditions that we have used for plant regeneration should provide a stringent test for any deleterious effect due to GFP. The 35S promoter was used to drive expression of the protein at high levels throughout the plant, including meristematic cells, and regeneration took place under continual illumination, allowing the possibility for GFP mediated phototoxicity. Despite poor regeneration of the brightest transformants, we

managed to obtain over 50 separate transgenic *Arabidopsis* lines, most of which contained levels of GFP that were easily detectable by microscopy.

Expression of GFP in other plants.

Some *gfp* expression has been seen in plant protoplasts of tobacco [20], *Citrus sinensis* [15] and maize [16, 17], and so aberrant splicing of *gfp* mRNA may not be as efficient in other plant species, as in transgenic *Arabidopsis*. However, the *mgfp4* gene has proved useful for expression studies in other plants which share features involved in intron recognition [22]. Experiments with tobacco and barley protoplasts [20] have demonstrated that *mgfp4* derived sequences are expressed at much higher levels than the wild-type gene in these species. Improved GFP expression is seen also in mammalian cells after alteration of gene codon usage [30, 31] . Increased levels of expression have been attributed to improved rates of translation due to optimised codon usage. However, this "humanization" of *gfp* also leads to alteration of the cryptic intron sequence, and has been shown to result in 20-fold increased expression of GFP in maize protoplasts [32], and the increased levels of expression may be due to an effect on RNA processing.

It is possible that altered mRNA sequences affect post-transcriptional processing in animal cells. However, introns found in animals, including *A. victoria* [6], share a conserved polypyrimidine tract adjacent to the 3' splice site, reviewed in [33], and introns in yeast cells possess a requirement for additional conserved sequences (UACUAAC) located at the branch point [34]. The lack of these additional features may help to minimise aberrant processing of *gfp* mRNA in fungal and animal cells.

Localization of GFP in plant cells.

GFP expression and localization can be visualised directly, without a prolonged and lethal staining procedure, and laser scanning confocal microscopy can be used to optically section GFP-expressing tissues. Confocal imaging allows precise visualisation of fluorescent signals within a narrow plane of focus, with exclusion of out-of-focus blur (Fig. 2), and the technique permits the reconstruction of three dimensional structures from serial optical sections. Intact plant tissue proves a difficult subject for fluorescence microscopy as it consists of deep layers of highly refractile cell walls and aqueous cytosol and contains various autofluorescent and light scatttering components. There are two approaches to the difficulties imposed by these conditions: (1) to fix and to clear the

tissue with a high refractive index mounting medium, or (2) to directly image living tissue using suitably corrected microscope optics. In our experience, has proved difficult to effectively clear *Arabidopsis* wholemounts without causing artifacts or losing GFP fluorescence, and there are considerable advantages to working with living tissues, so we have pursued the second approach. *Arabidopsis* seedlings can simply be mounted in water for microscopy, and examined using a long-working distance water immersion objective to minimise the effects of spherical aberration when focusing deep into an aqueous sample [18]. Even with the use of such a specialised objective (Nikon 60x planapochromat, N.A. 1.2 , working distance 220 µm), image quality degrades rapidly for optical sections deeper than 60-80 µM within the tissue. However, the small size of *Arabidopsis* seedlings allows very useful imaging despite this limitation and, for example, median longitudinal optical sections can be obtained from intact roots.

In transgenic *Arabidopsis* cells, GFP is found throughout the cytoplasm, but appears to predominate within the nucleoplasm. It is excluded from vacuoles, organelles and other bodies in the cytoplasm, and is excluded from the nucleolus (Fig 3). A similar subcellular distribution of GFP was seen in all *Arabidopsis* cell types examined, and red autofluorescent chloroplasts provide an effective counter fluor for GFP in the upper parts of the plant. Cytoplasmic streaming and the movement of organelles could be observed in these living cells. In addition to cell ultrastructure, the architecture of the intact tissue was also clearly discernible, and the arrangement of different cell types could be seen in longitudinal optical sections of root tips and cotyledons. For example, cells within the epidermis of the cotyledon contain few mature chloroplasts and could be distinguished from layers of neighbouring mesophyll cells (Fig. 2), and files of developing cells around the primary root meristem were clearly evident.

While the *mgfp4* gene was proving useful as a marker in transgenic *Arabidopsis*, it was also clear from the initial studies that it could bear improvement. While we were able to generate 35S-*mgfp4* transformed cells that were intensely fluorescent, and easily detectable by eye under long wavelength UV illumination, it proved difficult to regenerate fertile plants from the brightest transformants. It is possible that very high levels of GFP expression are mildly toxic or interfere with regeneration, perhaps due to the fluorescent or catalytic properties of the protein. In jellyfish photocytes, where high levels of GFP are well tolerated, the protein is found sequestered in cytoplasmic granules [2]. In contrast, the mature protein is found throughout the cytoplasm and accumulates within the nucleoplasm of transformed *Arabidopsis* cells. If GFP is a source of

fluorescence-related free radicals, for example, it might be advisable to target the protein to a more localised compartment within the plant cell.

ER localisation of GFP

We have fused several targeting peptides to GFP, and directed the protein to different subcellular compartments. The targeted forms of the *mgfp4* gene were initially tested by expression in *Saccharomyces cerevisiae*. The modified genes were introduced into yeast cells on a multicopy vector and expressed fluorescent protein was visualised using confocal microscopy. Unmodified protein is normally found throughout the cytoplasm and nucleoplasm of yeast cells, whereas targeted forms of the protein showed distinct limited distribution of the protein within the yeast cells. The modified forms of the protein were then placed behind the 35S promoter, introduced into *Arabidopsis* via *Agrobacterium* mediated transformation. It is possible to detect GFP fluorescence in primary transformants, and we examined the distribution and fluoresence intensity of GFP in over 25 transformed plantlets for each construction. The only variant that showed a substantial improvement over unmodified GFP was one that was targeted to the endoplasmic reticulum (ER). This targeted form of GFP contains an N-terminal signal peptide derived from an *Arabidopsis* vacuolar basic chitinase and the C-terminal amino acid sequence HDEL, to ensure secretion and retention of the protein within the lumen of the ER. Using this modified gene (*mgfp4-ER*), it has been possible consistently to regenerate fertile plantlets that are brightly fluorescent. GFP fluorescence within these plants could be readily observed by eye using a long wavelength UV lamp. The *mgfp4-ER* expressing plants were examined by confocal microscopy, and fluorescent protein was found mainly within the endomembrane system. The protein is excluded from the nucleus, shows a perinuclear distribution, and is found associated with the ER which forms a characteristic reticulate network in highly vacuolate cells (Fig. 3). In highly cytoplasmic meristematic cells, the nuclei and orientation of cell divisions can be clearly distinguished. Localisation of the modified protein to cytoplasmic organelles was also evident, to what appear to be leucoplasts or proplastids. For example, an optical section of a hypocotyl epidermal cell is shown in figure 3 and this includes a thin portion of cytosol which is pressed between the cell wall and vacuole. Such hypocotyl cells in *mgfp4-ER* transformed seedlings appear to contain a spectrum of developing plastids that range from the brightly green fluorescent to those which take on a yellow, orange or red appearance in dual channel confocal micrographs. We presume that this is due to increasing chlorophyll synthesis, and that the green fluorescent plastids may be the

maturing precursors of chloroplasts in these cells. These green fluorescent plastids are also found within the chloroplast-free epidermal cells of leaves and cotyledons, but are not found within the underlying mesophyll cells that are packed with mature chloroplasts. It seems likely that these organelles are proplastids and are capable of developing into chloroplasts, but we cannot exclude the possibility that they are some specialised form of plastid or other organelle.

The accumulation of *mgfp4-ER* protein within leucoplasts or developing proplastids, in addition to its entry into the secretory pathway and retention in the endoplasmic reticulum, may indicate misrecognition of the N-terminal signal peptide. Proplastid accumulation of GFP is not seen in the 35S-*mgfp4* transformed plants. If the *mgfp4-ER* encoded signal peptide is inefficiently recognised prior to docking and cotranslational transport of the protein into the lumen of the ER, a proportion of GFP bearing fused terminal sequences may be produced in the cytoplasm. If so, it is possible that the neglected signal peptide may act as a transit sequence for plastid entry. Alternatively, there may be some direct exchange between developing plastids and the endomembrane system. We see no free cytoplasmic fluorescence, and the protein is sorted very efficiently to the ER or to plastids.

It is unclear whether the beneficial effects of targeting GFP to the ER are due to increased levels or safer accumulation of mature GFP within cells. For example, if accumulation of fluorescent protein leads to the generation of free radicals in illuminated cells, it is conceivable that removing GFP from the nucleus could protect cells from DNA damage due to such short-range highly reactive species. However, it is also possible that the fusion of peptide targeting sequences may improve the properties of the protein itself, or that the localisation of GFP to the lumen of the ER may improve its maturation and accumulation. The maturation of the GFP apoprotein is sensitive to temperature, and the apoprotein readily misfolds under certain conditions [35]. The lumen of the ER is known to contain components, such as chaperones and peptidyl prolyl isomerases which aid protein folding [36], and secretion and retention of GFP within the ER may allow improved formation and accumulation of the mature fluorescent protein.

Thermosensitivity of GFP

Expression of GFP in mammalian cells has been described as highly variable [37], often requiring a strong promoter and decreased incubation temperature for good results [38-40]. Other researchers have found that development of fluorescence is similarly

favoured by a lower incubation temperature during expression of GFP in bacteria [41] and yeast [42]. These observations suggest that expression of GFP in cells that require higher incubation temperatures may be far from optimal. To date, a number of workers have isolated mutant forms that fluoresce many-fold more intensely than wild-type GFP when expressed in *E. coli* at 37°C. Below, I describe our own experiments that have resulted in the production of improved GFP forms, specifically tested for use in plants. We have demonstrated that increased fluorescence results primarily from suppression of a temperature-dependent defect in the folding of the GFP apoprotein. We have also manipulated the fluorescence spectra by site-directed mutagenesis, to produce a thermostable folding mutant, that can be efficiently excited using either long wave UV or blue light.

