
What stories can the *Frankia* genomes start to tell us?

LOUIS S TISA^{1,*}, NICHOLAS BEAUCHEMIN¹, MAHER GTARI², ARNAB SEN³ and LUIS G WALL⁴

¹Department of Molecular, Cellular, and Biomedical Sciences, University of New Hampshire, Durham, NH, USA

²Microorganisms and Active Biomolecules Laboratory, University of Tunis-El Manar, Tunisia

³Bioinformatics Facility, Department of Botany, University of North Bengal, Siliguri 734 013, India

⁴Department of Science and Technology, University of Quilmes, Bernal, Argentina

*Corresponding author (Email, louis.tisa@unh.edu)

Among the Actinobacteria, the genus *Frankia* is well known for its facultative lifestyle as a plant symbiont of dicotyledonous plants and as a free-living soil dweller. *Frankia* sp. strains are generally classified into one of four major phylogenetic groups that have distinctive plant host ranges. Our understanding of these bacteria has been greatly facilitated by the availability of the first three complete genome sequences, which suggested a correlation between genome size and plant host range. Since that first report, eight more *Frankia* genomes have been sequenced. Representatives from all four lineages have been sequenced to provide vital baseline information for genomic approaches toward understanding these novel bacteria. An overview of the *Frankia* genomes will be presented to stimulate discussion on the potential of these organisms and a greater understanding of their physiology and evolution.

[Tisa LS, Beauchemin N, Gtari M, Sen A and Wall LG 2013 What stories can the *Frankia* genomes start to tell us? *J. Biosci.* **38** 719–726] DOI 10.1007/s12038-013-9364-1

1. Introduction

Members of the genus *Frankia* are well known for their facultative lifestyle as plant symbionts and as free-living soil-dwelling actinobacteria (Benson and Silvester 1993; Chaia *et al.* 2010; Schwenecke and Carú 2001; Wall 2000). The symbiosis allows actinorhizal plants to colonize harsh environmental terrains under diverse ecological conditions. Based on several criteria including 16S rRNA gene (Normand *et al.* 1996), *glnII* (Cournoyer and Lavire 1999; Nouioui *et al.* 2011), *gyrB* (Nouioui *et al.* 2011) and 16S-23S rRNA intertranscribed spacer region (Ghodhbane-Gtari *et al.* 2010), phylogenetic analysis has identified four distinct clusters among the *Frankia* strains.

Elucidation of the first three *Frankia* genomes has revealed new potential in respect to metabolic diversity, natural product biosynthesis and stress tolerance, which may help aid the cosmopolitan nature of the actinorhizal symbiosis (Normand *et al.* 2007a, b; Udway *et al.* 2011). Genomes for representatives from each of the four clusters have been sequenced (Normand *et al.* 2007a; Persson *et al.* 2011; Sen *et al.* 2013, Ghodhbane-Gtari *et al.* 2013; Beauchemin *et al.* 2012; Wall *et al.* 2013; Nouioui

et al. 2013) and have provided vital baseline information for genomic approaches toward understanding these novel bacteria. At present, genome sequences are available for all four *Frankia* lineages: Strains CcI3, ACN14a, and QA3 represent Cluster 1 phylogeny (Normand *et al.* 2007a; Sen *et al.* 2013), while strains EAN1pec, EUN1f, BMG5.12, and BCU110501 provide genome databases for Cluster 3 phylogeny (Normand *et al.* 2007a; Beauchemin *et al.* 2012; Nouioui *et al.* 2013; Wall *et al.* 2013) Cluster 2 phylogeny is represented by a genome for the uncultured *Frankia* symbiont of *Datisca glomerata* (Persson *et al.* 2011). Furthermore, *Frankia* strain CN3, EuI1c and DC12 genomes represent Cluster 4 phylogeny or atypical *Frankia* (Beauchemin *et al.* 2012; Ghodhbane-Gtari *et al.* 2013; Tisa *et al.* unpublished). Atypical *Frankia* strains were isolated from actinorhizal nodules, but are unable to infect and re-establish effective nodulation on their host plant. *Frankia* sp. strain EuI1c is able to re-infect its host plant, but establishes an ineffective nodulation (Baker *et al.* 1980) The genome databases are providing a wealth of information and have been used in genome mining (Niemann and Tisa 2008; Perrine-Walker *et al.* 2010), comparative genomics (Normand *et al.* 2007a, b; Mastrorunzio *et al.*

