Stable Genetic Transformation and Heterologous Expression in the Nitrogen-fixing Plant Endosymbiont *Frankia alni* ACN14a

Isaac Gifford[a]#, Summer Vance[a], Giang Nguyen[a], and Alison M Berry[a]

Author affiliations: [a]Department of Plant Sciences, University of California, Davis, Davis, CA, United States

#Address correspondence to Isaac Gifford, isgifford@ucdavis.edu

Running Title: Genetic Transformation of *Frankia alni*

Word Counts:

Abstract: 221 words

Importance: 123 words

Text Body: 5315 words

Abstract

Genus *Frankia* is comprised primarily of nitrogen-fixing actinobacteria that form root nodule symbioses with a group of hosts known as the actinorhizal plants. These plants are evolutionarily closely related to the legumes, which are nodulated by the rhizobia. Both host groups utilize homologs of nodulation genes for root-nodule symbiosis, derived from common plant ancestors. However the corresponding endosymbionts, *Frankia* and the rhizobia, are distantly related groups of bacteria, leading to questions of their symbiotic mechanisms and evolutionary history. To date, a stable system of genetic transformation has been lacking in *Frankia*. Here, we report the

21 successful electrotransformation of *Frankia alni* ACN14a, by means of replicating
22 plasmids expressing chloramphenicol-resistance for selection, and the use of GFP as a
23 marker of gene expression. We have identified type IV methyl-directed restriction
24 systems, highly-expressed in a range of actinobacteria, as a likely barrier to *Frankia*
25 transformation and circumvented this barrier by using unmethylated plasmids, which
26 allowed the transformation of *F. alni* as well as the maintenance of the plasmid. During
27 nitrogen limitation, *Frankia* differentiates into two cell types: the vegetative hyphae and
28 nitrogen-fixing vesicles. When the plasmid transformation system was used with
29 expression of *egfp* under the control of the *nif* gene cluster promoter, it was possible to
30 demonstrate by fluorescence imaging the expression of nitrogen fixation in vesicles but
31 not hyphae in nitrogen-limited culture.

32 **Importance**

33 To date, the study of *Frankia*-actinorhizal symbioses has been complicated by the
34 lack of genetic tools for manipulation of *Frankia*, especially stable genetic transformation.
35 The transformation system reported here, particularly coupled with marker genes, can be
36 used to differentiate patterns of gene expression between *Frankia* hyphae and vesicles in
37 symbiosis or in free-living conditions. This will enable deeper comparisons between
38 *Frankia*-actinorhizal symbioses and rhizobia-legume symbioses in terms of molecular
39 signaling and metabolic exchange that will broaden understanding of the evolution of
40 these symbioses and potentially make possible their application in agriculture. The
41 development of transformation methods will allow further down-stream applications
42 including gene knock-outs and complementation that will, in turn, open up a much
43 broader range of experiments into *Frankia* and its symbioses.

Introduction

Bacteria in the genus *Frankia* form nitrogen-fixing root nodule symbioses with a group of host plants, the actinorhizal plants. The actinorhizal plants are evolutionarily closely related to the legumes within the Nitrogen-fixing Clade (NFC; Soltis *et al*., 1995). The bacterial symbionts are only distantly related however: *Frankia* belongs to the phylum Actinobacteria, comprised of high-GC-content gram-positive bacteria, whereas the rhizobia, symbionts of the legumes, are gram-negative proteobacteria. Additionally, *Frankia* is a multicellular, hyphal genus whereas the rhizobia are single-celled. In most of the actinorhizal symbioses, and in nitrogen-limiting conditions *in vitro*, *Frankia* differentiates into two cell types: vegetative hyphae and nitrogen-fixing vesicles. Vesicles are surrounded by a lamellar hopanoid lipid envelope (Berry *et al.,* 1993) that increases in number of layers in response to oxygen tension thereby likely reducing the flow of oxygen into the vesicle interior to protect the nitrogenase complex from deactivation (Benson and Silvester, 1993). Thus, unlike the rhizobia which require their hosts to regulate the flow of oxygen for nitrogen fixation in symbiosis (Appleby, 1984), *Frankia* vesicles are capable of fixing nitrogen in atmospheric oxygen conditions (Benson and Silvester, 1993).

Biological nitrogen fixation has potential applications in the development of minimal input strategies for more sustainable agriculture globally and in nutrient-deficient soils through genetic improvement of crops (Gutierrez, 2012). In efforts to understand the mechanisms of root nodule symbioses, phylogenetic studies of the Nitrogen-Fixing Clade have shown a shared predisposition underlying the evolution root nodule symbioses between the actinorhizal plants and legumes (Werner *et al*., 2014;

Battenberg *et al*., 2018). The actinorhizal and legume hosts share the Common Symbiotic Pathway, a signaling pathway that leads to the development of the nodule in response to the symbiont (reviewed in Oldroyd 2013) as well as several key gene orthologs in the process of nodule development (Battenberg *et al.,* 2018; Griesmann *et al*., 2018). Clear differences also exist in the genetics of nodulation in the two bacterial groups, leaving major questions concerning the evolution and development of root-nodule symbioses in the bacterial partners. The rhizobial signaling molecules that trigger nodulation in many of the legume hosts, called Nod factors, are synthesized by *nod* genes and have been extensively studied (Oldroyd, 2013). The majority of *Frankia* genomes do not contain clear sets of *nod* gene homologs with the exception of several of the cluster II *Frankia*: genomes of these *Frankia* contain homologs of the rhizobial *nodABC* genes that are expressed in symbiosis (Persson *et al*., 2015). Other mechanisms have been found to trigger legume nodulation by rhizobia, including an effector protein injected into the host by a type III secretion system (Okazaki *et al*., 2013); however genes responsible for these pathways lack homologs in *Frankia* genomes as well. Signals from *Frankia* in clusters I and III that elicit host root responses have been detected experimentally, but the chemical composition has not been defined (Ceremonie *et al*., 1999; Cissoko *et al*., 2018).

The major advances in knowledge of rhizobial symbioses have been made possible by the development of genetic tools for dissecting the metabolic and regulatory pathways in the microsymbiont-host interactions. Long *et al*. (1982) originally demonstrated the role of the *nod* genes in symbiosis by complementing nodulation-deficient mutants of *Sinorhizobium meliloti* with *nod* genes encoded on a replicating, broad host-range, plasmid. More recently, expression of marker genes in rhizobia has

enabled a wider-range of experiments including the tracking of microsymbionts during the host infection process (Gage, 2002) and the identification of regulatory networks involved in symbiotic interactions (Spaepen *et al*., 2009). However, to date, there has been no stable genetic transformation system for *Frankia* (Kucho *et al*., 2009). As reported here, a transformation system in *Frankia* will thus enable functional inquiries into the diverse mechanisms involved in establishing root-nodule symbiosis and maintaining biological nitrogen fixation in actinorhizal symbioses. These inquiries will, in turn, contribute to a wider understanding of the origins of root nodule symbioses and mechanisms of *Frankia* and rhizobia interaction with hosts in the NFC.

Frankia has several barriers to transformation. Actinobacteria in general have low rates of homologous recombination due to competition between their homologous recombination pathway and a Non-Homologous End-Joining pathway (Zhang *et al*., 2012). Additionally, natural vectors for *Frankia* transformation are limited as no *Frankia* phages have been discovered (Simonet *et al*., 1990). Finally, its multi-cellular hyphae require that all cells be transformed in order for experiments to be viable. Despite these barriers, it has been demonstrated that DNA can be electroporated into *Frankia* cells (Myers and Tisa, 2004), providing a potential avenue for genetic transformation. Furthermore, Kucho *et al.* (2009) reported initial success in integrating a non-replicating plasmid into the chromosome of *Frankia* sp. CcI3. However the recombined plasmid was lost in the following generation, limiting its use in experiments.

Restriction enzymes pose a barrier to successful transformation in bacteria due to their role in defense by digesting foreign DNA. In some actinobacteria, the use of unmethylated plasmids has increased transformation efficiency (Ankri *et al*., 1996; Molle

113 *et al*., 1999), potentially due to the presence of Type IV methyl-directed restriction
114 enzymes. In the majority of bacterial taxa, DNA is methylated during replication by the
115 methyltransferase Dam to mark parent strands for DNA repair and excision of
116 misincorporated bases from the daughter strand (Sanchez-Romero *et al*., 2015). The
117 majority of actinobacteria, however, lack *dam* homologs (Sanchez-Romero *et al*., 2015)
118 and an investigation of genomes of *Streptomyces*, *Rhodococcus*, and *Micromonospora*
119 spp. found that these genomes were not methylated in the canonical Dam pattern
120 (Novella *et al*., 1996). While actinobacteria have recently been shown to have a unique
121 mismatch repair pathway (Castaneda-Garcia *et al*., 2017) it is not yet known how this
122 pathway utilizes methylation, if at all.

123 In this study we report successful genetic transformation of *Frankia alni* ACN14a
124 by using an unmethylated plasmid to circumvent methylation-targeting restriction
125 enzymes encoded in the ACN14a genome. This permitted the maintenance of plasmids
126 derived from one with a very broad host-range origin of replication (Kurenbach *et al*.,
127 2003) in *F. alni* cultures. Additionally, we show that *gfp* expressed from a replicating
128 plasmid can be used to label gene expression differences between *Frankia* cell types
129 during nitrogen fixation, opening up molecular genetics-based experiments on *Frankia*
130 symbioses in the future.

131 **Results**

132 *Identification of Restriction Enzymes in* Frankia *and Transcriptome Analysis*

133 Type I, II and IV restriction enzymes were identified in several published *Frankia*
134 genomes (Figure 1). Examining the transcriptome of *Frankia alni* ACN14a in (+)N

135 culture (Alloisio *et al*., 2010), we found that three restriction enzyme genes were highly
136 expressed: one type I enzyme (FRAAL4992, 92nd percentile), one type II enzyme
137 (FRAAL0249, 91st percentile) and one type IV enzyme (FRAAL3325, 85th percentile)
138 (Figure 1). The type IV enzyme was annotated in REBASE as a "Type IV Methyl-
139 directed restriction enzyme" of the Mrr methyladenine-targeting family. Other *Frankia*
140 genomes contained likely homologs of this type IV enzyme as well. The genome of
141 *Frankia casuarinae* CcI3 contained five type IV enzymes, the most of any *Frankia*
142 genome examined (Figure 1). In the *Frankia* sp. CcI3 transcriptome, type IV Mrr
143 enzymes were very highly expressed, up to the 96th percentile. One of the five enzymes in
144 CcI3 (Francci3_2839) was annotated as a Mcr type IV enzyme, which targets 5-
145 methylcytosine instead (REBASE) and its expression was very low, in only the 18th
146 percentile of genes in the transcriptome. Of the complete genomes examined, only
147 *Candidatus* Frankia datiscae Dg1 did not contain any putative type IV restriction genes.