The *mgfp4* gene, which had been modified to remove a cryptic intron seen during expression in *Arabidopsis*, was subjected to random mutagenesis. A library of 10,000 mutant genes was introduced into *E. coli* and expressed overnight at 37°C and screened for increased fluorescence. Several mutants were isolated. The mutations responsible for bright phenotypes were mapped by recombination with the wild-type *mgfp4* gene. Sequencing of the brightest mutant (GFPA) revealed two amino-acid differences, V163A and S175G. The mutant GFP produces up to 35-fold increased fluorescence in bacterial cells, while the difference in protein levels is not nearly enough to account for this. The result suggested that a large proportion of GFP that is expressed in cells at 37°C is non-fluorescent.

Experiments with a GFP-nucleoplasmin fusion protein have indicated that maturation of GFP to the fluorescent form may be sensitive to temperature during expression in the yeast *Saccharomyces cerevisiae* [43]. To test whether the same could be true of expression in *E. coli* and whether the substitutions present in GFPA enhance maturation by suppressing any such sensitivity, we examined expression of GFP and GFPA over a range of different temperatures. Strains expressing GFP or GFPA were grown overnight at temperatures ranging between 25°C and 42°C. For each culture, the fluorescence values were measured and normalised against the amount of recombinant protein present in the cells to give a measure of the proportion of intracellular GFP that is fluorescent at different temperatures (Fig 4A). The proportion of GFP that is fluorescent steadily decreases with increasing incubation temperature, indicating that either mature GFP or the maturation pathway leading to its formation is temperature sensitive. Mature GFP is a highly stable molecule whose fluorescence *in vitro* is unaffected by temperatures up to 65°C [7], and we have confirmed that the fluorescence of the mature

protein is unaltered in bacterial cells at 42°C. Therefore, we concluded that higher incubation temperatures interfere with the post-translational maturation of GFP, rather than causing inactivation of the mature protein. Moreover, the primary effect of the substitutions present in GFPA is to enhance the proportion of GFP that is fluorescent at higher temperatures, rather than simply enhancing the intrinsic fluorescence properties of mature GFP. We confirmed that expression of GFP is also temperature-sensitive in yeast and demonstrated that this is suppressed by the substitutions present in GFPA. These results indicate that the thermosensitivity of GFP maturation may be a common phenomenon that can be suppressed by the amino acid substitutions present in GFPA [35].

Chromophore oxidation during maturation of GFP and GFPA

The post-translational maturation of GFP to the fluorescent form involves a number of steps [8, 9, 44]. In the first step, the GFP apoprotein presumably must fold into a catalytically active conformation that facilitates the novel reactions involved in formation of the chromophore. These reactions consist of cyclisation and oxidation of the tripeptide Ser65-Tyr66-Gly67 to give a p-hydroxybenzylidene-imidazolidinone structure. The mature protein must then be correctly folded to maintain its fluorescent properties, to protect the chromophore from solvent effects [43] In principle, any of these processes could be sensitive to temperature and thus be responsible for the observed thermosensitivity of GFP during maturation. Since the oxidation reaction involved in chromophore formation appears to require molecular oxygen, Heim *et al*. [9]were able to measure the reaction rate by expressing GFP in *E. coli* under anaerobic conditions and then monitoring the development of fluorescence after admission of air. To determine whether this reaction is temperature sensitive and whether the substitutions present in GFPA act by enhancing its rate at higher temperatures, we measured the rates of oxidation of GFP and GFPA expressed in yeast at both 25°C and 37°C. The time constant measured for the oxidation of GFP at 37°C (5.9 ± 0.1 min) was found to be approximately 3-fold faster than that measured at 25°C (16.2 ± 0.3 min), indicating that the post-translational oxidation of the GFP chromophore is not the step responsible for the temperature sensitivity of maturation. In confirmation of this conclusion, the time constants derived for GFPA at both 25°C and 37°C (22.5 ± 1.4 min and 18.1 ± 0.4 min, respectively) were actually slower than those measured for GFP [35].

Apo-GFP folds improperly at elevated temperatures

The improper folding of proteins often results in their aggregation into insoluble inclusion bodies during expression in *E. coli* [45]. To determine whether the proper folding of GFP might be temperature sensitive and whether the substitutions present in GFPA act by enhancing proper folding at increased temperatures, we examined the solubilities of the two proteins during expression in *E.coli* at 25°C and 37°C. Fluorescence was found almost exclusively in the soluble fraction. At 25°C, both GFP and GFPA were found predominantly in the soluble fraction, indicating that proper folding of both proteins is relatively efficient at this temperature. At 37°C, however, the majority of GFP was found as non-fluorescent protein in the insoluble fraction, whereas most of GFPA was still present in the soluble fraction. This result indicates that the temperature sensitivity of GFP maturation is due primarily to improper protein folding at higher temperatures, and that this defect is suppressed by the amino acid substitutions present in GFPA.

To obtain information on which species in the maturation pathway of GFP misfolds at higher temperatures, we examined the absorption spectrum of denatured protein isolated from inclusion bodies. If GFP undergoes cyclisation of the chromphore prior to aggregation, protein from inclusion bodies should show an absorption in the near UV/blue region that is characteristic of the GFP chromophore in either the mature or chemically reduced state [46, 47] On the other hand, if unmodified GFP (apo-GFP) is the aggregating species, no such absorption should be observable in this region. GFP was purified from the inclusion bodies of bacterial cells grown at 37°C and, as a positive control, from the soluble fraction of cells grown at 25°C. GFP extracts were heat denatured before absorbance measurements. Protein derived from cells grown at 25°C showed a characteristic absorption peak similar to that of acid-denatured GFP [46]. By contrast, protein purified from inclusion bodies of cells grown at 37°C showed no such absorption, indicating that the aggregating species has not formed a chromophore. Taken together, the results presented above indicate that the temperature sensitivity of GFP maturation is due primarily to the failure of the unmodified apoprotein to fold into its catalytically active conformation at higher temperatures. Furthermore, the amino acid substitutions present in GFPA suppress this defect by enhancing proper folding at elevated temperatures.

Modification of the spectral properties of GFP

Fluorescence spectroscopy of purified protein revealed that the fluorescence spectra of GFPA are essentially unchanged from those of GFP except for a small shift in the relative amplitudes of the 400 nm and 475 nm excitation peaks. Although this spectral shift might be advantageous for applications that utilise 400 nm excitation, the reduced relative amplitude of the 475 nm peak is detrimental for those that utilise blue light excitation. For many purposes, the ideal spectral variant would be a protein which could be efficiently excited at either of these wavelengths, a characteristic that would afford great flexibility with regard to the range of applications in which the protein could be used.

Recently, it has been demonstrated that the relative amplitudes of the excitation peaks of GFP can be altered by means of mutagenesis [9, 48-50]. A number of these mutations, like the substitutions present in GFPA, are located in the C-terminal region of the protein. It has been hypothesised that these mutations affect the microenvironment of the chromophore so as to influence the equilibrium between the two spectroscopic states of the chromophore that give rise to the two excitation peaks [9, 49]. One of these mutations, I167T, has been shown to increase the amplitude of the 475 nm excitation peak relative to that of the 400 nm peak [9]. In order to test whether the combination of the I167T substitution with the substitutions present in GFPA might increase the amplitude of the 475 nm peak relative to the 400 nm peak, the sequence of *mgfp4* was modified so as to code for the V163A, S175G and I167T substitutions, together with a number of additional changes to codon usage, to produce GFP5. This variant has two excitation peaks (maxima at 395 nm and 473 nm) of almost exactly equal amplitude and an emission spectrum (λ_{max} = 507 nm) largely unchanged from that of GFP (Fig. 4B). GFP5 retains a thermotolerant phenotype, and bacterial cells grown at 37ºC fluoresce 39-fold and 111-fold more intensely than cells expressing GFP, when excited at 395 nm and 473 nm, respectively. These results demonstrate that it is possible to manipulate the fluorescence spectra of GFPA by introducing additional substitutions into the protein without deleteriously affecting its improved folding characteristics.

Summary of alterations required for efficient *gfp* expression in plants.

The wild-type *gfp* sequence is inactive in *Arabidopsis*. I have described in some detail the various steps that our laboratory has taken in order to correct several defects within the marker gene. Other groups have been struggling with similar problems in other systems and a number of functionally similar *gfp* variants are now available. For

example, *gfp* variants with "humanised" or other optimised codon usage have been produced for better expression. These also show improved levels of expression in plants [21, 31, 32].. While these genes were expected to provide better translation efficiency, it is likely that these alterations also confer some degree of immunity from aberrant RNA processing. A number of workers have obtained GFP mutants that show brighter fluorescence in heterologous cell types, and it is likely that the improved properties result from better folding of the proteins. For example, the V163A mutation has been generated independently by four other groups [48, 51-53], and this residue may play a pivotal role in folding of the protein. In addition, Cormack et al. [54] have introduced large numbers of random amino acid substitutions into the 20 residues flanking the chromophore of GFP. They used fluorescence activated cell sorting to select variants that fluoresced 20- to 35-fold more intensely than wild-type. They also showed that the folding of these mutant proteins was more efficient during expression in bacteria. One of these variants (GFPmut1, [54]) contains two amino acid differences, F64L and S65T, located within the central a-helix of the protein, adjacent to the chromophore. The V163A and S175G mutations that we have isolated are positioned on the outer surface of the protein [55, 56]. and recombination of these two sets of mutations results in an exclusively blue light excited GFP with markedly improved fluorescence properties [57, 58] (J.H., unpublished results). The beneficial effect of both sets of mutations on protein folding, and their apparent additive effect suggests that they may play separate roles in the folding or maturation process.