Keywords. Actinobacteria; actinorhizal symbiosis; genomes; nitrogen fixation

2008; Sen et al. 2008; Bickhart et al. 2009; Udwary et al. 2011; Sen et al. 2012), transcriptomics (Alloisio et al. 2010; Popovici et al. 2010, 2011; Bickhart and Benson 2011) and proteomics approaches (Alloisio et al. 2007; Bagnarol et al. 2007; Mastronunzio et al. 2009; Mastronunzio and Benson 2010).

With the availability of 11 *Frankia* genome databases, an analysis of the basic characteristics and traits of the *Frankia* genome is possible. These genomes were sequenced to provide background information on the various cluster groups and to dissect genomic characteristics of *Frankia* sp. from diverse host plants and environments. An increased number of *Frankia* genomes also allows for the validation of the genome size and host range hypothesis. The goal of this study was to provide some insight on an overview of these databases. General characteristics of the genomes were examined and analysed at the broadest level. The gene details were not addressed in this study.

2. Materials and methods

2.1 Bioinformatics analysis

Sequences of the *Frankia* genomes were obtained from Integrated Microbial Genomes System from the Joint Genome Institute (www.img.doe.gov) (Markowitz et al. 2006). Comparative analysis of the *Frankia* genomes was performed using the IMG system (Markowitz et al. 2006). Putative horizontally transferred genes (HGT) were defined as genes that have the best hits to genes that are not in the phylogenetic group of the query sequence. To determine a HGT, not only was the best hit (highest bit score) used, but all of the hits that have bit score equal or greater than 95% of the best hit. Pseudogenes were identified from those characterized for each genome. Insert sequences and transposases were identified from search of the genome databases.

BlastP search were performed on the embedded IGM site and used the 70% and 50% identity levels. The query genome was BlastP searched against representative genome databases as described in the text. The algorithm was set for an absence or presence of homologs. The minimum taxon percent with homologs and without homologs were both set at 100%. Pseudogenes were not excluded.

Dot plots were generated via Mummer (Kurtz et al. 2004) on the IMG site. The DNA sequences for two genomes were chosen and the nucmer (NUCLEotide MUMmer) pipeline was used to generate the plots.

3. Results

3.1 Does the size relationship hold true?

Based on the first three *Frankia* genome sequences, a hypothesis was presented that genome size was related to host

specificity and biogeography ranges (Normand et al. 2007a). These three *Frankia* genomes represent two of the four major lineages within the genus. The most striking difference among the three genomes is their sizes, which varied from 5.4 Mb for a narrow-host-range *Casuarina* strain (*Frankia* sp. strain HFPCcI3) to 8.9 Mb for soil cosmopolitan and broad-host-range *Elaeagnus* strain (*Frankia* sp. strain EAN1pec). The smallest genome belongs to the narrow host range and geographically limited representative strain CcI3, and the largest is from strain EAN1pec that belongs to the broadest host range group. *Frankia alni* strain ACN14a is an intermediate host range being limited to *Alnus* and *Myricaceae* host infectivity and mostly restricted to alder stands. In recent years, several more *Frankia* strains have been sequenced (table 1), and analysis of the size of these genomes, in general, confirms the hypothesis. All of the broad-host-range Cluster 3 strains EUN1f, EAN1pec, BMG5.12 and BCU110501 had large genomes at 9.1, 8.9, 7.6 and 7.9 Mb, respectively. The Cluster 1 medium-host-range strains ACN14a and QA3 each were 7.5 Mb, while the narrow-host-range strain CcI3 had a 5.4 Mb sized genome. The Cluster 2 obligate symbiont strain Dg1 also had the smallest genome size (5.3 MB). Cluster 4 genomes varied in size with strains CN3, EuI1c, and DC12 at 9.9, 8.8 and 6.7 Mb, respectively.