148 A symbiotic transcriptome of *F. alni* ACN14a from root nodules is available in
149 addition to transcriptomes of free-living culture (Alloisio *et al*., 2010). In symbiosis with
150 *Alnus glutinosa* the *mrr* homolog was down-regulated relative to the rest of the
151 transcriptome, from the 85th percentile in culture to an expression level at maximum no
152 higher than 12 percent of transcriptome genes, while the other restriction enzymes
153 identified in the *F. alni* genome were not down-regulated in symbiosis (Figure 1). In the
154 other *Frankia sp*. with a symbiotic transcriptome available, the uncultured cluster II
155 species *Frankia* sp. Dg1 (Persson *et al*., 2015), expression of its two restriction enzymes,
156 neither of which was a Type IV enzyme, occurred at very low levels (Figure 1).

Of the actinobacteria outside of genus *Frankia* with published transcriptomes in pure culture, *Mycobacterium smegmatis*, *Streptomyces avermitilis*, and *Rhodococcus jostii* also highly expressed type IV restriction enzyme genes, in the upper-70th percentiles of their respective transcriptomes (Figure 2). In comparison, the transcriptomes of proteobacteria and firmicutes examined expressed their corresponding *mrr* genes around the 50th or 60th percentiles. One notable exception was the actinobacterium *Mycobacterium tuberculosis*, which expressed its single gene for a methyladenine-directed restriction enzyme (Mrr) at extremely low levels, around the 6th percentile in culture.

Genetic Transformation of Frankia alni *ACN14a with an Unmethylated Replicating Plasmid*

After electroporation *F. alni* cells formed visible hyphae in culture after approximately ten days (data not shown). When sub-cultured into chloramphenicol-selective media *F. alni* cultures transformed with unmethylated plasmid pSA3 were able to grow, whereas those transformed with methylated plasmid were not. When imaged under 488nm wavelength of excitation in the confocal microscope, green fluorescence typical of GFP was observed in hyphae as shown in Figure 3. Wild-type *F. alni* hyphae displayed autofluorescence around 575nm (data not shown), as observed by Hahn *et al.* (1993), but no fluorescence was observed in the 500-550nm range (Figure 3).

Differential Regulation of egfp *Under the Control of the* Frankia alni *ACN14a* nif *Cluster Promoter*

When grown in (–)N culture, transformants carrying the pIGSAFnif plasmid showed significant up-regulation of the *egfp* gene conjugated to the *nif* cluster promoter, approximately 100-fold relative to (+)N culture (Table 1). The *nifH* gene, used as a positive control for nitrogen fixation, was also significantly up-regulated approximately 9-fold in (–)N media compared with (+)N cultures. Thus, the up-regulation of the egfp gene is consistent with the previous report that plasmid pSA3 is maintained at approximately 10 copies per cell (Dao and Ferretti, 1985). Expression of the *rpoD* housekeeping gene, used as a negative control, was not significantly different between (+)N and (–)N cultures.

F. alni containing pIGSAFnif grown in (–)N media fluoresced predominantly in the vesicles (Figure 4). Little to no fluorescence was observed in hyphae. Fluorescence in the vesicles was present both in the spherical portion as well as in the stalk connecting the vesicle to the hyphae. No fluorescence was observed in hyphae or vesicles of wild-type *F. alni* grown in (–)N media.

Stability of Plasmid pIGSAF in F. alni ACN14a in the Absence of Selection

In cultures grown without chloramphenicol selection, qPCR analysis did not find a significant difference in the amount of plasmid relative to genomic DNA after one, two, or three rounds of sub-culturing (Figure 5). Only after the fourth round of sub-culturing was a significant decrease from the initial plasmid concentration detected by qPCR (Figure 5), however, fluorescence throughout the hyphae was still observed even after four weeks without selection (Figure 6).

Discussion

200 *Genetic Transformation of* Frankia alni *ACN14a with an Unmethylated Replicating*
201 *Plasmid*

202 We have shown that *F. alni* can be stably transformed with an unmethylated
203 plasmid introduced by electroporation. A homolog of an Mrr Type IV methyladenine-
204 targeting restriction enzyme was highly expressed in *F. alni* in culture (Figure 1),
205 suggesting that DNA with methylated adenine bases would be degraded in this organism.
206 Actinobacteria, especially *Frankia*, expressed Type IV methyl-directed restriction
207 enzyme genes more highly than proteobacteria and firmicutes (Figure 2), a finding that
208 correlates with previous reports of higher transformation efficiencies with unmethylated
209 plasmids in *Corynebacterium* (Ankri *et al*., 1996) and *Streptomyces spp*. (Molle *et al*.,
210 1999). Genomes of the majority of actinobacteria are missing homologs of the *dam*
211 methyltransferase gene (Sachez-Romero *et al*., 2015) whose product is used to mark
212 parent DNA strands during replication, and also are missing *mutS* and *mutL* that form a
213 complex for the removal and repair of mismatched bases on the daughter strand
214 determined by the methylation of adenine residues (Sachadyn, 2010). Together, these
215 factors suggest a difference in preference for unmethylated over methylated DNA among
216 most of the actinobacteria relative to other bacterial phyla.

217 Type IV restriction enzymes have been suggested to have evolved as a counter to
218 phage methylation systems that themselves evolved to evade host restriction systems by
219 methylating restriction sites (Westra *et al*., 2012). Phage genomes adopt the methylation
220 patterns of their previous host (Loenen and Raleigh, 2014) thus increasing the likelihood
221 of digestion by actinobacterial enzymes if replicated in a *dam*+ host. The expression of
222 type IV restriction enzymes in actinobacteria therefore could represent an adaptation to

prevent infection by DNA phages based on the methylation state of their genomes. Differences in methylation patterns between actinobacteria and other bacterial phyla (Novella *et al*., 1996) potentially constitute a barrier to horizontal gene transfer between these groups, including phage-mediated gene transfer. Of particular interest to the evolution of root nodule symbioses is the transfer of genes between *Frankia* and the rhizobia and vice versa. It has been suggested that the *nodA* gene involved in Nod factor synthesis evolved in the actinobacteria, including some *Frankia*, and was then horizontally transferred to the rhizobia (Persson *et al*., 2015). With type IV restriction enzymes creating a barrier to horizontal transfer into actinobacteria from *dam*+ bacteria including proteobacteria, this direction would be more likely than the reverse.

However, *F. alni* was observed to down-regulate its type IV *mrr* gene substantially in symbiosis (Figure 1). As roots contain much lower concentrations of bacteriophage than the surrounding soil (Ward and Mahler, 1982) this likely represents a decreased necessity for restriction enzymes as a defense mechanism during symbiosis. A potential side-effect of this down-regulation, however, is that the barrier to horizontal transfer posed by type IV enzymes is likely lowered during symbiosis, promoting horizontal transfer to *Frankia* from other endophytic bacteria.

M. tuberculosis showed much lower transcription of its annotated type IV methyladenine targeting restriction enzyme than other actinobacteria. *M. tuberculosis* expresses an adenine methyltransferase in hypoxic conditions that regulates the expression of genes likely involved with survival during macrophage infection (Shell *et al*., 2013). For this reason it is likely that *M. tuberculosis* responds to methylated DNA differently than other actinobacteria, indeed electrotransformation of *M. tuberculosis* can

246 be readily achieved with methylated plasmids replicated in *E. coli* DH5α (Pelicic *et al*.,
247 1997), suggesting that methylated DNA is not digested in *M. tuberculosis*.

248 In this study plasmid pSA3 and its derivatives were capable of replication in *F.*
249 *alni*. This shows that the broad host-range origin is capable of replication in *Frankia* and
250 supports its use as a vector for the manipulation of *Frankia* spp. The parent plasmid of
251 pSA3, pIP501, replicates in a very broad range of bacteria including *Streptomyces*
252 *lividans* and *E. coli* (Kurenbach *et al*., 2003) indicating the potential for transformation of
253 additional actinobacteria with these plasmids.

254 *Differential Regulation of* egfp *Under the Control of the* Frankia alni *ACN14a* nif *Cluster*
255 *Promoter*

256 The expression of the *egfp* gene of plasmid pIGSAFnif was up-regulated in (–)N
257 media compared with expression in (+)N media, at proportional levels similar to the
258 expression of the *nifH* nitrogenase gene (Table 1), demonstrating for the first time that
259 expression of reporter genes can be manipulated in *Frankia*. This transformation system
260 resulted in the ability to visualize the expression of nitrogen-fixation genes *in vitro* by
261 fluorescence microscopy (Figure 4). Interestingly, fluorescence was detected in both the
262 spherical portion of the vesicle as well as in the stalk that connects to the hyphae,
263 suggesting that nitrogen fixation genes are expressed in both parts of the vesicle. Previous
264 studies have shown that the vesicle envelope is deposited around the stalk as well
265 (Lancelle *et al*., 1985), supporting the observation that nitrogen fixation occurs in the
266 stalk.

Although the fluorescence observed when *egfp* was expressed under the control of the *nif* cluster promoter was predominantly in the vesicles, some fluorescence was occasionally observed in hyphae under nitrogen-fixing conditions, whereas in (+)N media there was no observable fluorescence (Figure 4). This suggests that there is some expression of *nif* genes in the hyphae as well as the vesicles induced by nitrogen limitation. *Frankia spp*. in symbiosis with members of the Casuarinaceae have been reported to fix nitrogen in hyphae, since no vesicles are differentiated (Murry *et al*., 1985); this correlated with the formation of a lignified host cell wall in the symbiotic tissue that likely excludes oxygen (Berg and McDowell, 1987). In liquid culture, there may be zones of low pO_2 that develop in portions of the *Frankia* hyphal colony where nitrogen fixation could be induced.

Frankia spp. in symbiosis have been suggested to be more autonomous than rhizobial microsymbionts due to their ability to control the flow of oxygen with the formation of vesicles as well as the expression of more metabolic pathways in the microsymbiont. These factors allow *Frankia* to be more metabolically independent from their hosts (Alloisio *et al*., 2010; Berry *et al*., 2011). As an example, both *Frankia* and the rhizobia share a second glutamine synthetase gene, named *glnII*, which is not present in most other bacteria (Ghoshroy *et al*., 2010). *glnII* is up-regulated in *F. alni* in symbiosis relative to nitrogen-replete culture (Alloisio *et al*., 2010) but not differentially expressed or required for effective nodule formation by rhizobia such as *Sinorhizobium meliloti* (Becker *et al*., 2004; de Bruijn *et al*., 1989). The development of genetic tools for the manipulation of *Frankia* will allow further experimentation into these and other

289 distinctive molecular aspects of actinorhizal symbioses, which will, in turn, inform
290 analyses of the evolution of root nodule symbiosis.