In our case, a number of improvements have been combined in a single, highly active form of the gene named *mgfp5-ER*. The gene has been optimised for expression in plants, particularly Arabidopsis, and includes:

i) Removal of a cryptic intron. We discovered that the GFP mRNA sequence is efficiently mis-spliced in transgenic *Arabidopsis* plants, resulting in the removal of 84 nucleotides from within the coding sequence, between residues 380 to 463. We removed the cryptic intron by mutagenesis, allowing proper expression.

ii) Subcellular localisation of GFP. We found that GFP accumulates within the nucleoplasm of plant cells, as it does in other organisms, and that it was difficult to regenerate plants from very brighly transformed tissues. This apparent mild toxicity may be due to the generation of fluorescence-related free radicals during growth under light, which might then lead to DNA damage. We have therefore targeted GFP to different subcellular compartments in transgenic plants, and assayed for improved regeneration and fluorescence. One of our constructions, which is excluded fron nuclei and is retained

within the endoplasmic reticulum in *Arabidopsis*, consistently produces bright and healthy transformants.

iii) Isolation of thermotolerant GFP mutants. We have shown that wild-type GFP is thermosensitive, and fluoresces poorly at 37ºC. We have subjected our modified GFP coding sequence to PCR-based mutagenesis, and have isolated a thermotolerant mutant with improved fluorescence. The mutant contains two altered amino acids (V163A, S175G) which greatly improve folding of the apoprotein at elevated temperatures.

iv) Alteration of spectral properties. The fluorescence excitation of wild-type GFP peaks at wavelengths of 400nm and 475nm, with the 400nm peak predominating. We have recombined a published mutant of GFP (I167T, Heim *et al*., PNAS 91:12501-12504, 1994) with our improved mutant (V163A, S175G) and produced a variant which has dual excitation peaks of almost equal amplitude, and which is highly fluorescent *in vivo*. This allows the efficient use of techniques which require either UV or blue light excitation of the protein, for example when screening GFP-expressing plants with a UV lamp, or when using blue laser light excited confocal microscopy, respectively.

This variant has proved to be a bright fluorescent marker, and has found routine use for live monitoring of gene expression and marking cells.

Imaging of cellular dynamics

The expression of GFP within an organism produces an intrinsic fluorescence that colours normal cellular processes, and high resolution optical techniques can be used non-invasively to monitor the dynamic activities of these living cells. Using coverslip-based culture vessels, specialised microscope objectives and the optical sectioning properties of the confocal microscope [18], it is possible to monitor simply and precisely both the three dimensional arrangement of living cells within a meristem, and their behaviour through long time-lapse observations . *In vivo* detection of GFP can also be used as a simple genetic screening procedure for plants growing in normal culture, and we have used the marker to simplify a large enhancer-trap screen.

For time-lapse studies, it is very important that GFP fluorescence be bright, to minimise high levels of illumination that can cause phototoxicity and photobleaching during observation. In this way the efforts required to produce improved forms such as the *mgfp5-ER* gene have been rewarded, and we now routinely use this gene for work in *Arabidopsis*. While we do not expect that misfolding of GFP apoprotein is a substantial

problem at the normal temperatures for *Arabidopsis* growth (around 22ºC), the versatile spectral properties of the *mgfp5-ER* product allow both efficient UV and blue light excitation. This allows simple inspection of transformed material using a hand-held UV lamp, since the illuminating wavelength is poorly detected by the human eye, and an improved signal for microscopy with commonly available filter sets and blue lasers.

During confocal microscopy experiments, we have routinely observed high rates of cytoplasmic streaming within living specimens, and we have used short term time-lapse observations to gain a better understanding of the relative movements of cellular components. *Arabidopsis* seedlings that expressed the *mgfp5-ER* gene at high levels were simply mounted in water for confocal microscopy, which allowed observation for up to two hours. Hypocotyl epidermal cells form ideal specimens for viewing the various components of the cytoplasm. The cells are large, highly vacuolate and surface borne. An extremely thin layer of cytoplasm is squeezed between the wall and the vacuole of these cells. This greatly limits the movement of cytoplasmic components to within a single plane of focus of the microscope, and objects can be rapidly tracked across a portion of the cell without the need for refocusing. A seedling can easily be mounted so that the hypocotyl is pressing closely against the microscope coverslip, and the layer of cytoplasm beneath the outer wall of an epidermal cell will be only a few microns from the surface, allowing maximal optical resolution.

To follow rapid movements in cells it is necessary to use a correspondingly fast sampling rate. We have collected time-lapse confocal images at up to 2 frames per second, which requires almost continual laser scanning. Living specimens have been examined for up to an hour without appreciable phototoxic or bleaching effects, but this is only possible with bright samples which allow attenuation of the exciting laser light. A short segment of a time-lapse experiment is shown in figure 6. A section of hypocotyl epidermis was monitored at a rate of 0.5 frames per second for about 20 minutes and representative confocal images are shown for a 1 minute 10 second period. Cellular components are clearly recognisable in the optical sections, and their identity is indicated in a schematic diagram (Fig. 5A). The cells contain green fluorescent proplastids and highly reticulate endomembranes. The nuclei are outlined due to the peripheral distribution of the ER, and the reticulate surface of a nucleus can be seen in the cell that is central to the field of view. A cross-section of a nucleus can also be seen in the adjacent lower cell. Chloroplasts are red autofluorescent, and characteristically small and spheroid in these hypocotyl epidermal cells.

When a time course of images is played at video rate proplastids and what appears to be vesicular material move vigorously and erratically through the cells. The plastids move with uneven velocities, up to 20 microns per second, along irregular paths that may correspond to underlying cytoskeletal elements such as actin. In contrast, the endoplasmic reticulum, which is presumably associated with cortical microtubules, undergoes relatively slower rearrangement. A relatively stable feature of the ER is indicated with an arrow in figure 5, panels 0:00 to 0:25, while nearby proplastids undergo substantial movement. Chloroplasts and nuclei moved only slowly during the 20 minute time course of the experiment. These cells contain an ER retained form of GFP, and we expect the protein to be cycled in vesicles between the lumen of the ER and the *cis* golgi. A rapid and irregular movement of small vesicle-like particles is seen throughout cells during the time course. Although these small movements are difficult to see in still images, we also see the transient formation of extended filamentous structures, which are comprised of a larger amount of this fluorescent material, and are associated with rapid movement of both vesicular-like material and proplastids. The location of the *cis* golgi in these micrographs is unclear, although small regions of punctate fluorescence can be seen associated with endomembranes. GFP can also be readily targeted to different subcellular compartments or cytoskeletal elements, to selectively highlight different structures within the cell.

The behaviour of GFP-labelled cells within intact plants can be observed using longer time-lapse techniques. For example, *mgfp5-ER* expressing seedlings can be planted in sterile agar media and grown in coverslip-based vessels. The roots of these plants grow down through the media and then along the surface of the coverslip. The roots are then ideally positioned for microscopic imaging through the base of the vessel. A series of images are shown in Fig. 5B that illustrate cell division within an *Arabidopsis* root tip. Confocal optical sections were collected at 2 minute intervals. The localization of GFP to the endoplasmic reticulum, and its consequent perinuclear distribution, ensures that the cell nuclei are clearly evident in these highly cytoplasmic meristem cells. The processes of cell division can be visualised within the living plant. The breakdown of the nuclear membrane, segregation of chromosomes and formation of the daughter nuclei and cell wall plate are reflected in changes of the distribution of the ER-localized GFP. Also, the cell nuclei appear to possess a larger volume prior to cell division, consistent with an extra, newly replicated DNA complement. This may be a useful character for scoring DNA replication within living tissues.

Spectral variants of GFP are now available that allow double-labelling [48, 55] . The precision with which particular cellular structures can be decorated with GFP, and the ease with which subcellular traffic can be monitored indicates that this approach will be very useful for cell biological and physiological observations, particularly for detailed examination of plant mutant phenotypes.

Marking different cell types in Arabidopsis.

In vivo detection of GFP can also be used as a simple genetic screening procedure for plants growing in normal culture, and used the marker to greatly speed and improve a large enhancer-trap screen. We have developed a scheme for targeted gene expression in plants, that is based on a method widely used in *Drosophila* [59]. This requires the use of a heterologous transcription factor, such as the yeast GAL4 protein. We have found it necessary to alter the codon usage of the gene, and to use a derivative, GAL4-VP16, to ensure efficient expression in *Arabidopsis*. Elevated A/U content plays a major role in intron recognition during plant pre-mRNA splicing and we have found that this can poison expression of heterologous A/T rich genes such as GAL4 and GFP. We have randomly inserted the modified GAL4-VP16 gene into the *Arabidopsis* genome, using *Agrobacterium tumefaciens*-mediated transformation. Expression of the GAL4-VP16 gene is dependent upon the presence of adjacent genomic enhancer sequences, and so different patterns of expression are generated. The inserted DNA also contains a GAL4-responsive GFP gene (Fig. 6), and so patterns of GAL4-VP16 gene expression are immediately detectable, with each GAL4-expressing cell marked by green fluorescence.

The *in vivo* detection of GFP allowed us to develop a new and efficient screening procedure. As our particular interest is in the cells of the *Arabidopsis* root tip, we have modified the plant transformation protocol to include an auxin induction of roots from regenerating shootlets. More than 7500 transformants were then generated, planted in grid patterns in sterile culture dishes and directly screened for GAL4 -mediated GFP expression within roots. Several hundred lines with interesting patterns of root expression were choosen, documented, transferred to soil and grown to seed, to both amplify and self-hybridise the lines. As a result, we have a collection of 250 *Arabidopsis* lines with distinct and stable patterns of GAL4-VP16 and GFP expression in the root (Fig 7). These are being made available through the *Arabidopsis* stock centre, and we have created a graphical database of the expression patterns to allow easy computer and web access (http://brindabella.mrc-lmb.cam.ac.uk). These lines provide a valuable set of markers,

where particular cell types are tagged and can be visualised with unprecedented ease and clarity in living plants. Image reconstruction techniques can be used to measure the three dimensional arrangement of cells and common fields of gene expression within the living tissue, as shown for an *Arabidopsis* root tip in Fig. 8. More importantly, GAL4-VP16 expression within these same lines will allow precise targeted gene misexpression.

The major advantage of the GAL4 system is that a chosen target gene can be cloned under the control of GAL4 upstream activation sequences (UAS), transformed, and maintained silently in the absence of GAL4. Genetic crossing of this single line with any of the library of GAL4-containing lines allows specific activation of the target gene in particular tissue and cell types. The phenotypic consequences of mis-expression, including those lethal to the organism can be conveniently studied. We are now proceeding to target the expression of toxic and regulatory proteins to cells of the root meristem.