3.2 What about genome plasticity?

The *Frankia* genomes were analyzed for their distribution of horizontal gene transfer (HGT), IS elements and pseudogenes content (figure 1). The percentages of these mobile elements are shown in figure 1. Among the *Frankia* genomes, HGT genes content was divided between two groups: Group 1 consisting of strains ACN14a, CcI3, QA3, EAN1pec, DC12 and CN3 that had 1.5 to 4.5 % of the gene content. Group 2 consisting of strains EUN1f, EuI1c Dg1, BMG5.12 and BCU110501 that contained 7, 3, 1, 0 and 0 HGT genes, respectively. The absence of HGT in the BMG5.12 and BCU110501 genomes may be reflection of the draft status of the genome. The numbers of IS elements found among these genomes ranged from 0.5 % for ACN14a to 4.0% for Dg1 except for the EuI1c genome that had only 10 IS elements (0.1%). This result would indicate that *Frankia* sp. strain EuI1c genome was stable. Analysis of the pseudogene content of these genomes indicates that the level of these genes ranged from 1 to 7%. The exceptions were the ACN14a, EAN1pec, BMG5.12 and BCU110501 genomes, which had 12, 1, 0 and 0 pseudogenes, respectively. The nature of the minimum draft sequence data may explain the BMG5.12 and BCU110501 results.

3.3 What is the composition of the *Frankia* core genome?

The *Frankia* core genome was identified by BlastP search of the protein content of the *Frankia* genomes (table 2). The

Table 1. Properties of *Frankia* genomes

<i>Frankia</i> strain	Clade	Genome size	GC%	No. of genes	rRNA genes	tRNA genes	Accession number	Reference
ACN14a	I	7,497,934	72.83	6795	6	46	CT573213	Normand <i>et al.</i> 2007a
CcI3	I	5,433,628	70.08	4621	6	46	CP000249.1	Normand <i>et al.</i> 2007a
QA3	I	7,590,853	72.59	6546	4	46	CM001489.1	Sen <i>et al.</i> 2013
Dg1	II	5,323,336	70.04	4579	6	44	CP002801	Persson <i>et al.</i> 2011
EAN1pec	III	8,982,042	71.15	7250	9	47	CP000820.1	Normand <i>et al.</i> 2007a; Beauchemin <i>et al.</i> 2012
EUN1f	III	9,322,173	70.82	7833	9	47	NC014666	Beauchemin <i>et al.</i> 2012
BMG5.12	III	7,589,313	71.67	6342	5	51	ARFH0100000	Nouioui <i>et al.</i> 2013
BCU110501	III	7,891,711	72.39	6839	8	47	ARDT0100000	Wall <i>et al.</i> 2013
EuI1c	IV	8,815,781	72.31	7262	9	46	CP002299.1	Beauchemin <i>et al.</i> 2012
CN3	IV	9,978,592	71.72	8412	5	68	AGJN0100000	Ghodhbane-Gtari <i>et al.</i> 2013
DC12	IV	6,884,336	71.92	5858	9	46		Tisa, unpublished

presence of shared common genes was determined at the high stringency level (70% identity) and at the lower threshold level (50% identity). With these parameters, 688 and 1442 common core genes among all 11 *Frankia* genomes were identified at the high and low stringency levels, respectively. The core genomes among the nitrogen-fixing strains were 807 and 1645 genes at the high and low stringency levels, respectively. The higher numbers reflect a representative picture of the nitrogen-fixing symbionts. The profile of individual *Frankia* clusters were examined and Cluster 3 including Cluster 4 strain EuI1c had 1117 and 747 genes at the high and low stringency levels, respectively. If strain EuI1c genome, as an infective and non-nitrogen-fixing strain, is removed and placed in Cluster 4, these values drop to 45 and 74 genes at high and low stringency, respectively. Most of these genes consisted of *nif* genes. Because of the ability of strain EuI1c to re-infect its host plant, the exception of this genome from the Cluster 4 was preferred to detect infectivity genes.