291 *Future Directions*

292 Even in the absence of selection plasmid pIGSAF was found to be stable in *F.*
293 *alni* cultures for at least three weeks (Figure 5) and cultures continued to fluoresce after
294 at least four weeks (Figure 6). Three to four weeks is sufficient to nodulate *F. alni* hosts
295 (Alloisio *et al*., 2010), suggesting that transformants can be used to inoculate plants in
296 studies of the role of *Frankia* and its interactions with hosts during nodule establishment
297 and symbiosis. In future this system could be modified using recombination or viral
298 integrases to anchor genes within the genome. Differential regulation of reporter genes
299 such as *egfp* can be used to localize the expression of genes identified by genomics and
300 transcriptomics in specific *Frankia* cell types, in different growth conditions, and in
301 symbiosis. Replicating plasmids may also enable the study of gene function by
302 constitutive expression of selected genomic genes, by promoter switching or by knock-
303 down experiments expressing anti-RNAs to genes of interest (Gillaspie *et al*., 2009).
304 Circumventing the natural restriction systems of *Frankia* will also increase the
305 transformation rate of non-replicating plasmids and enable higher efficiency
306 recombination for gene knock-out experiments as attempted by Kucho *et al.* (2009).

307 **Materials and Methods**

308 *Restriction Enzyme Analysis*

309 Genes annotated as restriction enzymes (Restriction enzyme types I, II, III, or IV)
310 present in the annotations of all completely sequenced actinobacterial genomes and their

311 annotations were downloaded from the REBASE database (Roberts *et al*., 2010).
312 Bacterial transcriptomes used in this study were downloaded from the NCBI GEO
313 database (Supplementary Table 1). Transcriptomes were chosen based on the following
314 criteria: 1) transcriptomes were made from pure cultures in log-phase growth, 2)
315 organism did not have genetic manipulations including mutations or exogenous plasmids,
316 and 3) cultures did not have additional experimental compounds added including
317 antibiotics or complex carbon sources. For comparisons between transcriptomes,
318 expression levels for each gene were calculated as the percent of genes with lower
319 expression than the gene of interest in each transcriptome.

320 *Culture Conditions*

321 *E. coli* strains DH5α and GM48 (*dam*- and *dcm*-) were grown in 50ml Difco 1.5%
322 (w/v) Luria Broth (LB) (Catalog #241420, Becton Dickinson, Franklin Lakes, NJ), pH
323 6.8, in 250mL flasks at 37°C with shaking at 150 rpm overnight. Plates were made with
324 1.5% (w/v) Bacto Agar (Becton Dickinson, Franklin Lakes, NJ, catalog #214010) in
325 1.5% LB, and incubated overnight at 37°C. All media were first sterilized by autoclaving
326 for 30 minutes at 121°C.

327 *Frankia alni* ACN14a (Normand and Lalonde, 1982) was cultured in 5ml liquid
328 BAPP media modified from Murry *et al*., 1984 by the addition of 5mM pyruvate and
329 5mM MOPS, adjusted to pH 6.7, in sterile 25ml glass test tubes with tetracycline added
330 to a final concentration of 10µg/ml to prevent contamination. (+)N media included 5mM
331 ammonium chloride as a nitrogen source whereas (–)N media had no added nitrogen
332 source. For sub-culturing, hyphae were collected in sterile 2ml microcentrifuge tubes,

centrifuged at 10,000rpm for ten minutes in a tabletop microcentrifuge (model #5424, Eppendorf, Hamburg, Germany), and re-suspended in 1mL fresh BAPP media. Cultures were then homogenized by passage through a 21G needle six times, then homogenate equivalent to 50ul packed-cell volume as measured after centrifugation was added to 4ml of fresh media and incubated at 28°C without shaking. Cultures were routinely sub-cultured once per week. For selection of transformants, chloramphenicol was added to a final concentration of 25µg/ml.

Frankia Transformation

F. alni cells were grown in culture for one week prior to transformation. Two milliliters of growth media with hyphae was then pelleted as above and the pellet was re-suspended in 500µL of ice-cold deionized (DI) water that had been sterilized by autoclaving. This was repeated two more times, but on the third round of centrifugation the pellet was re-suspended in 300µL ice-cold sterile 10% glycerol instead. Hyphae in the cell suspension were then homogenized by passage through a 21G needle twice.

300µL of the *F. alni* cell suspension was pipetted into an electroporation cuvette with a 2mm gap (Molecular BioProducts Catalog #5520, San Diego, CA) and mixed with 10µg of plasmid DNA. The cuvette was then incubated on ice for five minutes. Electroporation was carried out in a Bio-Rad Gene Pulser™ with Pulse Controller at 2.5kV, 200Ω resistance, and 25µF of capacitance. The cuvette was then immediately filled with 1ml of ice-cold BAPP media. The cuvette was sealed with Parafilm® (Pechiney Plastic Packaging, Menasha, WI) and incubated overnight without shaking at 28°C.

355 The following day the *F. alni* culture was removed from the cuvette and added to
356 3.5ml of sterile BAPP media with tetracycline as above, in a glass test tube. The culture
357 was incubated at 28°C without shaking until visible hyphae were observed
358 (approximately 10 days after electroporation). At this point *F. alni* hyphae were sub-
359 cultured into fresh media as above and then once more one week later. The following
360 week (two weeks after visible hyphae were observed) the hyphae were sub-cultured again
361 into BAPP media, this time with chloramphenicol added for selection to a final
362 concentration of 25μg/ml. Chloramphenicol was chosen as the selective antibiotic
363 because all *Frankia* strains tested in a study by Tisa *et al*. (1998) were susceptible to it.
364 This process was repeated the following week for an additional round of selection.

365 *DNA Extraction*

366 Plasmids were purified from *E. coli* using a QIAprep® Spin Miniprep Kit
367 (Catalog #27106, Qiagen). Two milliliters of overnight cultures were pelleted at
368 13,000rpm for five minutes and used for extraction. Plasmids were resuspended in EB
369 buffer (Qiagen) and quantified on a NanoDrop Microvolume Spectrophotometer
370 (ThermoFisher). Plasmids were further purified by running on a 0.7% agarose gel and
371 extracted with a Zymoclean Gel DNA Recovery Kit (Catalog #11-300, Genesee
372 Scientific, San Diego, CA). To synthesize unmethylated plasmids, methylated plasmids
373 were first extracted from *E. coli* DH5α and transformed into *E. coli* GM48 by heat shock
374 in $CaCl_2$ (Sambrook *et al.*, 1989) then cultured and re-extracted as above.

375 *F. alni* DNA extraction was carried out with a CTAB protocol (Feil *et al*., 2012).
376 Briefly, cells were pelleted at 10,000rpm for five minutes in a tabletop centrifuge. Cells

were then re-suspended in TE buffer and lysed with lysozyme, SDS, Proteinase K, and CTAB. DNA was extracted with 24:1 chloroform:isoamyl alcohol, and 25:24:1 phenol:chloroform:isoamyl alcohol, washed with isopropanol followed by ethanol, resuspended in 50ul TE buffer, and quantified by NanoDrop.

Plasmid Synthesis

A plasmid designated as pIGSAF was synthesized by ligating a PCR-amplified fragment containing the *egfp* gene and promoter from plasmid pDiGc (Helaine *et al*., 2010, Addgene, Cambridge, MA) to plasmid pSA3 (Dao and Ferretti, 1985). Plasmid pSA3 is a broad host-range replicating plasmid developed as a shuttle vector from pIP501. The pIP501 origin of replication has been shown to replicate in bacteria of diverse phyla including Firmicutes, from which it was originally isolated (Horodniceanu *et al*., 1976), as well as Actinobacteria and Proteobacteria (Kurenbach *et al*., 2003). Plasmid pSA3 contains chloramphenicol and tetracycline resistance genes and an *E. coli*-specific origin of replication for propagation (Dao and Ferretti, 1985).

Primers for PCR, listed in Table 2, were designed using NCBI Primer-BLAST with default settings (Ye *et al*., 2012). For cloning, primers were designed with linkers adding target restriction sites onto their 5' ends (Table 2) as well as 4-6 additional bases to aid in restriction digestion of the ends. A diagram of the addition of linkers by PCR is shown in Figure 7A. For the synthesis of pIGSAF, primers were designed targeting the *egfp* coding region as well as the promoter region 200 base pairs upstream. PCR was performed in a Bio-Rad S1000 Thermal Cycler (Bio-Rad, Hercules, CA) using a Qiagen Taq PCR kit (catalog #201223). Products were synthesized as in Figure 7A, first

399 amplified for 10 cycles using the annealing temperature of the region corresponding to
400 the binding site on the target DNA and then an additional 30 cycles using the annealing
401 temperature of the full primer including the linker (Table 2).

402 The *egfp* fragment was synthesized with SalI and BamHI restriction sites on the 5'
403 and 3' ends, respectively, with primers EGFP_SalI_F and EGFP_BamHI_R. The PCR
404 product and plasmid pSA3 were then digested with both enzymes. The fragments were
405 purified on an agarose gel as above and then mixed together, denatured by heating for
406 five minutes at 65°C, and then ligated together by incubation with T4 ligase (catalog
407 #M0202S, Qiagen) at 16°C overnight. The resulting ligation was transformed into *E. coli*
408 DH5α by heat shock (Sambrook *et al.*, 1989). Transformants were selected on LB plates
409 with chloramphenicol at a final concentration of 25μg/ml. Transformed colonies were
410 inoculated into liquid LB media with chloramphenicol and cultured as above. Plasmid
411 pIGSAF was then re-extracted from transformed *E. coli* and its composition was
412 confirmed by digestion with SalI and BamHI as above.

413 The presence of plasmids pSA3 and pIGSAF electroporated into *F. alni* cultures
414 was confirmed with PCR. Genomic DNA was extracted and amplified with primers pairs
415 pSA3_Cm_F/pSA3_CmR for the chloramphenicol resistance gene of plasmid pSA3 and
416 GFP_qPCR_F/GFP_qPCR_R for the *egfp* gene of plasmid pIGSAF. PCR products were
417 separated on an agarose gel and extracted as above, and sequenced by the UCDNA
418 Sequencing Facility (Davis, CA). Sequences were compared by BLAST against gene
419 sequences obtained from the original plasmids (Supplementary Table 2).