Conclusions

In order to overcome problems with the expression of GFP in plant cells, and with the safe accumulation and detection of GFP in whole *Arabidopsis* plants, we have engineered many improvements to the *gfp* gene. The modified gene contains (i) altered codon usage to remove a cryptic plant intron, (ii) added peptide sequences to allow targeting of the protein to the lumen of the endoplasmic reticulum, and mutations which (iii) improve folding of the apoprotein during post-translational maturation (V163A, S175G), and (iv) provide equalised UV and blue light excitation (I167T). This highly modified variant (*mgfp5-ER*) is proving useful as a safe and bright marker in transgenic plants.

A major use for green fluorescent protein will be as a replacement for the ß-glucuronidase gene, which is widely used as a reporter for promoter and gene fusions in transformed plants. The GUS gene product can be localised or quantified using histochemical techniques, but these are generally destructive tests [1]. In contrast, GFP can be directly seen in living tissues. For example, high levels of fluorescence intensity are obtained in GFP-transformed bacterial and yeast colonies, allowing simple screening for GFP expression with the use of a hand-held UV lamp. Such an assay for gene expression in living plants will be a very useful tool for plant transformation and breeding experiments. Many transformation techniques give rise to regenerating tissues which are variable or chimeric, and require testing of the progeny of the primary transformants.

Potentially, GFP expressing tissues could be monitored using *in vivo* fluorescence, avoiding any need for destructive testing, and the appropriate transformants could be rescued and directly grown to seed. Similarly, *in vivo* fluorescence will be an easily scored marker for field testing in plant breeding, allowing transgenes linked to the GFP gene to be easily followed.

Unlike enzyme markers, green fluorescent protein can be visualised at high resolution in living cells using confocal microscopy. The images are not prone to fixation or staining artifacts, and can be of exceptional clarity. Moreover, the activities of living cells, such as cytoplasmic streaming, are clearly evident during microscopy. Ordinarily, movement within a sample is a nuisance, placing constraints on the use of sometimes lengthy techniques for noise reduction during confocal microscopy, such as frame averaging. However,we have shown that it is also possible to monitor dynamic events by time-lapse confocal microscopy, and this combination of a vital fluorescent reporter with high resolution optical techniques shows much promise for use in cell biological and physiological experiments.

Genetic systems such as that of *Arabidopsis* provide a large resource of potentially informative mutants, and there has been much recent improvement in techniques for determining the molecular basis of a particular phenotype. The use of fluorescent proteins will provide a further tools for examining the biology of mutant cells. The ability to simply and precisely monitor both particular cells and subcellular structures that have been highlighted with a fluorescent signal will improve both the screening for particular abnormal phenotypes and the characterisation of dynamic processes.

References.

1. Jefferson, R.A., T.A. Kavanagh, and M.W. Bevan, *GUS fusions: ß-glucuronidase as a sensitive and versitile gene fusion marker in higher plants.* The EMBO Journal, 1987. **6**(13): p. 3901-3907.

2. Davenport, D. and J.A.C. Nichol, *Luminescence in Hydromedusae.* Proceedings of the Royal Society, Series B., 1955. **144**: p. 399-411.

3. Morin, J.G. and J.W. Hastings, *Energy transfer in a bioluminescent system.* Journal of Cellular Physiology, 1971. **77**: p. 313-318.

4. Anderson, J.M. and M.J. Cormier, *Lumisomes, the cellular site of bioluminescence in coelenterates.* The Journal of Biological Chemistry., 1973. **248**(8): p. 2937-2943.

5. Morise, H., *et al.*, *Intermolecular energy transfer in the bioluminecent system of Aequorea.* Biochemistry, 1974. **13**(12): p. 2656-2662.

6. Prasher, D.C., *et al.*, *Primary structure of the aequorea-victoria green-fluorescent protein.* Gene, 1992. **111**(2): p. 229-233.

7. Bokman, S.H. and W.W. Ward, *Renaturation of Aequorea green fluorescent protein.* Biochemical and Biophysical Research Communications, 1981. **101**(4): p. 1372-1380.

8. Cody, C.W., *et al.*, *Chemical-structure of the hexapeptide chromophore of the aequorea green-fluorescent protein.* Biochemistry, 1993. **32**(5): p. 1212-1218.

9. Heim, R., D.C. Prasher, and R.Y. Tsien, *Wavelength mutations and posttranslational autoxidation of green fluorescent protein.* Proceedings Of the National Academy Of Sciences Of the United States Of America, 1994. **91**(26): p. 12501-12504.

10. Chalfie, M., *et al.*, *Green fluorescent protein as a marker for gene-expression.* Science, 1994. **263**(5148): p. 802-805.

11. Inouye, S. and F.I. Tsuji, *Aequorea green fluorescent protein - expression of the gene and fluorescence characteristics of the recombinant protein.* Febs Letters, 1994. **341**(2-3): p. 277-280.

12. Wang, S.X. and T. Hazelrigg, *Implications for bcd messenger-rna localization from spatial- distribution of exu protein in drosophila oogenesis.* Nature, 1994. **369**(6479): p. 400-403.

13. Baulcombe, D.C., S. Chapman, and S.S. Cruz, *Jellyfish green fluorescent protein as a reporter for virus- infections.* Plant Journal, 1995. **7**(6): p. 1045-1053.

14. Heinlein, M., *et al.*, *Interaction of tobamovirus movement proteins with the plant cytoskeleton.* Science, 1995. **270**(5244): p. 1983-1985.

15. Niedz, R.P., M.R. Sussman, and J.S. Satterlee, *Green fluorescent protein - an in-vivo reporter of plant gene- expression.* Plant Cell Reports, 1995. **14**(7): p. 403-406.

16. Sheen, J., *et al.*, *Green-fluorescent protein as a new vital marker in plant-cells.* Plant Journal, 1995. **8**(5): p. 777-784.

17. Hu, W. and C.L. Cheng, *Expression of Aequorea green fluorescent protein in plant cells.* FEBS Letters, 1995. **369**(2-3): p. 331-334.

18. Haseloff, J. and B. Amos, *GFP in plants.* Trends in genetics, 1995. **11**: p. 328-329.

19. Haseloff, J., *et al.*, *Removal of a cryptic intron and subcellular localisation of green fluorescent protein are required to mark transgenic Arabidopsis plants brightly.* Proc. Natl. Acad. Sci. U.S.A., 1997.

20. Reichel, C., *et al.*, *Enhanced green fluorescence by the expression of an Aequorea victoria green fluorescent protein mutant in mono- and dicotyledonous plant cells.* Proceedings of the National Academy of Sciences USA, 1996. **93**: p. 5888-5893.

21. Pang, S.Z., *et al.*, *An improved green fluorescent protein gene as a vital marker in plants.* Plant Physiol., 1996. **112**: p. 893-900.

22. Luehrsen, K.R., S. Taha, and V. Walbot, *Nuclear pre-mRNA processing in higher plants.* Progress in Nucleic Acid Researc and Molecular Biology, 1994. **47**: p. 149-193.

23. Simpson, C.G., *et al.*, *Mutation of putative branchpoint consensus sequences in plant introns reduces splicing efficiency.* The Plant Journal, 1996. **9**(3): p. 369-380.

24. Liu, H.X. and W. Filipowicz, *Mapping of branchpoint nucleotides in mutant pre-mRNAs expressed in plant cells.* The Plant Journal, 1996. **9**(3): p. 381-389.

25. Wiebauer, K., J.-J. Herrero, and W. Filipowicz, *Nuclear pre-mRNA processing in plants: distinct modes of 3'-splice-site selection in plants and animals.* Molecular & Cellular Biology, 1988. **8**: p. 2042-2051.

26. Goodall, G.J. and W. Filipowicz, *The AU-rich sequences present in the introns of plant nuclear pre-mRNAs are required for splicing.* Cell, 1989. **58**: p. 473-483.

27. Goodall, G.J. and W. Filipowicz, *Different effects of intron nucleotide composition and secondary structure on pre-mRNA splicing in monocot and dicot plants.* The EMBO Journal, 1991. **10**(9): p. 2635-2644.

28. Hanley, B.A. and M.A. Schuler, *Plant intron sequences: evidence for distinct groups of introns.* Nucleic Acids Research, 1988. **16**: p. 7159-7176.

29. Valvekens, D., M. Van Montagu, and M. Van Lijsebettens, *Agrobacterium tumefaciens-mediated transformation of Arabidopsis thaliana root explants by using kanamycin selection.* Proceedings of the National Academy of Sciences U.S.A., 1988. **85**: p. 5536-5540.

30. Zolotukhin, S., *et al., A "humanized" green fluorescent protein cDNA adapted for high level expression in mammalian cells.* Journal of Virology, 1996. **70**: p. 4646-4654.

31. Haas, J., E.C. Park, and B. Seed, *Codon usage limitation in the expression of hiv-1 envelope glycoprotein.* Current Biology, 1996. **6**(3): p. 315-324.

32. Chiu, W.L., *et al., Engineered GFP as a vital reporter in plants.* Current Biology, 1996. **6**(3): p. 325-330.

33. Green, M.R., *Biochemical mechanisms of constitutive and regulated pre-mRNA splicing.* Annual Reviews of Cell Biology, 1991. **7**: p. 559-599.

34. Langford, C.J., *et al., Point mutations identify the conserved, intron-contained TACTAAC box as an essential splicing signal in yeast.* Cell, 1984. **36**: p. 645-653.

35. Siemering, K.R., *et al., Mutations that supress the thermosensitivity of green fluorescent protein.* Current Biology, 1996.

36. Fischer, G., *Peptidyl-prolyl isomerases and their effectors.* Anrew. Chem. Int. Ed. Engl., 1994. **33**: p. 1415-1436.

37. Rizzuto, R., *et al., Chimeric green fluorescent protein as a tool for visualizing subcellular organelles in living cells.* Current Biology, 1995. **5**: p. 635-642.

38. Kaether, C. and H.H. Gerdes, *Visualization of protein transport along the secretory pathway using green fluorescent protein.* Federation of European Biochemical Societies, 1995. **369**: p. 267-271.

39. Ogawa, H., *et al., Localization, trafficking, and temperature-dependence of the Aequorea green fluorescent protein in cultured vertebrate cells.* Proceeding of the National Academy of Science USA, 1995. **92**: p. 11899-11903.