The *Frankia* genomes were analyzed for strain-specific unique genes (table 3). At a 70% identity high stringency level, the number of unique strain-specific genes ranged from 1359 for strain BCU110501 to 3401 for the atypical strain CN3. At a lower stringent level (50% identity), these values decreased and ranged from 757 for strain CcI3 to 1605 for strain CN3. In both cases, the largest sized genome (Strain CN3) contained the most unique strain-specific genes.

3.4 What about major shifts and realignments in the *Frankia* genomes?

To understand the presence of major shifts and realignments, the genomes were analysed by the use of DNA-DNA dot plots. Analysis of the Cluster 1 genomes indicates that the genomes aligned well, and the major area of inversions and

deletions was at the terminus region (figure 2). With the Cluster 3 genomes, a similar pattern occurs, but with greater variability along the length of the genome (figure 3). The pattern of strains EUN1f and EAN1pec shows a greater numbers of gaps (figure 3) than the patterns with strains ACN14a and QA3 and CcI3 and ACN14a (Fig 2), which are focused at the terminus regions. The draft genomes of BCU110501 (figure 3) and BMG5.12 (data not shown) aligned very well with the EAN1pec genome. The EuI1c genome which is a Cluster 4 strain aligns well with both EUN1f and EAN1pec. This strain is infective toward *Eleaegmus* host plant, but forms ineffective nodules. The Cluster 4 genomes have major gaps when aligned together (figure 4). These patterns show major rearrangements between EuI1c genome and DC12 or CN3 genomes. The Cluster 2 genome (Dg1) does not align very well with any genomes of the other three Clusters (figure 5), while Clusters 1 and 3 genomes aligned well except in the terminus region (figure 5).

4. Discussion

A correlation between host-plant range and genome size was seen before on 3 genomes (Normand *et al.* 2007a). In general, our study of 11 genomes confirms that observation and hypothesis. However, the size variability of the promiscuous Cluster 3 genomes was large, ranging from 7.6 to 9.3 Mb. Furthermore, the atypical Cluster 4 genomes showed the greatest range 6.8 to 9.9 Mb, which may reflect the diversity of this Cluster. The small-sized genomes of the host-restricted *Casuarina* isolate CcI3 and the Cluster 2 uncultured isolate Dg1 indicates that their genomes are reduced. Confirmation of this hypothesis requires more genomes for these two groups. Our current efforts have focused on sequencing five *Casuarina* isolates and preliminary data indicate that these genomes are also reduced in size, ranging from 5.4 to 5.6 Mb (Tisa, unpublished). More genomes