For the differential expression of *egfp* by nitrogen limitation under the control of the *nif* cluster promoter region of *F. alni* ACN14a, plasmid pIGSAFnif was synthesized. Ligation of the *nif* cluster promoter region to the coding region of *egfp* was carried out as outlined in Figure 7B. The *egfp* coding region of pDiGc was amplified without the upstream promoter region using primers GFP_CDS_EcoRI_F and GFP_CDS_SalI_R (Table 2). An EcoRI restriction site was added before the start codon and a SalI site was added 200 bases downstream of the stop codon. Separately, the 300 bases upstream of the *nif* nitrogenase cluster in the *F. alni* ACN14a genome were amplified with an XbaI site upstream and an EcoRI site downstream using primers nif_promoter_XbaI_F and nif_promoter_EcoRI_R (Table 2). These two PCR products were digested with EcoRI (catalog #R0101S, New England Biolabs) and ligated together. The ligation product was then re-amplified by PCR using the *nif* promoter forward primer and the *egfp* reverse primer. The amplified ligation product and plasmid pSA3 were then each digested with SalI and XbaI, ligated together, transformed, and selected on chloramphenicol as above.

qPCR Verification of Differential Regulation

Quantifications of gene expression and fold-changes were obtained based on qPCR amplification. Cultures of transformed cells were grown in 4.0 ml (+)N and (–)N media in sterile six-well plates (catalog #353046, Corning Inc, Corning, NY) with shaking at 50rpm at 28°C. After five days, RNA was extracted by bead-beating, following a protocol adapted from Dietrich *et al*. (2000): *F. alni* hyphae were pelleted at 9000rpm for fifteen minutes, resuspended in 1050ul Buffer RLT (Qiagen), and transferred to 2ml tubes containing Lysing Matrix B (catalog #6911-100, MP Biomedicals, Burlingame, CA). Samples were processed with a FastPrep FP120 (Thermo

Fisher Scientific, Waltham, MA) for 45 seconds at setting 6.5, then placed on ice for 45 seconds. The processing step was then repeated twice more with cooling on ice between each step. Supernatants were then transferred to Qiagen RNeasy spin columns and purified with a Qiagen RNeasy Mini Kit (catalog #74104). RNA was eluted in RNase-free water and then contaminating DNA was digested with an Invitrogen TURBO DNA-free Kit (catalog #AM1907, Waltham, MA). Finally, cDNA was synthesized with an Invitrogen Superscript III Kit (catalog #18080051) with random hexamer primers. qPCR was performed on a 7500 Fast Real-Time PCR System (Applied Biosystems, Foster City, CA) with Fast SYBR Green qPCR Master Mix (catalog #4385612, Applied Biosystems) and primer pairs rpoD_qPCR_F/rpoD_qPCR_R, nifH_qpCR_F/nifH_qpCR_R, infC_qPCR_F/infC_qPCR_R, and gfp_qPCR_F/gfp_qPCR_R (Table 1). *egfp* and *nifH* were used as experimental targets to confirm differential regulation of *egfp* by nitrogen limitation; housekeeping gene *infC* was used for normalization as in Alloisio *et al*. (2010); and sigma factor gene *rpoD* was used as a negative control.

Confocal Microscopy

For visualization of hyphae and GFP fluorescence, *Frankia* cells were grown either in (+)N or (–)N BAPP medium as above and immobilized on glass slides with a drop of 3% molten agarose solution: 1.5g molecular biology-grade agarose (catalog #A9539, Sigma, St. Louis, MO) was dissolved in 50ml DI H_2O and kept warm in a water bath at 50°C. Slides were pre-heated on a slide warmer (Fisher) at 50°C. Then 15ul of *Frankia* hyphae were pipetted onto each slide and covered with 35ul of 3% molten agarose. A #1.5 coverslip was added to the *Frankia* cells in agarose, which were allowed to cool to room temperature. The *Frankia* preparations were visualized on a Leica TCS

SP8 STED 3X confocal microscope with a 100X oil-immersion objective and a HyD detector. For fluorescence imaging, samples were excited with 488nm light. The emission wavelengths were collected from 500-550nm. Images were stored as .lif files from Leica LAS X and then viewed in FIJI (Schindelin *et al*., 2012). To visualize hyphal three-dimensional structure, Z-stack images were taken and then combined using the highest fluorescent intensity of each pixel (FIJI MAX setting).

Plasmid Stability

In order to test the persistence of a plasmid in transformed *F. alni*, cultures containing plasmid pIGSAF were grown without selection in non-selective media lacking chloramphenicol. The cells were first pelleted and re-suspended in fresh BAPP media, then sub-cultured as above without the addition of chloramphenicol to the media. These cultures were grown in six-well plates as above for one week. At that time each culture was pelleted and re-suspended in 500ul fresh BAPP media. This suspension was homogenized by passage through a 21G needle twice. 250ul of the homogenate was transferred to fresh BAPP media without chloramphenicol and the remaining 250ul was used for total genomic DNA extraction as above. This process was repeated once per week for four weeks. The amount of plasmid in each sample was quantified by qPCR, performed in duplicate for each of three biological replicates, using *egfp* primers GFP_qPCR_F and GFP_qPCR_R (Table 2). Relative fold-change of plasmid between each time point was calculated using the ΔΔCt method (Lee *et al*., 2006) with the *infC* gene as a control to normalize the amount of DNA in each sample (amplified with primers infC_qPCR_F and infC_qPCR_R, Table 1). To determine significant changes in plasmid abundance, two-tailed Welch's t-tests were performed in R on normalized ΔCt

489 values calculated ($p<.05$). Cultures grown with selection and without selection for four
490 weeks were also imaged with fluorescence as above.

491 **Acknowledgements**

492 We thank Dr. Rebecca Parales for providing plasmid pSA3 and Dr. Wolf Heyer
493 for *E. coli* strain GM48. We also thank Dr. Ingrid Brust-Mascher for technical assistance
494 with confocal microscopy. IG was supported by a UC Davis Department of Plant
495 Sciences graduate student research fellowship. Research funding for this project was
496 provided by UC Davis Henry A. Jastro Research Grants to IG, and by USDA-NIFA-CA-
497 D-PLS-2173-H (AB).

498 **References**

499 1) Alloisio N, Queiroux C, Fournier P, Normand P, Vallenet D, Medigue C,
500 Yamaura M, Kakoi K, and Kucho K. 2010. The *Frankia alni* symbiotic
501 transcriptome. Molecular Plant Microbe Interactions 23(5): 593-607.
502 2) Ankri S, Reyes O, and Leblon G. 1996. Electrotransformation of highly DNA-
503 restrictive *Corynebacteria* with synthetic DNA. Plasmid 35: 62-66.
504 3) Appleby CA. 1984. Leghemoglobin and *Rhizobium* respiration. Ann. Rev. Plant
505 Physiol. 35: 443-78.
506 4) Battenberg K, Potter D, Tabuloc C, Chiu JC, and Berry AM. 2018. Comparative
507 transcriptomics of two actinorhizal plants and the legume *Medicago truncatula*
508 support the homology of root nodule symbioses and is congruent with a two-step
509 process of evolution in the nitrogen-fixing clade of angiosperms. Frontiers in
510 Plant Science, in press.

5) Becker A, Berges H, Krol E, Bruand C, Ruberg S, Capela D, Lauber E, Meilhoc E, Ampe F, de Bruijn FJ, Fourment J, Francez-Charlot A, Kahn D, Kuster H, Liebe C, Puhler A, Weidner S, and Batut J. 2004. Global changes in gene expression in *Sinorhizobium meliloti* 1021 under microoxic and symbiotic conditions. Molecular Plant-Microbe Interactions 17(3): 292-303.

6) Benson DR and Silvester WB. 1993. Biology of *Frankia* strains, actinomycete symbionts of actinorhizal plants. Microbiological Reviews 57(2): 293-319.

7) Berg RH and McDowell L. 1987. Endophyte differentiation in *Casuarina* actinorhizae. Protoplasma 136: 104-117.

8) Berry AM, Harriott OT, Moreau RA, Osman SF, Benson DR. and Jones AD. 1993. Hopanoid lipids compose the Frankia vesicle envelope, presumptive barrier of oxygen diffusion to nitrogenase. PNAS 90(13): 6091-6094.

9) Berry AM, Mendoza-Herrera A, Guo Y, Hayashi J, Persson T, Barabote R, Demchenko K, Zhang S, and Pawlowski K. 2011. New perspectives on nodule nitrogen assimilation in actinorhizal symbioses. Functional Plant Biology 38: 645-652.

10) Bickhart DM and Benson DR. 2011. Transcriptomes of *Frankia* sp. strain CcI3 in growth transitions. BMC Microbiology 11(1): 1.

11) Castaneda-Garcia A, Prieto AI, Rodriguez-Beltran J, Alonso N, Cantillon D, Costas C, Perez-Lago L, Zegeye ED, Plocinski P, Tonjum T, Garcia de Viedma D, Paget M, Waddell SJ, Rojas AM, Doherty AJ, and Blazquez J. 2017. A non-canonical mismatch repair pathway in prokaryotes. Nature Communications 8: 14246.

12) Ceremonie H, Debelle F, and Fernandez MP. 1999. Structural and functional comparison of Frankia root hair deforming factor and rhizobia Nod factor. Canadian Journal of Botany 77(9): 1293-1301.

13) Cissoko M, Hocher V, Gherbi H, Gully D, Carre-Mlouka A, Sane S, Pignoly S, Champion A, Ngom M, Pujic P, Fournier P, Gtari M, Swanson E, Pesce C, Tisa LS, Sy MO, and Svistoonoof S. 2018. Actinorhizal signaling molecules: *Frankia* root hair deforming factor shares properties with NIN inducing factor. Front. Plant Sci. 9: 1494.

14) Claunch KM, Bush M, Evans CR, Malmquist JA, Hale MC, and McGillivray SM. 2018. Transcriptional profiling of the clpX mutant in *Bacillus anthracis* reveals regulatory connection with the lrgAB operon. Microbiology 164(4):659-669.

15) Dao ML and Ferretti JJ. 1985. *Streptococcus-Escherichia coli* shuttle vector pSA3 and its use in the cloning of streptococcal genes. Applied and Environmental Microbiology 49(1): 115-119.

16) Daugherty A, Powers KM, Standley MS, Kim CS, and Purdy GE. 2011. *Mycobacterium smegmatis* RoxY is a repressor of oxyS and contributes to resistance to oxidative stress and bactericidal ubiquitin-derived peptides. J Bacteriol 193(24):6824-33.

17) De Bruijn FJ, Rossbach S, Schneider M, Ratet P, Messmer S, Szeto WW, Ausubel FM, and Schell J. 1989. *Rhizobium meliloti* 1021 has three differentially regulated loci involved in glutamine biosynthesis, none of which is essential for symbiotic nitrogen fixation. Journal of Bacteriology 171(3): 1673-1682.