40. Pines, J., *GFP in mammalian cells.* Trends in Genetics, 1995. **11**: p. 326-327.

41. Webb, C.D., *et al., Use of green fluorescent protein for visualization of cell-specific gene expression and subcellular protein localization during sporulation in Bacillus subtilis.* Journal of Bacteriology, 1995. **177**: p. 5906-5911.

42. Lim, C.R., *et al., Thermosensitivity of green fluorescent protein fluorescence utilized to reveal novel nuclear-like compartments in a mutant nucleoporin NSP1.* Journal of Biochemistry, 1995. **118**: p. 13-17.

43. Ward, W.W., *et al., Spectrophotomeric identity of the energy transfer chromophores in Renilla and Aequorea green fluorescent proteins.* Photochemistry and Photobiology, 1980. **31**: p. 611-615.

44. Cubitt, A.B., *et al., Understanding, improving and using green fluorescent proteins.* Trends Biochem. Sci., 1995. **20**: p. 229-233.

45. Kane, J.F. and D.L. Hartley, *Formation of recombinant protein inclusion bodies in Escherichia coli.* Trends in Biotechnology, 1988. **6**: p. 95-101.

46. Ward, W.W. and S.H. Bokman, *Reversible denaturation of Aequorea green-fluorescent protein: physical separation and characterization of the renatured protein.* Biochemistry, 1982. **21**: p. 4535-4540.

47. Inouye, S. and F.I. Tsuji, *Evidence for redox forms of the aequorea green fluorescent protein.* Febs Letters, 1994. **351**(2): p. 211-214.

48. Heim, R. and R.Y. Tsien, *Engineering green fluorescent protein for improved brightness, longer wavelengths and fluorescence resonance energy-transfer.* Current Biology, 1996. **6**(2): p. 178-182.

49. Ehrig, T., D.J. Okane, and F.G. Prendergast, *Green-fluorescent protein mutants with altered fluorescence excitation-spectra.* Febs Letters, 1995. **367**(2): p. 163-166.

50. Delagrave, S., *et al., Red-shifted excitation mutants of the green fluorescent protein.* Bio-Technology, 1995. **13**(2): p. 151-154.

51. Crameri, A., *et al., Improved green fluorescent protein by molecular evolution using dna shuffling.* Nature Biotechnology, 1996. **14**(3): p. 315-319.

52. Davis, S. and R. Viestra, Weeds World, 1996.

53. Kohler, R.H., *et al., The green fluorescent protein as a marker to visualize plant mitochondria in vivo.* Plant J., 1997. **11**: p. 613-621.

54. Cormack, B.P., R.H. Valdivia, and S. Falkow, *FACS-optimized mutants of the green fluorescent protein (GFP).* Gene, 1996. **173**: p. 33-38.

55. Yang, F., L.G. Moss, and G.N.J. Phillips, *The molecular structure of green fluorescent protein.* Nature Biotech., 1996. **14**: p. 1246-1251.

56. Ormo, M., *et al., Crystal structure of the Aequoria victoria green fluorescent protein.* Science, 1996. **273**: p. 1392-1395.

57. Zoenicka-Goetz, M., *et al.*, *An indelible lineage marker for Xenopus using a mutated green fluorescent protein.* Development, 1996. **122**: p. 3719-3724.

58. Zoenicka-Goetz, M., *et al.*, *Following cell fate in the living mouse embryo.* Development, 1997. **124**: p. 1133-1137.

59. Brand, A.H. and N. Perrimon, Development, 1993. **118**: p. 401-415.

Figure Legends

Fig. 1 Cryptic splicing of *gfp* transcripts in transgenic *Arabidopsis thaliana*.

(*A*) Schematic diagram of the *gfp* gene sequence which shows the positions of restriction endonuclease sites used for gene analysis, and the location of the cryptic intron, shown with dark shading. Sequences that are similar to those normally found at splice sites and branchpoints of plant introns are shown below. Splice sites are arrowed, and the putative lariat branchpoint is shown in reverse type [19].

(B)Restriction endonuclease digestion of PCR fragments derived from *gfp* DNA and mRNA sequences. Sequences corresponding to the integrated *gfp* gene and to mRNA transcripts were isolated from transgenic plantlets, separately amplified using PCR techniques, and incubated with various restriction endonucleases. The radiolabelled fragments were fractionated by electrophoresis in a 5% polyacrylamide gel, and are shown labelled with the source of the amplified sequences (DNA or mRNA) and the name of the restriction endonuclease used for digestion, or not (uncut). The mRNA derived sequences appeared to lack sites for *Dra* I and *Acc* I, and to contain a corresponding deleted region of 80-90 nucleotides. Restriction endonuclease fragments that are smaller than those expected of the gene sequence have been indicated with a white asterisk.

(C) Sequence analysis of cloned *gfp* mRNAs. Autoradiograph and sequence of the amplified *gfp* mRNA sequence. Nucleotides 380-463 are absent from the transcribed sequence, and the site of this 84 nucleotide deletion is arrowed.

Fig. 2 Confocal optical sectioning of intact plants.

Transgenic Arabidopsis seedlings that expressed a 35S promoter-driven *mgfp4* gene were subjected to microscopic examination. A series of images is shown to exemplify the benefits of laser scanning confocal microscopy for observation of GFP expressing whole mount samples.

(A) The shoot apex as imaged using conventional epifluourescence optics. The depth of the sample, combined with a large fluorescence signal from outside the plane of focus, result in a poorly resolved image. The outlines of cells and punctate nuclear accumulation of the mgfp4 gene product are obscured by out-of-focus blur.

The remaining images were collected using a Biorad MRC-600 instrument equipped with Nikon Optiphot microscope and Nikon 10x and planapo 60x water immersion lens. GFP and chlorophyll were excited using the 488nm and 568nm lines, respectively, of a 25nm krypton-argon ion laser. The green and red emissions were collected in separate channels and combined using Adobe Photoshop.

(B) The shoot apex can be more clearly resolved, even using a low magnification (10x) objective .

(C) When higher power objective (60x) with larger numerical aperture (1.2) is used, the optical sectioning properties of the confocal microscope are clearly evident. An optical section through the shoot apex of a transgenic seedling is shown. GFP (green channel) is found distributed throughout the cytoplasm. The red autofluorescent chloroplasts (red channel) are clearly resolved from the GFP fluorescence.

(D) Individual cells can be clearly imaged. Nuclear accumulation of GFP is seen in these cells, with exclusion from the nucleolus. The cytoplasm is pressed to the cell wall in these highly vacuolate cells, and sandwich non-fluorescent cell walls between adjacent cells. Chloroplasts are shown in red.

(E) The cytosol forms a thin layer at the periphery of highly vacuolate cells. A highly magnified optical section through an outer layer of cytoplasm in a hypocotyl epidermal cell is shown. Unmodified GFP is excluded from endomembrane components and plastids within the cytosol, and forms a negative stain for these components. A single chloroplast, with its red fluorescent chlorophyll contents, can be seen in this image. Several non-fluorescent cigar-shaped bodies that appear to be some kind of plastids are also evident.

Fig. 3 Localisation of GFP to the endoplasmic reticulum

35S promoter-driven GFP transformed seedlings were grown in sterile agar culture and mounted intact in water for confocal microscopy. Panels A and C show expression of 35S-mgfp4, and panels B and D show expression of mgfp4-ER, which is localised to the endoplasmic reticulum. Images were collected using a Biorad MRC-600 microscope, as described in Fig. 2.

(A) An optical section of a single hypocotyl cortex cell, showing mature chloroplasts, and distribution of GFP throughout the cytoplasm.

(B) A view of epidermal cells within the hypocotyl, showing the reticulate distribution of GFP-ER within the endomembrane system and the appearance of green fluorescence within maturing plastids. Mature chloroplasts are brightly red autofluorescent in these cells.

(C) Cells from within the root meristem. GFP accumulates within nuclei, but is excluded from nucleoli, and is found throughout the cytoplasm where various endomembrane compartments are shown in negative relief.

(D) Cells within the root meristem clearly display the characteristic perinuclear distribution expected for the ER-localised GFP. This is also seen in the shoot (panel B).

Fig. 4 Improved mutants of GFP.

(A). Improved thermotolerance of GFP. Bacterial cells expressing GFP and GFPA (V163A, S175G) were grown at different temperatures. GFP fluorescence values were measured and normalized with respect to the amount of intracellular recombinant protein for cultures grown at 25°C, 30°C, 37°C and 42°C. [35]

(B). Excitation and emission spectra of GFP, GFPA (V163A, S175G) and GFP5 (V163A, I167T, S175G). Protein concentrations were 23.5 μg/ml in PBS (pH 7.4). All spectra have been normalised to a maximum value of 1.0. [35]

Fig. 5 Time-lapse confocal microscopy of subcellular processes.

(A) A 5-day old transgenic *Arabidopsis* seedling expressing the *mgfp5-ER* gene was mounted in water and a small segment of the hypocotyl epidermis was observed using a Biorad MRC-600 laser scanning confocal microscope. The laser light was attenuated by 99% using a neutral density filter, and the confocal aperture was stopped down. Two channel, single scan images were collected at the rate of 1 per 2 seconds for 20 minutes, and transferred to an Apple Macintosh computer. The large data file was then converted to full-colour numbered PICT files using the program PicMerge, and finally converted to a Quicktime movie for analysis and video rate playback. A section corresponding to 4.5 minutes of the original observation was chosen and representative frames are presented here. Each frame is marked with the time (minutes:seconds) that had elapsed from the first chosen frame. The images 0:00 to 0:25 each contain an arrow which indicates a ring-like feature within the ER that provides a morphological landmark.

(B) Seeds of the *Arabidopsis* enhancer trap line J0571 (J.H. & Sarah Hodge, unpublished results) were germinated and grown in agar medium on a coverglass. This line expressed mgfp5-ER within the cortex and endodermal cell layers of the root. After 10 days of growth, an emerging lateral root was visualized by confocal time-lapse microscopy. The root tip was imaged through the coverglass of the tissue culture vessel. A median longitudinal optical section was collected every 2 minutes over a 6 hour period. Frames from a 160 minute period are shown, labelled with the time of collection (min: sec). Cell division of one cell in the endodermis layers is shown. The ER-localization of the GFP marker allows clear visualization of nuclear division and phragmoplast formation in these cells.