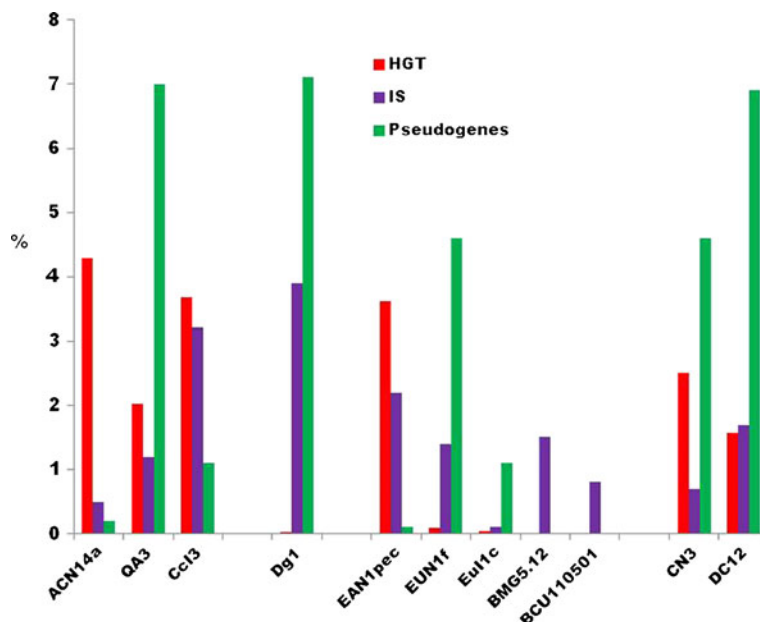


Figure 1. Identification of the number of IS elements, HGT and pseudogenes among the *Frankia* genomes. Numbers of IS elements, HGT and pseudogenes are expressed as a percentage of the number of genes on the genome.

sequences from Clusters 2 and 4 are required to confirm the Cluster 2 finding and tease out information on the Cluster 4 group.

The presence of large numbers of HGT and IS elements in some of the *Frankia* genome is an indication of the genome plasticity. However, the *Frankia* sp. strain EuI1c genome appears to be the exception and contained only 10 IS elements and 3 HGTs. These results would indicate this genome was the most stable. Although this strain groups within Cluster 4 (Nouioui *et al.* 2011), it is infective on its host plant and establishes an ineffective nodule structure (Baker *et al.* 1980). The absence of any nitrogen fixation genes within the genome is mostly the cause of its ineffective nodulation (Beachemin *et al.* 2012). The large genome size of *Frankia* sp. strain EuI1c allows this strain to live a cosmopolitan lifestyle within the host plant

and outside in the soil environment. This genome may provide insight on the genes needed to infect a host plant and should serve as a model system.

An analysis of the DNA-DNA dot plots provides a picture of the major alignments and shifts among the *Frankia* genomes. The Cluster1 genomes aligned very well with an inversion of the terminus region in strain QA3 compared to *F. alni* ACN14a. The terminus of the CcI3 genome was where most deletions occurred. With Cluster 3 genomes,

Table 2. *Frankia* core genomes based on gene lineage

Condition	70% Identity	50% Identity
11 Genomes	688	1442
All N ₂ Fixing (–CN3, DC12, EuI1c)	807	1645
Cluster III (not in CN3 or DC12)	1117	747
Cluster I (not in CN3 or DC12)	902	542
Cluster III (not in EuI1c)	1283	982
Cluster III (not in cluster I)	218	427
Cluster I (not in cluster III)	525	232

Homologs were identified by BlastP analysis with a cut-off level of 70% and 50% protein identity.

Table 3. Identification of strain-specific unique genes among the *Frankia* genomes

<i>Frankia</i> strain	70% Identity	50% Identity
ACN14a	1953	1433
CcI3	1568	757
QA3	1523	912
Dg1	2712	1471
EAN1pec	1486	844
EUN1f	2930	1496
EuI1c	1875	1018
BMG5,12	1765	813
BCU110501	1359	839
CN3	3401	1605
DC12	1373	885

The presence of a unique strain-specific gene was determined by the use of BlastP compared to all *Frankia* genomes and with a 70% and 50% protein identity cut-off values.