18) Dietrich G, Schaible UE, Diehl KD, Mollenkopf H, Wiek S, Hess J, Hagens K, Kaufmann SH, and Knapp B. 2000. Isolation of RNA from mycobacteria grown under *in vitro* and *in vivo* conditions. FEMS Microbiol Lett. 186(2): 177-80.

19) Feil WS, Feil H, and Copeland A. 2012. Bacterial genomic DNA isolation using CTAB. Joint Genome Institute.

20) Fu LM and Shinnick TM. 2007. Genome-wide exploration of the drug action of capreomycin on *Mycobacterium tuberculosis* using Affymetrix oligonucleotide GeneChips. J Infect 54(3):277-84.

21) Gage DJ. 2002. Analysis of infection thread development using Gfp- and DsRed-expressing *Sinorhizobium meliloti*. Journal of Bacteriology 184(24): 7042-7046.

22) Ghoshroy S, Binder M, Tartar A, and Robertson DL. 2010. Molecular evolution of glutamine synthetase II: Phylogenetic evidence of a non-endosymbiotic gene transfer event early in plant evolution. BMC Evolutionary Biology 10:198.

23) Gillaspie D, Perkins I, Larsen K, McCord A, Pangonis S, Sweger D, Seleem MN, Sriranganathan N, and Anderson BE. 2009. Plasmid-based system for high-level gene expression and antisense gene knockdown in *Bartonella henselae*. Applied and Environmental Microbiology 75(16): 5434-5436.

24) Griesmann M, Chang Y, Liu X, Song Y, Haberer G, Crook MG, Billault-Penneteau B, Lauressergues D, Keller J, Imanishi L, Roswanjaya YP, Kohlen W, Pujic P, Battenberg K, Alloisio N, Liang Y, Hilhorst H, Salgado MG, Hocher V, Gherbi H, Svistoonoff S, Doyle JJ, He S, Xu Y, Xu S, Qu J, Gao Q, Fang X, Fu Y, Normand P, Berry AM, Wall LG, Ane J, Pawlowski K, Xu X, Yang H, Spannagl M, Mayer KFX, Wong GK, Parniske M, Delaux P, and Cheng S. 2018.

579 Phylogenomics reveals multiple losses of nitrogen-fixing root nodule symbiosis.
580 Science 24: eaat1743.
581 25) Gutierrez RA. 2012. Systems biology for enhanced plant nitrogen nutrition.
582 Science 336: 1673-1675.
583 26) Hahn D, Amann RI, and Zeyer J. 1993. Whole-cell hybridization of *Frankia*
584 strains with fluorescence- or digoxigenin-labeled, 16S rRNA-targeted
585 oligonucleotide probes. Applied and Environmental Microbiology 59(6): 1709-
586 1716.
587 27) Helaine S, Thompson JA, Watson KG, Liu M, Boyle C, and Holden DW. 2010.
588 Dynamics of intracellular bacterial replication at the single cell level. Proceedings
589 of the National Academy of Sciences of the United States of America 107(8):
590 3746-51.
591 28) Horodniceanu T, Bouanchaud DH, Bieth G, and Chabbert YA. 1976. R plasmids
592 in *Streptococcus agalactiae* (Group B). Antimicrobial Agents and Chemotherapy
593 10(5): 795-801.
594 29) Indest KJ, Eberly JO, Hancock DE, Jung CM, Carr MR, and Blakeney GA. 2016.
595 *Rhodococcus jostii* RHA1 TadA-homolog deletion mutants accumulate less
596 polyhydroxyalkanoates (PHAs) than the parental strain. The Journal of General
597 and Applied Microbiology, 62(4), 213-216.
598 30) Kucho KI, Kakoi K, Yamaura M, Higashi S, Uchiumi T, and Abe M. 2009.
599 Transient transformation of *Frankia* by fusion marker genes in liquid culture.
600 Microbes and Environments 24(3): 231-240.

601 31) Kurenbach B, Bohn C, Prabhu J, Abudukerim M, Szewzyk U, and Grohmann E.
602 2003. Intergeneric transfer of the *Enterococcus faecalis* plasmid pIP501 to
603 *Escherichia coli* and *Streptomyces lividans* and sequence analysis of its *tra* region.
604 Plasmid 50(1): 86-93.
605 32) Lancelle SA, Torrey JG, Hepler PK, and Callaham DA. 1985. Ultrastructure of
606 freeze-substituted *Frankia* strain HFPCcI3, the actinomycete isolated from root
607 nodules of *Casuarina cunninghamiana*. Protoplasma 127: 64-72.
608 33) Lee C, Kim J, Shin SG, and Hwang S. 2006. Absolute and relative QPCR
609 quantification of plasmid copy number in *Escherichia coli*. Journal of
610 Biotechnology 123: 273-280.
611 34) Li L, Guo J, Wen Y, Chen Z, Song Y, and Li J. 2010. Overexpression of ribosome
612 recycling factor causes increased production of avermectin in *Streptomyces*
613 *avermitilis* strains. Journal of industrial microbiology & biotechnology, 37(7),
614 673-679.
615 35) Liu M, Durfee T, Cabrera JE, Zhao K, Jin DJ, and Blattner FR. 2005. Global
616 transcriptional programs reveal a carbon source foraging strategy by *Escherichia*
617 *coli*. *Journal of Biological Chemistry 280*(16), 15921-15927.
618 36) Loenen WAM and Raleigh EA. 2014. The other face of restriction: modification-
619 dependent enzymes. Nucleic Acids Research 42(1): 56-69.
620 37) Long SR, Buikema WJ, and Ausubel FM. 1982. Cloning of *Rhizobium meliloti*
621 nodulation genes by direct complementation of Nod- mutants. Nature 298: 485-
622 488.

623 38) Molle V, Palframan WJ, Findlay KC, and Buttner MJ. 1999. WhiD and WhiB,
624 homologous proteins required for different stages of sporulation in *Streptomyces*
625 *coelicolor* A3(2). Journal of Bacteriology 182(5): 1286-1295.
626 39) Murry MA, Fontaine MS, and Torrey JG. 1984. Growth kinetics and nitrogenase
627 induction in *Frankia* sp. ArI3 grown in batch culture. Plant and Soil 78: 61-78.
628 40) Murry MA, Zhongze Z, and Torrey JG. 1985. Effect of O_2 on vesicle formation,
629 acetylene reduction, and O_2-uptake kinetics in *Frankia* sp. HFPCcI3 isolated from
630 *Casuarina cunninghamiana*. Canadian Journal of Microbiology 31(9): 804-809.
631 41) Myers AK and Tisa LS. 2003. Effect of electroporation conditions on cell
632 viability of *Frankia* EuI1c. Plant and Soil 254(1): 83-88.
633 42) Normand P and Lalonde M. 1982. Evaluation of *Frankia* isolated from
634 provenances of two *Alnus* species. Can. J. Microbiol. 28: 1133-1142.
635 43) Novella IS, Marin I, and Sanchez J. 1996. Restriction analysis of actinomycetes
636 chromosomal DNA. Can. J. Microbiol. 42: 201-206.
637 44) Okazaki S, Kaneko T, Sato S, and Saeki K. 2013. Hijacking of leguminous
638 nodulation signaling by the rhizobial type III secretion system. PNAS: 201302360.
639 45) Oldroyd GED. 2013. Speak, friend, and enter: signaling systems that promote
640 beneficial symbiotic associations in plants. Nature Reviews Microbiology 11:
641 252-263.
642 46) Pelicic V, Jackson M, Reyrat J, Jacobs WR, Gicquel B, and Guilhot C. 1997.
643 Efficient allelic exchange and transposon mutagenesis in *Mycobacterium*
644 *tuberculosis*. Proc. Natl. Acad. Sci. USA 94: 10955-10960.

645 47) Persson T, Battenberg K, Demina IV, Vigil-Stenman T, Vanden Heuvel B, Pujic
646 P, Facciotti MT, Wilbanks EG, O'Brien A, Fournier P, Hernandez MAC, Herrera
647 AM, Medigue C, Normand P, Pawlowski K, and Berry AM. 2015. *Candidatus*
648 Frankia datiscae Dg1, the actinobacterial microsymbiont of *Datisca glomerata*,
649 expresses the canonical *nod* genes *nodABC* in symbiosis with its host plant. PLoS
650 One 10(5): e0127630.
651 48) Roberts RJ, Vincze T, Posfai J, and Macelis D. 2009. REBASE-a database for
652 DNA restriction and modification: enzymes, genes, and genomes. Nucleic Acids
653 Research 38(suppl_1): D234-D236.
654 49) Sachadyn P. 2010. Conservation and diversity of MutS proteins. Mutation
655 Research 694: 20-30.
656 50) Sambrook J, Fritsch EF, and Maniatis T. 1989. Molecular Cloning: A Laboratory
657 Manual. Second Edition. Cold Spring Harbor Laboratory Press.
658 51) Sanchez-Romero MA, Cota I, and Casadesus J. 2015. DNA methylation in
659 bacteria: from the methyl group to the methylome. Current Opinion in
660 Microbiology 25: 9-16.
661 52) Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T,
662 Preibisch S, Rueden C, Saalfeld S, Schmid B, Tinevez J, White DJ, Hartenstein V,
663 Eliceiri K, Tomancak P, and Cardona A. 2012. Fiji: an open-source platform for
664 biological-image analysis. Nature Methods 9: 676-682.
665 53) Shell SS, Prestwich EG, Baek S, Shah RR, Sassetti CM, Dedon PC, and Fortune
666 SM. 2013. DNA methylation impacts gene expression and ensures hypoxic
667 survival of Mycobacterium tuberculosis. PLoS Pathogens 9(7): e1003419.

668 54) Simonet P, Normand P, Hirsch AM, and Akkermans ADL. 1990. The genetics of
669 the *Frankia*-Actinorhizal symbiosis. Molecular Biology of Symbiotic Nitrogen
670 Fixation. Edited by Peter M. Gresshoff. 77-109.
671 55) Spaepen S, Das F, Luyten E, Michiels J, and Vanderleyden J. 2009. Indole-3-
672 acetic acid-regulated genes in *Rhizobium etli* CNPAF512. FEMS Microbiology
673 Letters 291(2): 195-200.
674 56) Soltis DE, Soltis PS, Morgan DR, Swensen SM, Mullin BC, Dowd JM, and
675 Martin PG. 1995. Chloroplast gene sequence data suggest a single origin of the
676 predisposition for symbiotic nitrogen fixation in angiosperms. Proc. Natl. Acad.
677 Sci. USA 92: 2647-2651.
678 57) Tisa LS, Chval MS, Krumholz GD, and Richards J. 1998. Antibiotic resistance
679 patterns of *Frankia* strains. Can. J. Bot. 77: 1257-1260.
680 58) Ward RL and Mahler RJ. 1982. Uptake of bacteriophage f2 through plant roots.
681 Applied and Environmental Microbiology 43(5): 1098-1103.
682 59) Werner GDA, Cornwell WK, Sprent JI, Kattge J, and Kiers ET. 2014. A single
683 evolutionary innovation drives the deep evolution of symbiotic N_2-fixation in
684 angiosperms. Nature Communications 5: 4087.
685 60) Westra ER, Swarts DC, Staals RHJ, Jore MM, Brouns SJJ, and van der Oost J.
686 2012. The CRISPRs, they are a-changin': how prokaryotes generate adaptive
687 immunity. Annual Review of Genetics 46:311-339.
688 61) Ye J, Coulouris G, Zaretskaya I, Cutcutache I, Rozen S, and Madden TL. 2012.
689 Primer-BLAST: A tool to design target-specific primers for polymerase chain
690 reaction. BMC Bioinformatics 13(134): 1-11.