Fig. 6 Scheme for GAL4-VP16 enhancer trapping

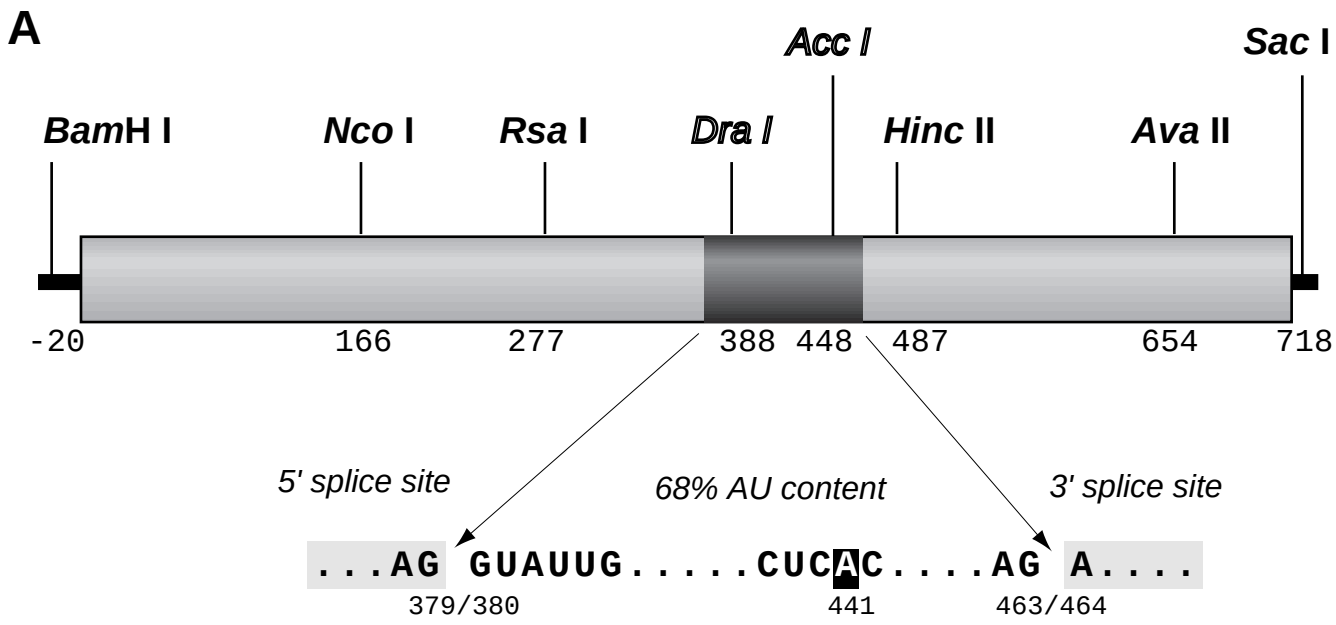
An enhancer trap vector was randomly introduced into the Arabidopsis genome by Agrobacterium-mediated transformation. The transformation vector contains a GAL4-VP16 gene with modified codon usage adjacent to a naive promoter at a border of the

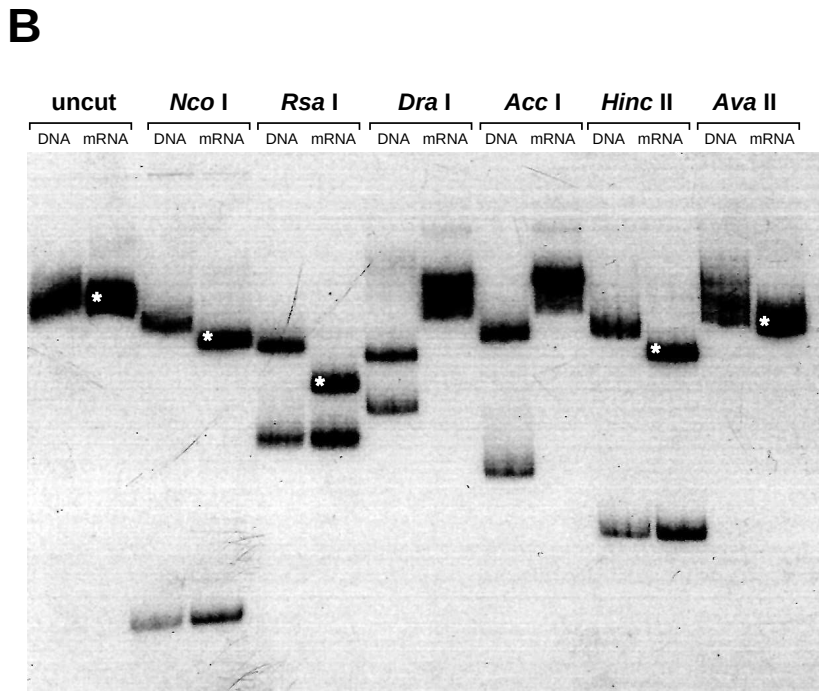
transferred DNA, a kanamycin resistance selection marker and a GAL4-responsive mGFP5-ER gene. The modified GFP has improved fluorescent properties and is targeted to the endoplasmic reticulum. Cell-specific activation of the GAL4-VP16 gene by a cellular enhancer results in the expression of the GFP marker gene, allowing the simple characterisation of expression patterns. Targeted expression of another gene (X) can be induced by a genetic cross with a GAL4-VP16 line.

Fig. 7 *Arabidopsis* GAL4-VP16 enhancer trap lines.
Confocal micrographs of selected enhancer trap lines showing root tip specific expression. The roots were counterstained with propidium iodide, which outlines all live cells. The GAL4-VP16 driven GFP signal is shown superimposed. A schematic diagram of the different types of cell in the root tip is shown (upper left).

Fig. 8 Three dimensional arrangement of cells inside living tissues.
(A) and (B) Transverse longitudinal sections of root tips from a GAL4-VP16 expressing Arabidopsis line are shown. GAL4-driven GFP expression is seen in the root cap and endodermis. These panels show roots from different 5 day old seedlings that have been stained with propidium iodide (red channel) to reveal the outliness of cells in the tip, wih the GFP signal shown in the green channel. The root tip labelled B possesses a dead cell within the meristem that has filled with red fluorescent propidium iodide. The dead cell is positioned as an initial within the vascular bundle and is in direct contact with cells of the quiescent centre.
(C) Image reconstruction techniques were used to build a three dimensional model of the cells within the root tip from a series of optical sections. The model was then rotated to view the relative three dimensional arrangement of the dead and the GFP expressing cells. Four frames from this rotation series are shown. GFP expression within the endodermal cell adjacent to the dead cell has been extinguished. Death of the vascular initial appears to have triggered cell fate changes in its immediate neigbourhood.

Figure 1


A
Acc I
Sac I
BamH I
Nco I
Rsa I
Dra I
Hinc II
Ava II
-20
166
277
388
448
487
654
718
5' splice site
68% AU content
3' splice site
...AG GUAUUG.....CUCAC....AG A....
379/380
441
463/464


B
uncut
Nco I
Rsa I
Dra I
Acc I
Hinc II
Ava II
DNA mRNA

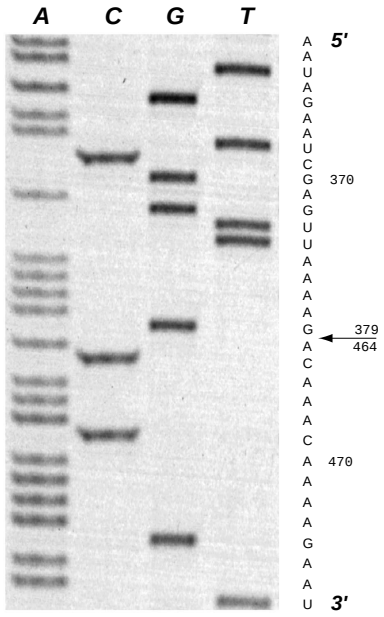

C
A
C
G
T
5'
370
379
464
470
3'