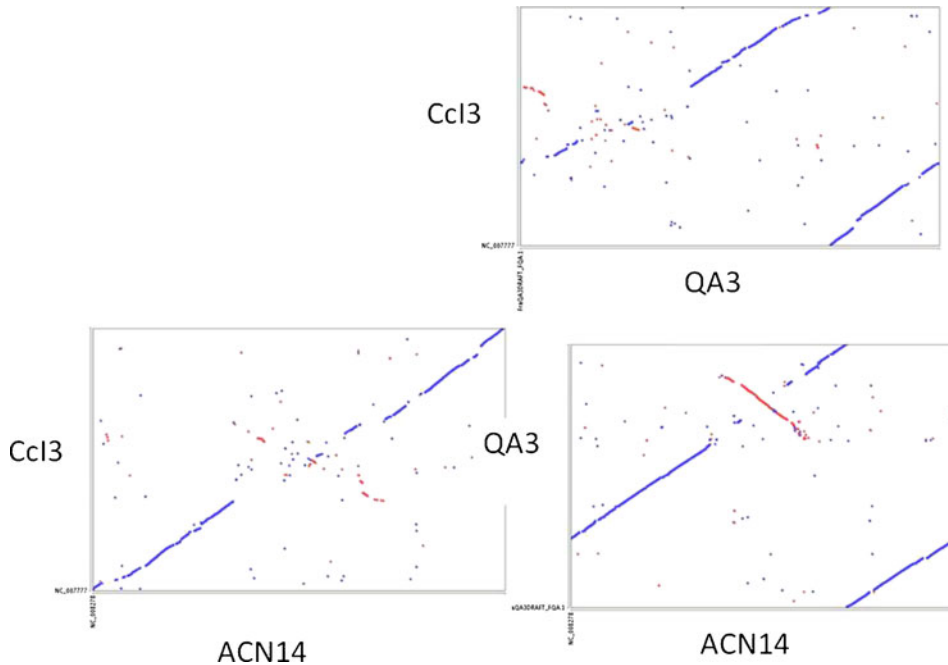


Figure 2. DNA-DNA dot plot of *Frankia* genomes from Cluster 1.

the aligned genomes were similar, but showed more gaps. The terminus region had the most gaps, but many occurred in other regions of the genome. This result would indicate that the genomes of the Cluster 3 isolates were more fluid.

Not surprisingly, the strain EuI1c genome aligned well with the Cluster 3 genomes. In fact, it aligned better with Cluster 3 genomes than the two Cluster 4 genomes. The Cluster 4 genomes had diverse alignment patterns compared to the Cluster 1 and 3

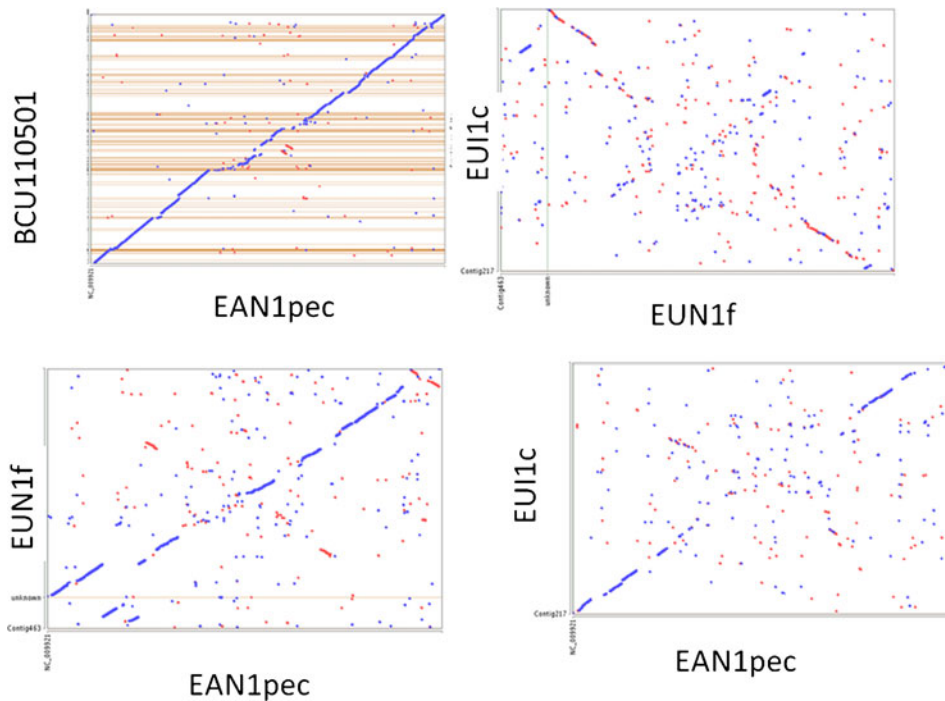


Figure 3. DNA-DNA dot plot of *Frankia* genomes from Cluster 3.