691 62) Zhang X, Chen W, Zhang Y, Jiang L, Chen Z, Wen Y, and Li J. 2012. Deletion of
692 *ku* homologs increases gene targeting frequency in *Streptomyces avermitilis*.
693 Journal of Industrial Microbiology and Biotechnology 39(6): 917-925.

694 **Figures and Tables**

695 **Figure 1** Restriction enzymes annotated in the *Frankia* genomes and expression levels of
696 each gene in the transcriptome in culture when available (Alloisio *et al*., 2010 and
697 Bickhart and Benson, 2011). NA: Transcriptome not available.

698 **Figure 2** Restriction enzymes annotated in the genomes of *Frankia*, other actinobacteria,
699 and proteobacteria and firmicutes for comparison. For each, expression levels for
700 available transcriptomes growing in pure culture are provided.

701 **Figure 3** Transformation of *Frankia alni* ACN14a with a plasmid expressing *egfp*. Top:
702 Wild-type *F. alni* ACN14a control. Bottom: Transformed *F. alni* ACN14a with plasmid
703 pIGSAF. Images were obtained on a confocal microscope with both brightfield and
704 epifluorescence and then overlaid.

705 **Figure 4** Transformation of *Frankia alni* ACN14a with a plasmid expressing *egfp* under
706 the control of the *nif* cluster promoter region from the *F. alni* genome. Transformed
707 cultures were grown in both (+)N and (–)N media to compare fluorescence with and
708 without nitrogen fixation. Top: *F. alni* ACN14a transformed with plasmid pIGSAFnif
709 grown in (+)N media. Middle: *F. alni* ACN14a with plasmid pIGSAFnif grown in (–)N
710 media. Bottom: Close-up of a vesicle showing fluorescence in the stalk. Images were
711 obtained on a confocal microscope with both brightfield and epifluorescence and then
712 overlaid.

713 **Figure 5** Maintenance of plasmid pIGSAF in *Frankia alni* in culture without
714 chloramphenicol selection over the course of four weeks. Each week the culture was sub-
715 cultured into fresh media and genomic DNA was extracted to measure relative plasmid
716 concentrations via qPCR. Groups 'a' and 'b' indicate time points that are not significantly
717 different from each other ($p<.05$).

718 **Figure 6** Fluorescence of a *Frankia alni* hyphal colony transformed with plasmid
719 pIGSAF grown with chloramphicol selection (top) and without selection for four weeks
720 (bottom).

721 **Figure 7** PCR amplification and digestion and ligation overview for plasmid synthesis
722 used in this study. Linkers including restriction sites are colored red, blue, or green. A)
723 PCR products for cloning were amplified by amplification at a lower initial temperature
724 to allow the incorporation of additional restriction sites on the 5' end of the primers
725 (shown in red and blue). The annealing temperature was then increased to amplify full-
726 length products with the added restriction sites. B) Addition of the *F. alni* ACN14a *nif*
727 cluster promoter region to the *egfp* coding sequence. Both the promoter region and *egfp*
728 coding sequence were amplified with EcoRI sites (blue). These were then digested and
729 ligated together followed by amplification of the ligation product. Restriction sites on the
730 ends (red and green) were digested, allowing incorporation of the *nif* promoter:*egfp*
731 product into the digested plasmid.

732 **Table 1** qPCR verification of differential regulation of *egfp* in (+)N and (–)N media.
733 qPCR tested fold-change of *egfp* as well as *nifH* as a nitrogen-fixation positive control
734 and *rpoD* as a housekeeping negative control.

735 **Table 2** Primers designed for this study. Restriction sites added to primers by linker PCR
736 for use in cloning are bolded in the primer sequence and annealing temperatures are listed
737 both for the initial reaction ("Without Linker") and the main amplification phase of the
738 PCR ("Full Sequence").

739 **Supplementary Figures and Tables**

740 **Supplementary Table 1** List of genomes and transcriptomes used in this study.
741 Transcriptomes are listed with their GEO accession number and original reference.

742 **Supplementary Table 2** Sequences from PCR products of *Frankia alni* transformants
743 corresponding to the chloramphenicol resistance gene (*camr*) from pSA3 and *egfp* gene
744 of plasmid plasmid pIGSAF and pDiGc.

Gene ID	Enzyme Type	Percentile (Culture)	Percentile (Symbiosis)
***Frankia alni* ACN14a**			
FRAAL1375	I	29.2	12.1
FRAAL4992	I	92.6	95.8
FRAAL0249	II	91.2	92.8
FRAAL3915	II	14.7	14.7
FRAAL6331	II	28.6	48.9
FRAAL3325	IV	85.3	12.8
***Frankia* sp. EuI1c**			
FraEuI1c_5666	II	NA	NA
FraEuI1c_7150	II	NA	NA
***Frankia* sp. CcI3**			
Francci3_0194	I	88.8	NA
Francci3_4011	I	41.6	NA
Francci3_1102	II	83.6	NA
Francci3_2728	II	83.2	NA
Francci3_2953	II	77.1	NA
Francci3_0156	IV	65.0	NA
Francci3_1123	IV	73.7	NA
Francci3_1132	IV	96.0	NA
Francci3_2839	IV	18.7	NA
Francci3_2840	IV	37.3	NA
***Frankia* sp. EAN1pec**			
Franean1_5336	II	NA	NA
Franean1_5421	II	NA	NA
Franean1_6769	II	NA	NA
Franean1_2349	IV	NA	NA
Franean1_7063	IV	NA	NA
***Candidatus* Frankia datiscae Dg1**			
FsymDg_3166	I	NA	43.0
FsymDg_0581	II	NA	57.4

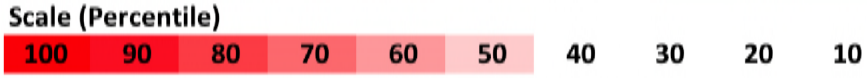

Figure 1 Restriction enzymes annotated in the *Frankia* genomes and expression levels of each gene in the transcriptome in culture when available (Alloisio *et al.*, 2010 and Bickhart and Benson, 2011). NA: Transcriptome not available.

Organism	Phylum	Gene ID	Percentile (Culture)
***Frankia alni* ACN14a**	Actinobacteria	FRAAL3325	**85.3**
***Frankia* sp. CcI3**	Actinobacteria	FranCcI3_1132	**96.1**
		FranCcI3_1123	**73.7**
		FranCcI3_0156	**65**
***Mycobacterium smegmatis* MC2 155**	Actinobacteria	MSMEG_1765	**79.8**
***Mycobacterium tuberculosis* H37Rv**	Actinobacteria	Rv2528c	**5.7**
***Streptomyces avermitilis* MA-4680**	Actinobacteria	SAV_5202	**77.4**
***Rhodococcus jostii* RHA1**	Actinobacteria	RHA1_ro05906	**79.9**
***Escherichia coli* K-12 MG1655**	Proteobacteria	ECK4341	**63.1**
***Bacillus anthracis* Sterne**	Firmicutes	BAS2162	**8.8**
		BAS2317	**50.7**

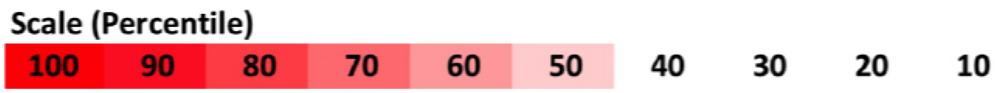

Figure 2 Restriction enzymes annotated in the genomes of *Frankia*, other actinobacteria, and proteobacteria and firmicutes for comparison. For each, expression levels for available transcriptomes growing in pure culture are provided.