Figure 2

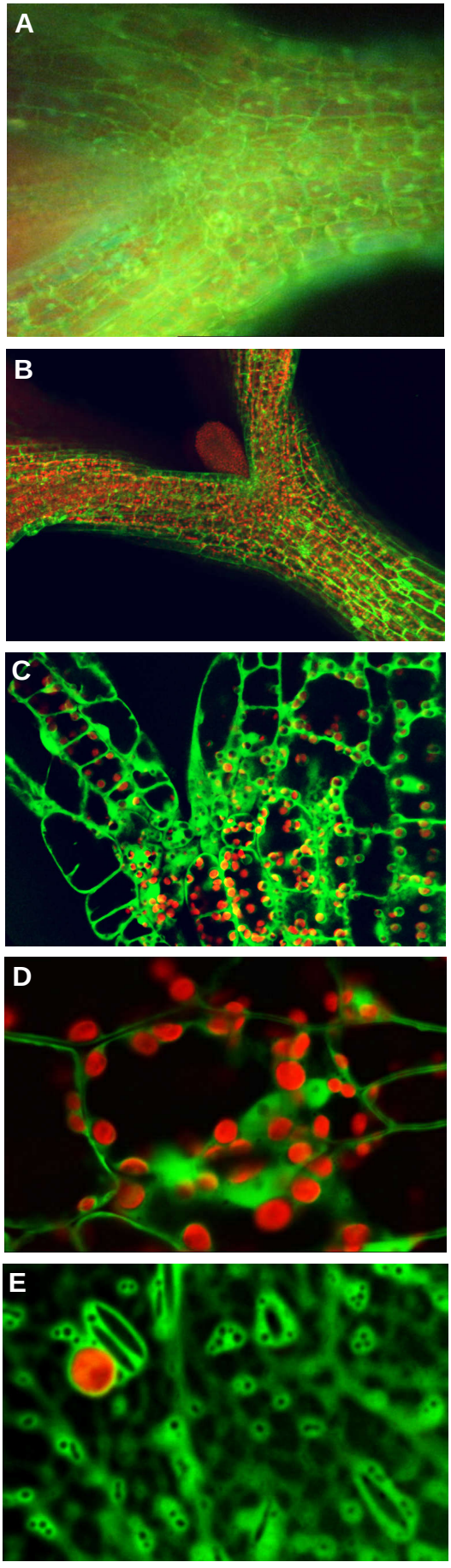
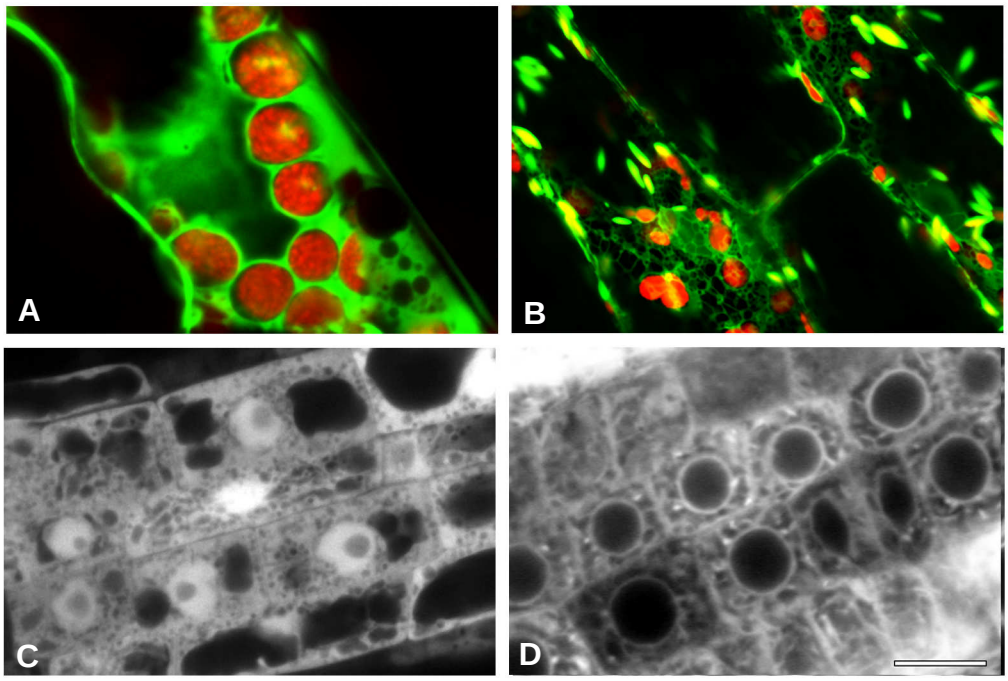
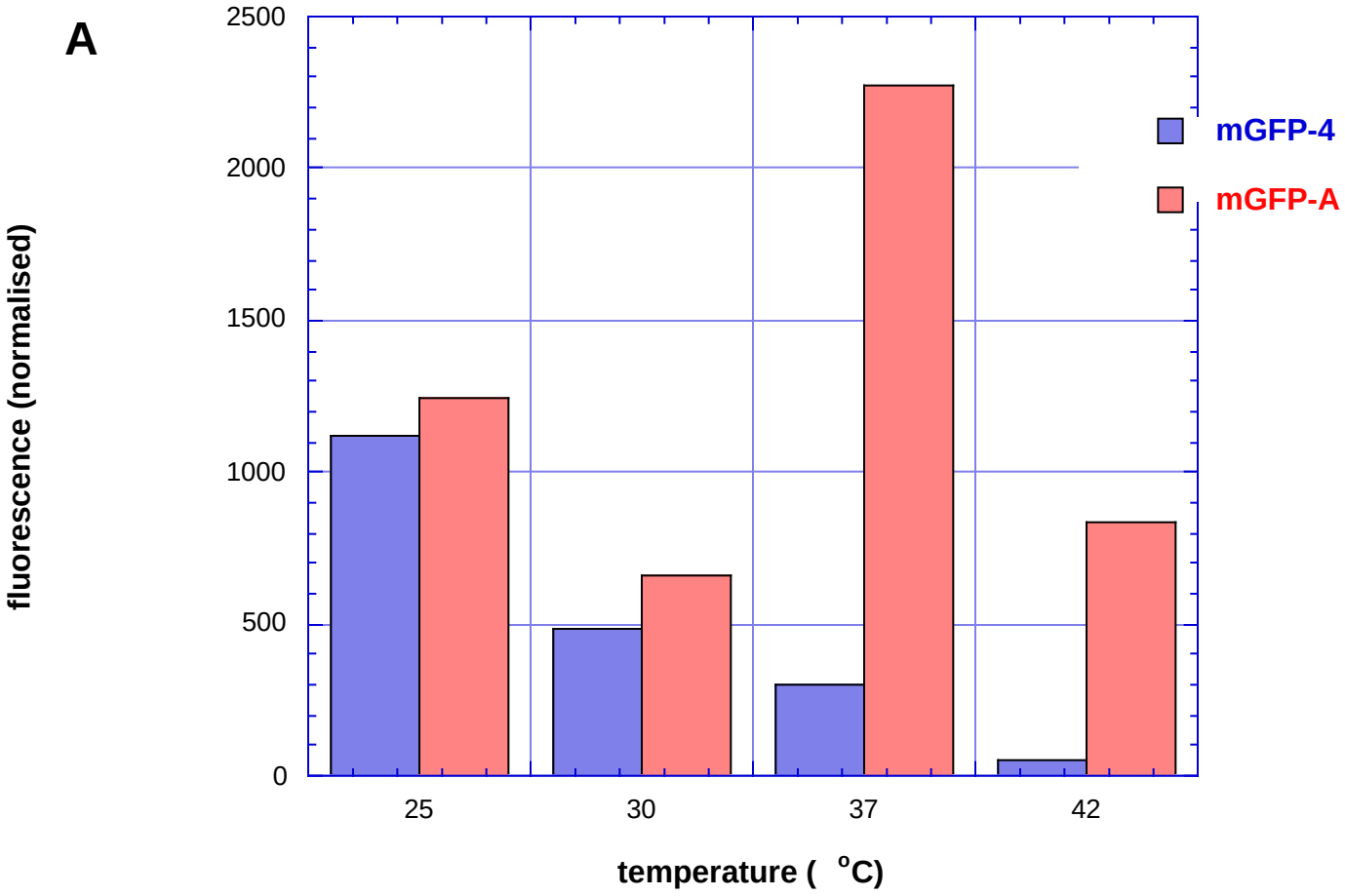


Figure 3




A
2500
2000
1500
1000
500
0
fluorescence (normalised)
25
30
37
42
temperature (°C)
mGFP-4
mGFP-A

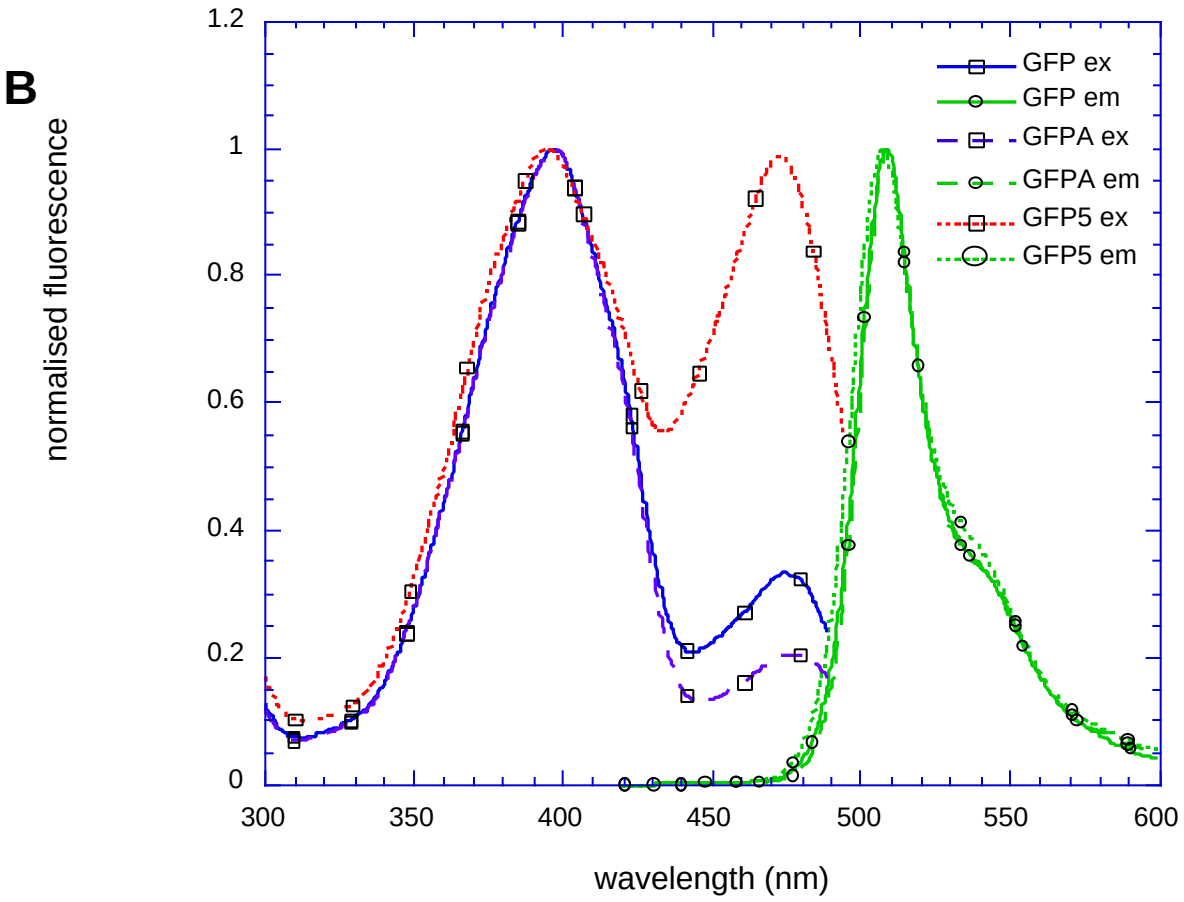

B
1.2
1
0.8
0.6
0.4
0.2
0
normalised fluorescence
300
350
400
450
500
550
600
wavelength (nm)
GFP ex
GFP em
GFPA ex
GFPA em
GFP5 ex
GFP5 em