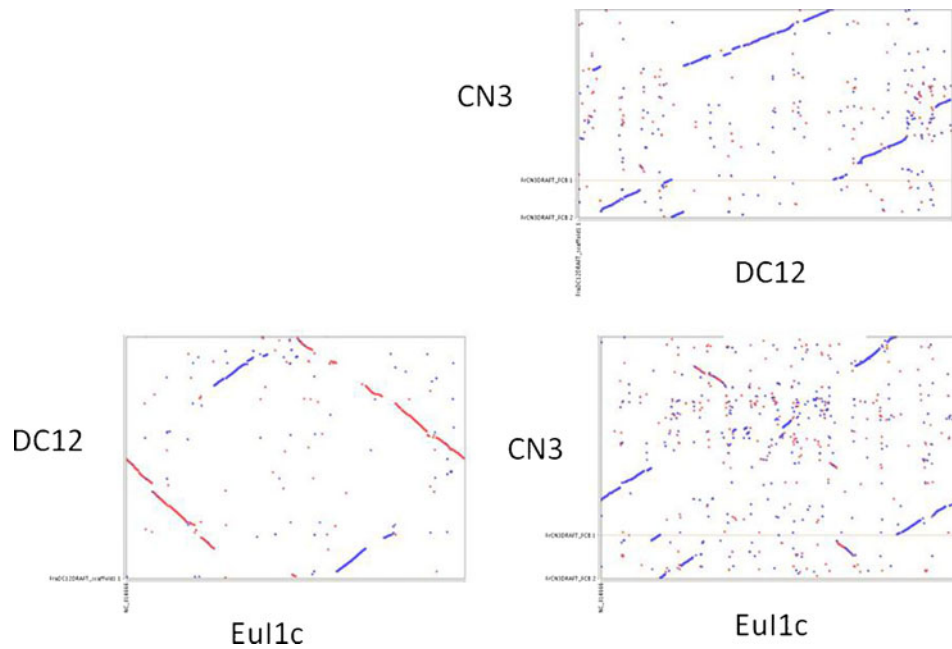


Figure 4. DNA-DNA dot plot of *Frankia* genomes from Cluster 4.

genome patterns. These results would suggest that the Cluster 4 group is much more diverse and undergoing major changes.

The Cluster 2 Dg1 genome did not align well with either the Cluster 1 or 3 genomes. However, Cluster 1 and 3 genomes aligned well with the major changes in the terminus region. This

observation would suggest that the Cluster 2 genomes are very different from the genomes of the other Clusters. The availability of more Cluster 2 genomes would help clarify that situation.

The absence of *Frankia* genetics including gene inactivation and/or complementation methods has been a major