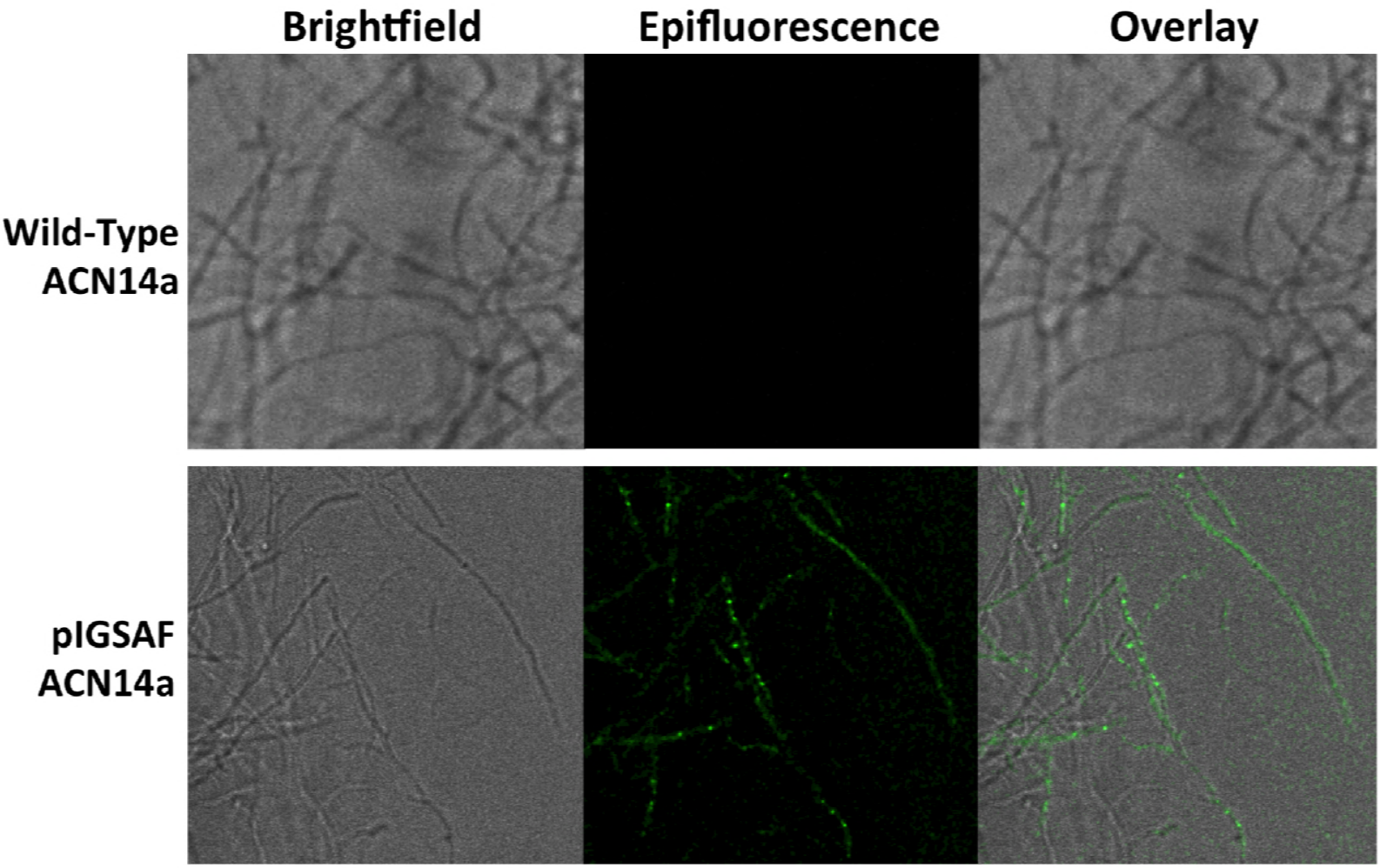

Figure 3 Transformation of *Frankia alni* ACN14a with a plasmid expressing *egfp*. Top: Wild-type *F. alni* ACN14a control. Bottom: Transformed *F. alni* ACN14a with plasmid pIGSAF. Images were obtained on a confocal microscope with both brightfield and epifluorescence and then overlaid.

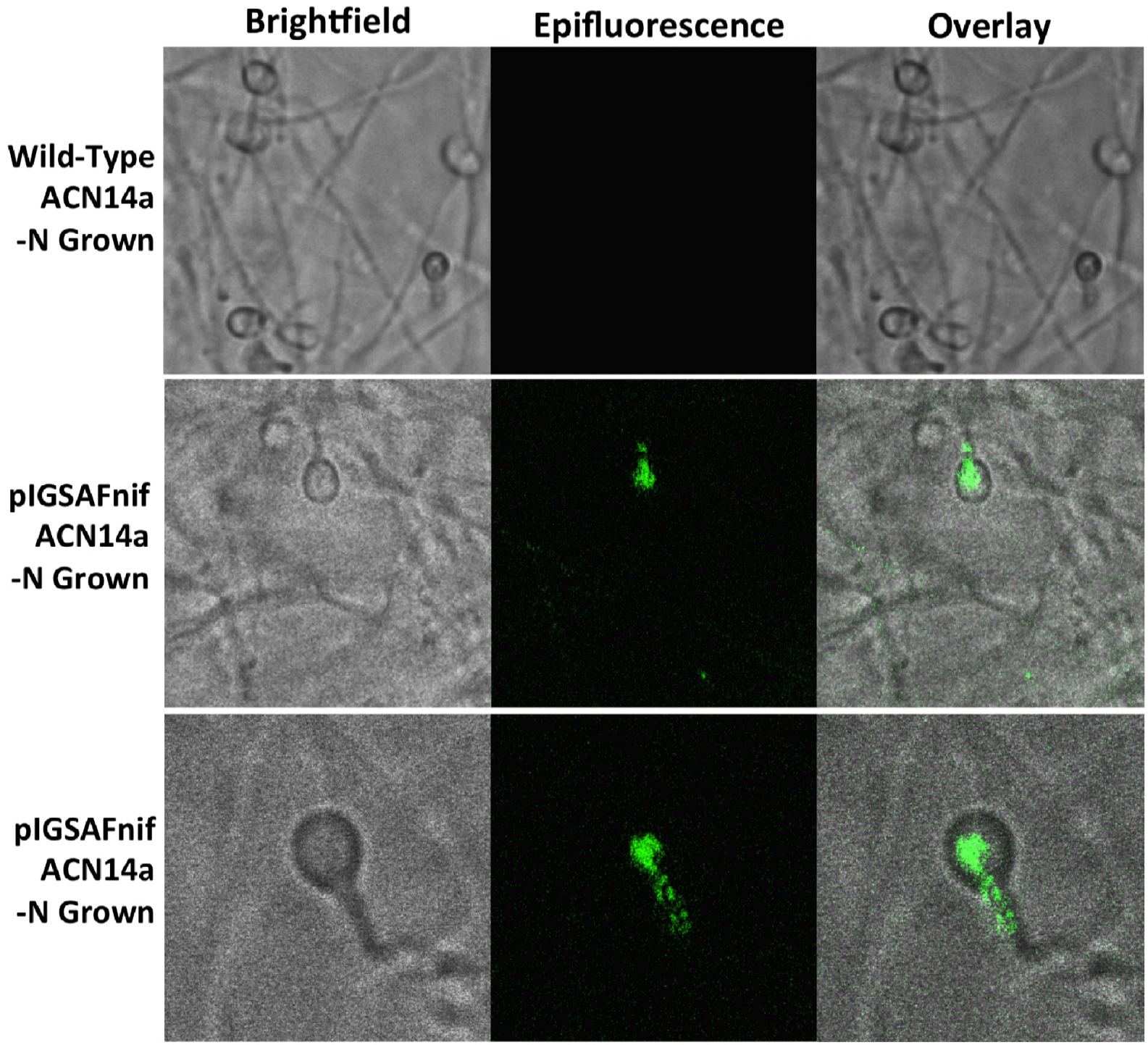

Figure 4 Transformation of *Frankia alni* ACN14a with a plasmid expressing *egfp* under the control of the *nif* cluster promoter region from the *F. alni* genome. Transformed cultures were grown in both (+)N and (–)N media to compare fluorescence with and without nitrogen fixation. Top: *F. alni* ACN14a transformed with plasmid pIGSAFnif grown in (+)N media. Middle: *F. alni* ACN14a with plasmid pIGSAFnif grown in (–)N media. Bottom: Close-up of a vesicle showing fluorescence in the stalk. Images were obtained on a confocal microscope with both brightfield and epifluorescence and then overlaid.

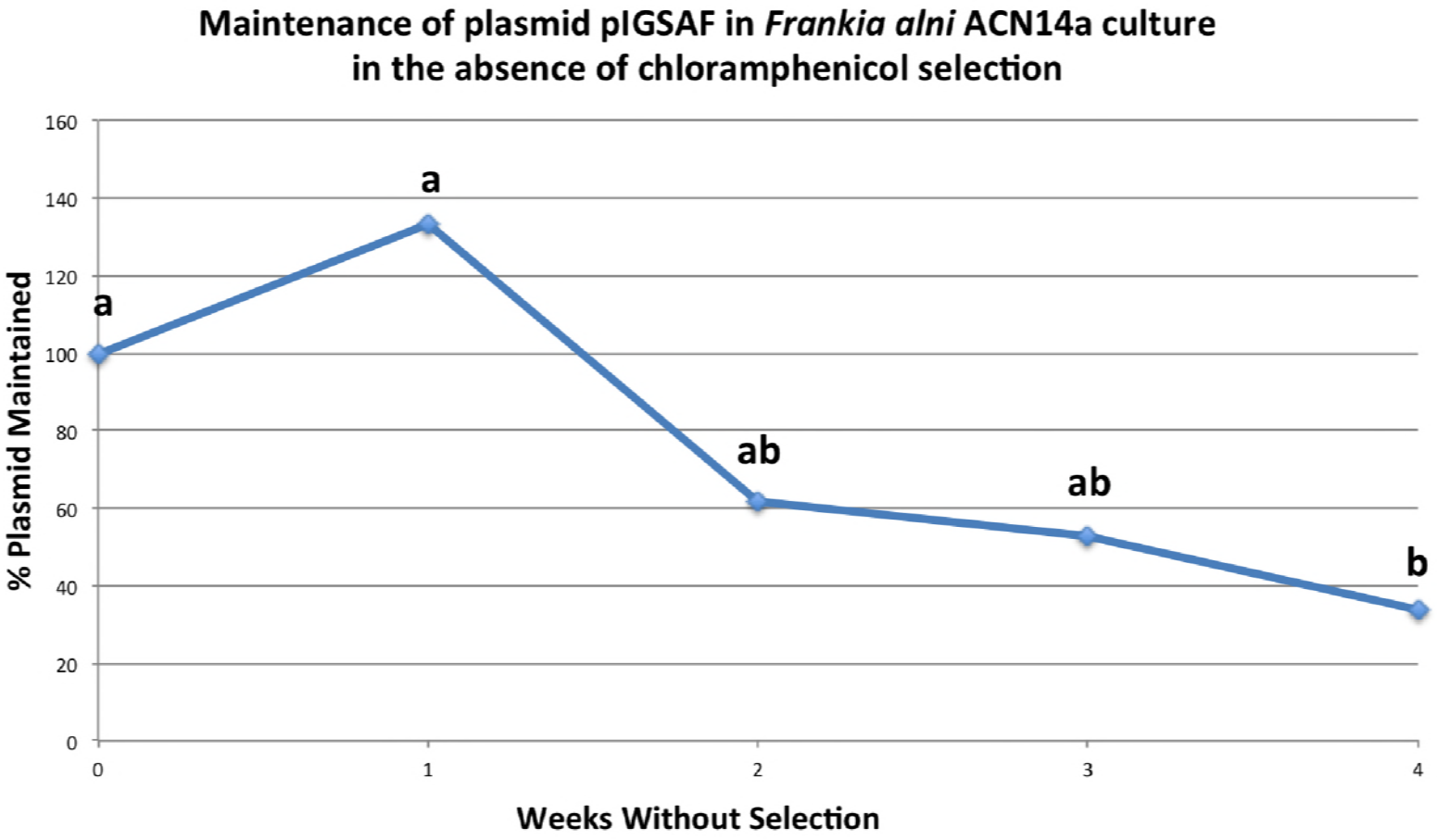

Figure 5 Maintenance of plasmid pIGSAF in *Frankia alni* in culture without chloramphenicol selection over the course of four weeks. Each week the culture was sub-cultured into fresh media and genomic DNA was extracted to measure relative plasmid concentrations via qPCR. Groups 'a' and 'b' indicate time points that are not significantly different from each other ($p<.05$).