Figure 5

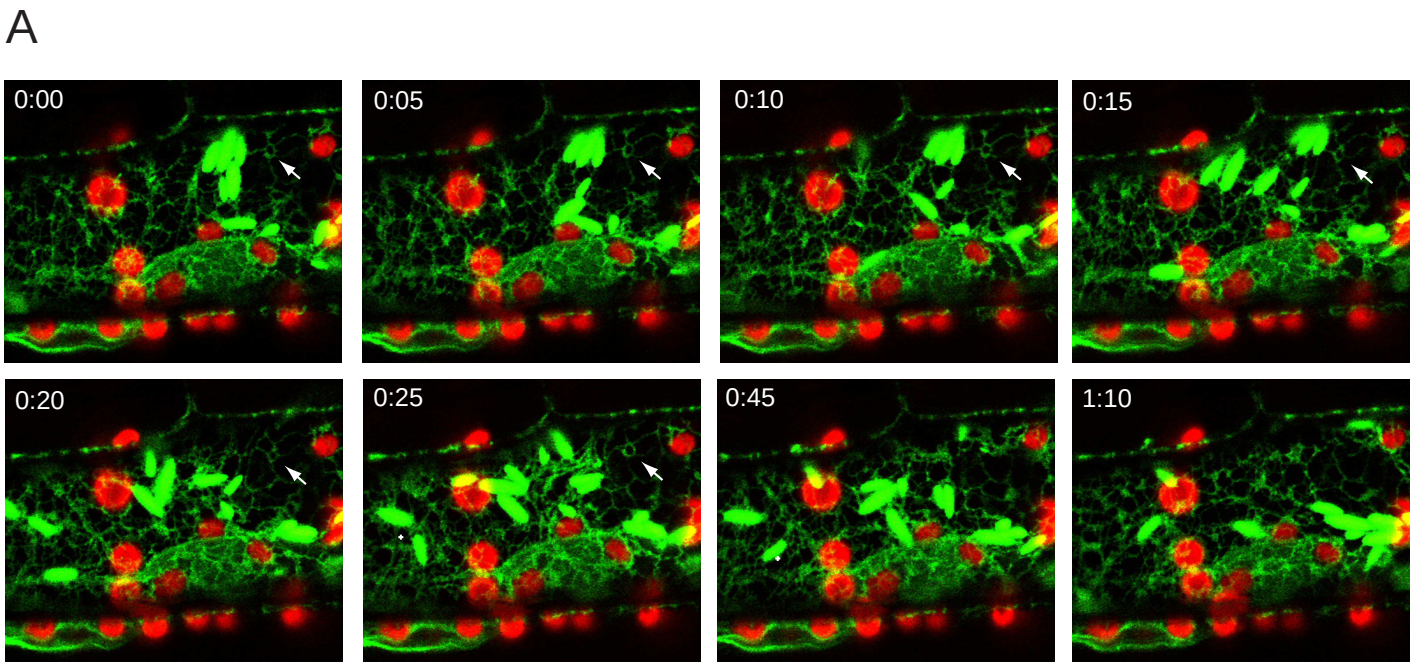

A
0:00
0:05
0:10
0:15
0:20
0:25
0:45
1:10
B
0:00
4:00
8:00
12:00
16:00
20:00
28:00
160:00

Figure 6

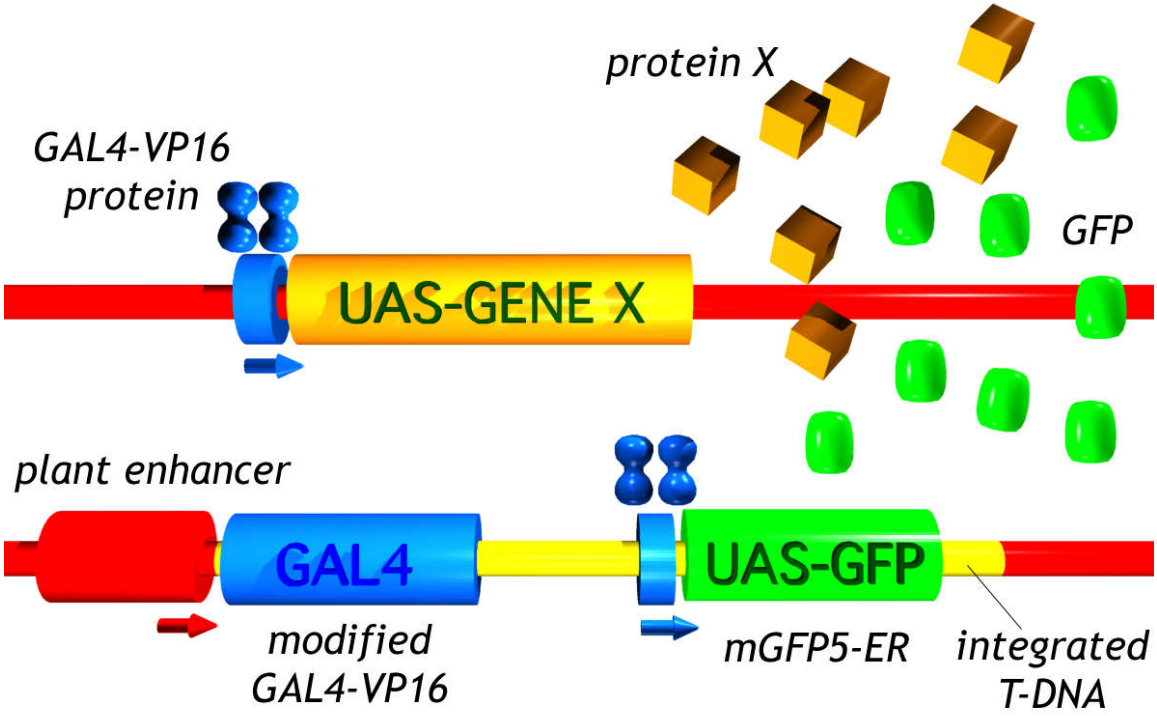

protein X
GAL4-VP16
protein
GFP
UAS-GENE X
plant enhancer
GAL4
UAS-GFP
modified
GAL4-VP16
mGFP5-ER
integrated
T-DNA

Figure 7

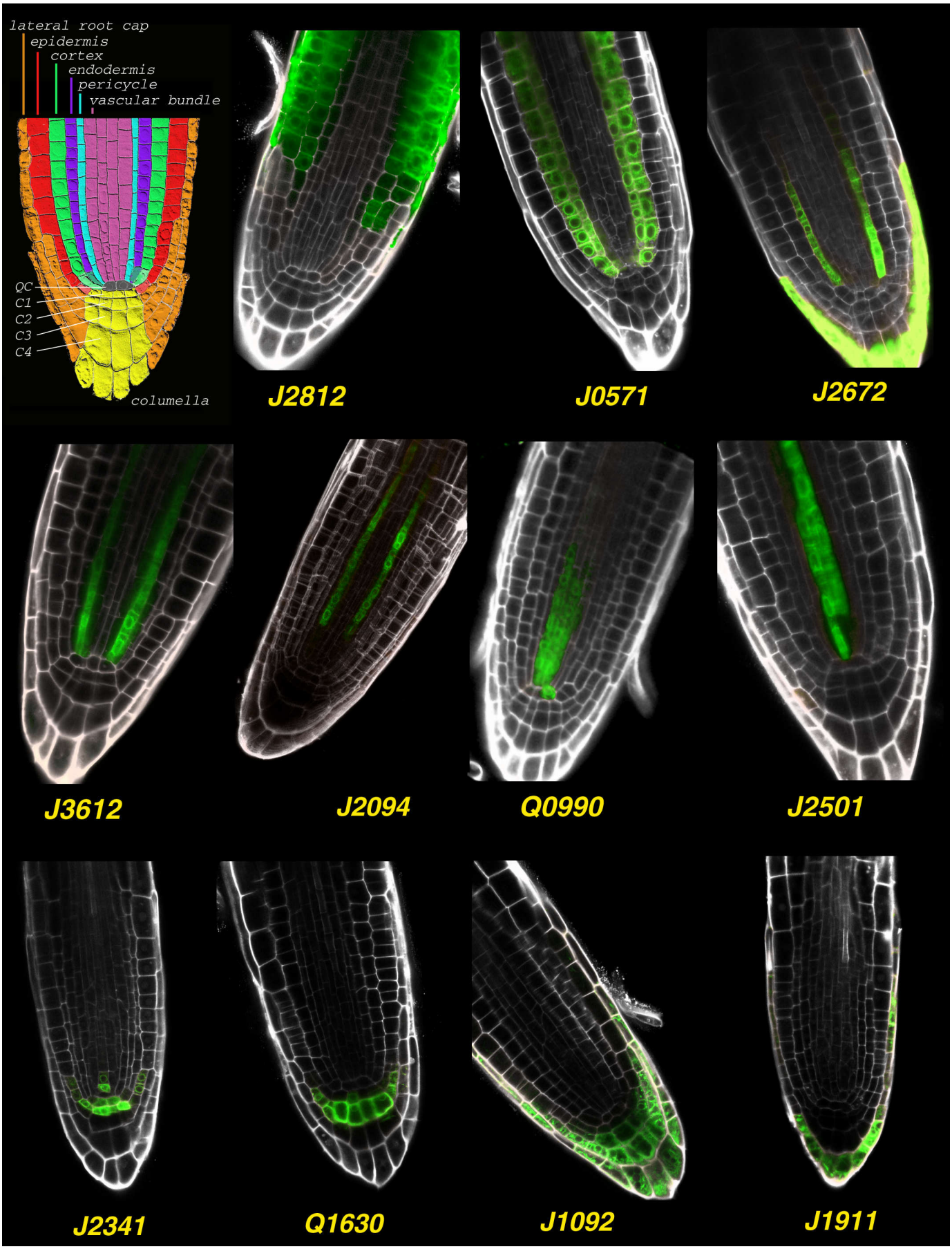

lateral root cap
epidermis
cortex
endodermis
pericycle
vascular bundle
QC
C1
C2
C3
C4
columella
J2812
J0571
J2672
J3612
J2094
Q0990
J2501
J2341
Q1630
J1092
J1911

Figure 8