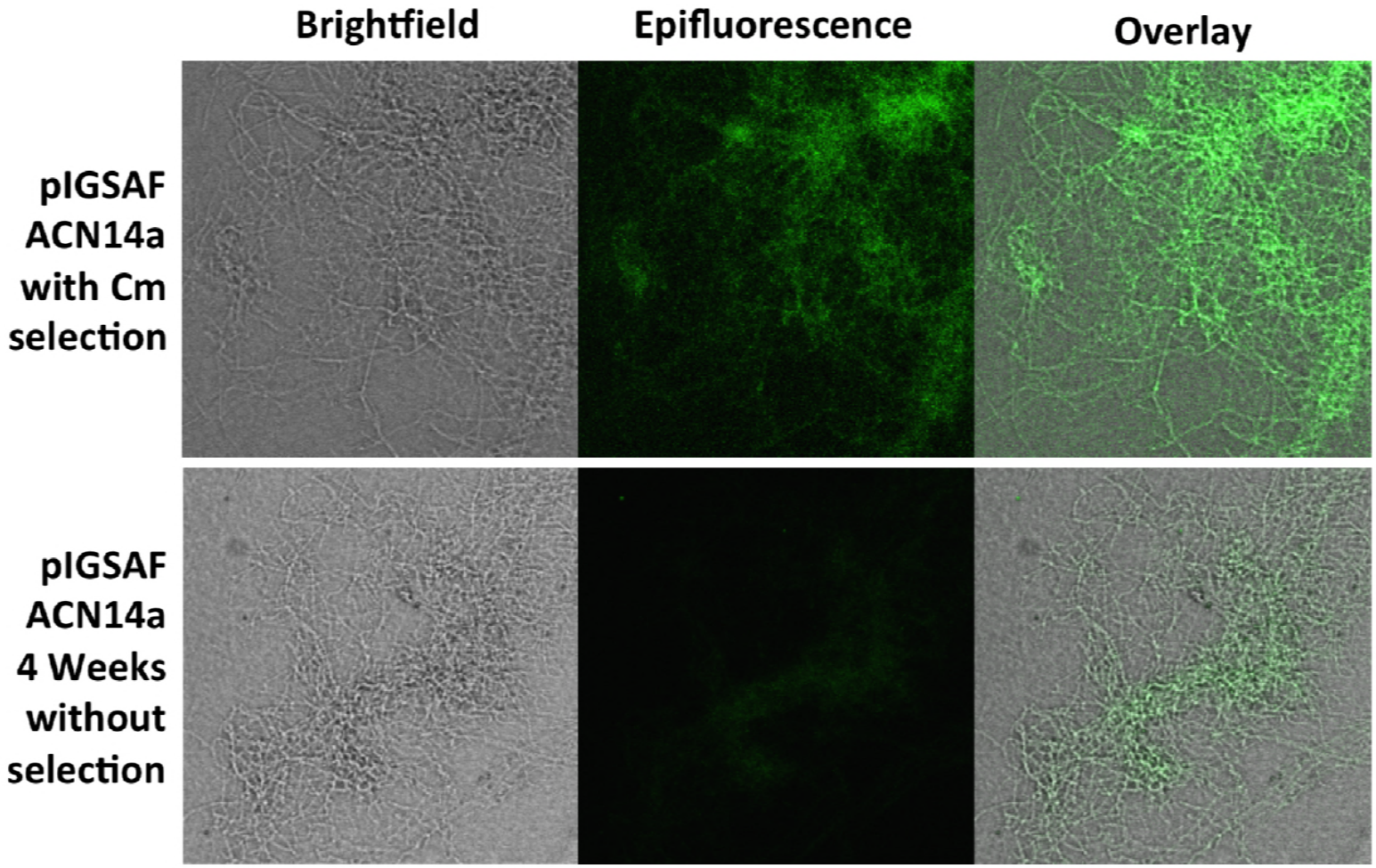

Figure 6 Fluorescence of a *Frankia alni* hyphal colony transformed with plasmid pIGSAF grown with chloramphicol selection (top) and without selection for four weeks (bottom).

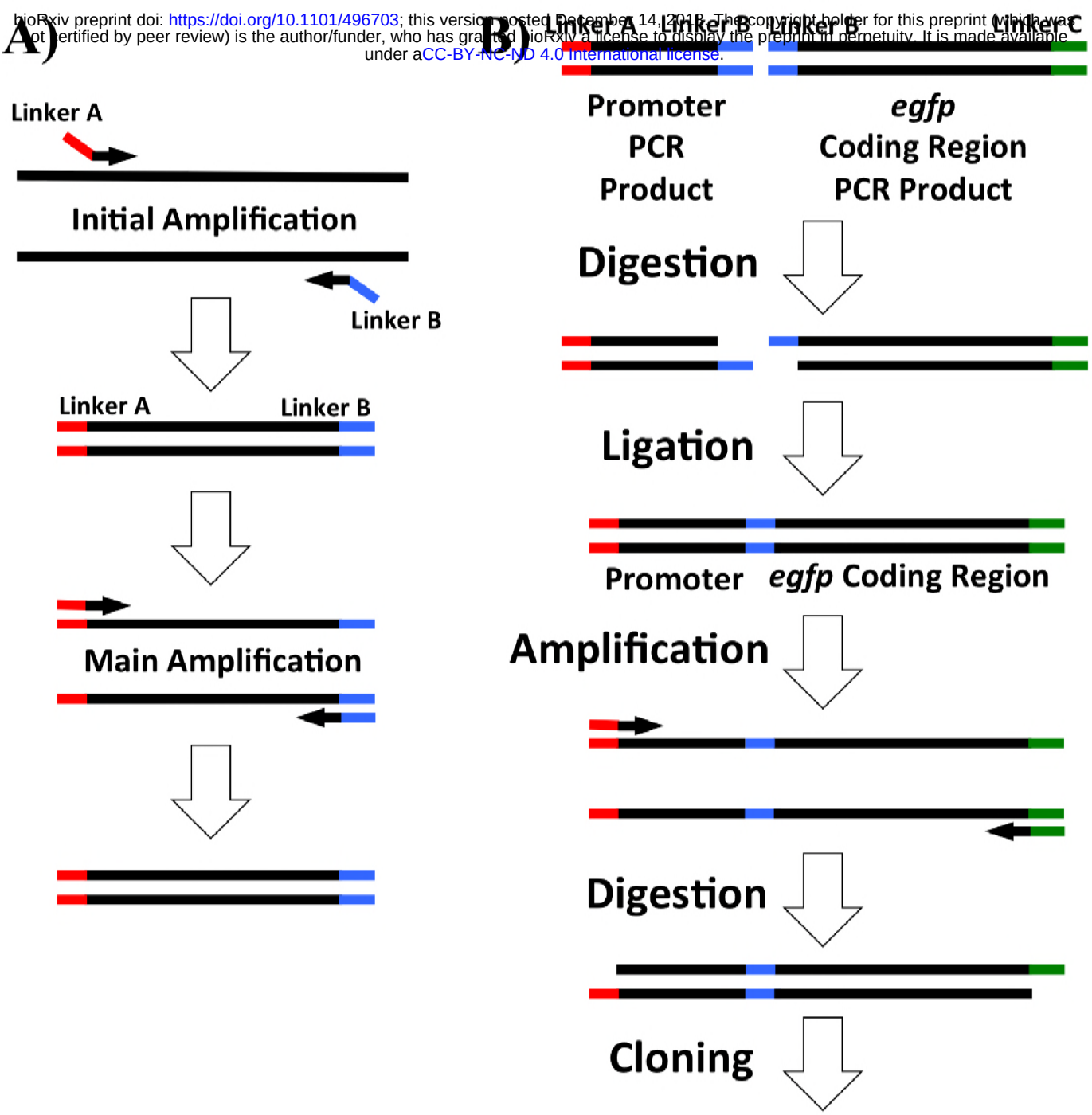

Figure 7 PCR amplification and digestion and ligation overview for plasmid synthesis used in this study. Linkers including restriction sites are colored red, blue, or green. A) PCR products for cloning were amplified by amplification at a lower initial temperature to allow the incorporation of additional restriction sites on the 5' end of the primers (shown in red and blue). The annealing temperature was then increased to amplify full-length products with the added restriction sites. B) Addition of the *F. alni* ACN14a *nif* cluster promoter region to the *egfp* coding sequence. Both the promoter region and *egfp* coding sequence were amplified with EcoRI sites (blue). These were then digested and ligated together followed by amplification of the ligation product. Restriction sites on the ends (red and green) were digested, allowing incorporation of the *nif* promoter:*egfp* product into the digested plasmid.

Gene	Role	Fold Change (-N/+N)	p-value
nif:egfp	Fluorescence	101.6	<.05
nifH	Nitrogen fixation	8.5	<.05
rpoD	Housekeeping	0.4	>.05

Primer Name	Linker Restriction Site	Sequence	Melting Temperatures (°C)	
			Full Primer	Without Linker
pSA3_Cm_F	-	TATTCAGGCGTAGCACCAGG	60.5	-
pSA3_Cm_R	-	TGTTGATACCGGGAAGCCCT	60.5	-
EGFP_SalI_F	SalI	TAA**GTCGAC**GAGCCGAAGCATAAACAGCG	71.9	60.5
EGFP_BamHI_R	BamHI	ATTAA**GGATCC**GTACCGGCATAACCAAGCCT	72.1	60.5
nif_promoter_XbaI_F	XbaI	AAGTT**TCTAGA**GTGCTCCTATTCGTTCGGC	70.8	59.5
nif_promoter_EcoRI_R	EcoRI	AAGAA**GAATTC**TTGCTCCGGGACTGAAGAC	70.8	59.5
GFP_CDS_EcoRI_F	EcoRI	AAGCT**GAATTC**ATGAGTAAAGGAGAAGAAC	66.7	50.9
GFP_CDS_SalI_R	SalI	AATTC**GTCGAC**TGCCTGACTGCGTTAGCAA	72.1	57.5
rpoD_qPCR_F	-	ATGCTGTTCCTGGACCTCATC	61.2	-
rpoD_qPCR_R	-	GTGGCGTAGGTCGAGAACTT	60.5	-
nifH_qPCR_F	-	GCGTACTTCAGGATGCCTCG	62.5	-
nifH_qPCR_R	-	GACGTTGTGTGTGGTGGGTT	60.5	-
infC_qPCR_F	-	ACGACGTGACCCTTCTTGGT	60.5	-
infC_qPCR_R	-	TCGGGAAGCTCGGAAGAAC	59.5	-
gfp_qPCR_F	-	TGCTTTGCGAGATACCCAGA	58.4	-
gfp_qPCR_R	-	ACGTGTCTTGTAGTTCCCGTC	61.2	-