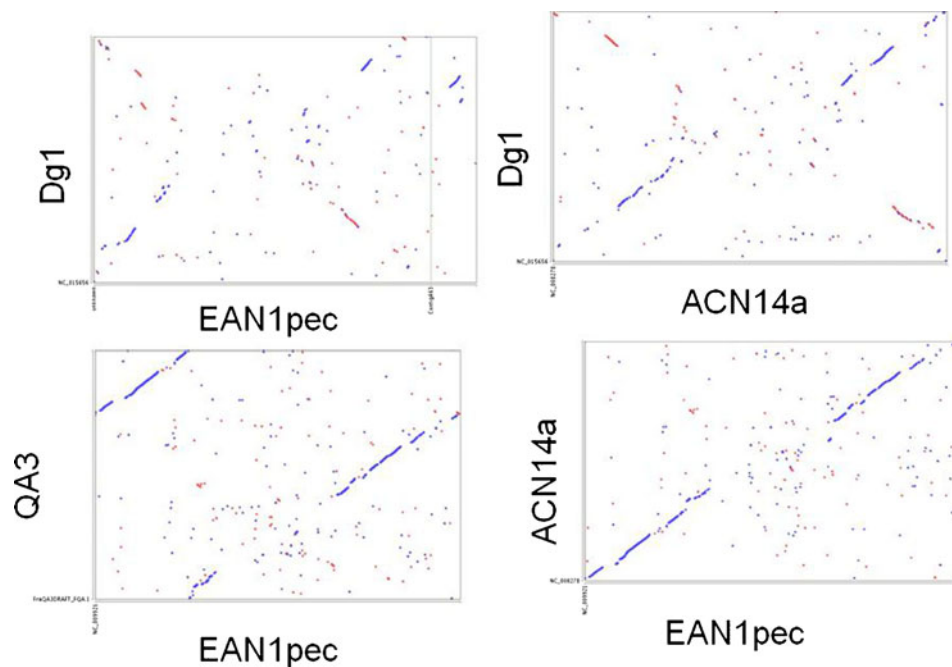


Figure 5. DNA-DNA dot plot of *Frankia* genomes from across clusters.

hindrance in our progress on the biology of this actinobacterium and the genetic determinants of plant root invasion and symbiotic establishment. The availability of these *Frankia* genomes greatly improved our understanding and has provided wealth of information on the facultative plant symbiotic and nitrogen-fixing actinobacteria. For example, the absence of common rhizobial *nod* genes involved cell signaling in the legume symbiosis suggests that the actinorhizal symbiosis uses novel signal compounds during the infection process. Several cores of *Frankia* genes are identifiable that may provide tools to further access *Frankia* diversity including *Frankia*-unique genes that may serve for host specificity and infectivity, and the adaptation to diverse soil conditions and taxonomically disparate host plants.

Acknowledgements

This project (LST) was supported in part by Agriculture and Food Research Initiative Grant 2010–65108–20581 from the USDA National Institute of Food and Agriculture, Hatch Grant NH530, and The College of Life Sciences and Agriculture at the University of New Hampshire, Durham, NH. MG was supported in part by a Visiting Scientist Program administered by the NH AES at the University of New Hampshire.

References

- Alloisio N, Felix S, Marechal J, Pujic P, Rouy Z, Vallenet D, Medigue C and Normand P 2007 *Frankia alni* proteome under nitrogen-fixing and nitrogen-replete conditions. *Physiol. Plantarum* **130** 440–453
- Alloisio N, Queiroux C, Fournier P, Pujic P, Normand P, Vallenet D, Medigue C, *et al.* 2010 The *Frankia alni* Symbiotic Transcriptome. *Mol. Plant Microbe In.* **23** 593–607
- Bagnarol E, Popovici J, Alloisio N, Marechal J, Pujic P, Normand P, and Fernandez MP 2007 Differential *Frankia* protein patterns induced by phenolic extracts from *Myricaceae* seeds. *Physiol. Plantarum* **130** 380–390
- Baker D, Newcomb W, and Torrey JG 1980 Characterization of an ineffective actinorhizal microsymbiont, *Frankia* sp. Eu11c (Actinomycetales). *Can. J. Microbiol.* **26** 1072–1089
- Beauchemin N, Gtari M, Ghodhbane-Gtari F, Furnholm T, Sen A, Wall L, Tavares F, *et al.* 2012 What can the genome of an infective ineffective (Fix-) *Frankia*. Strain (Eu11c) that is able to form nodules with its host plant tell us about actinorhizal symbiosis and *Frankia* evolution. The 112th General Meeting of the American Society for Microbiology American Society for Microbiology, San Francisco, CA
- Benson DR and Silvester WB 1993 Biology of *Frankia* strains, actinomycete symbionts of actinorhizal plants. *Microbiol. Mol. Biol. Rev.* **57** 293–319
- Bickhart DM, and Benson DR 2011 Transcriptomes of *Frankia* sp strain CcI3 in growth transitions. *BMC Microbiol.* **11** 192
- Bickhart DM, Gogarten JP, Lapierre P, Tisa LS, Normand P and Benson DR 2009 Insertion sequence content reflects genome plasticity in strains of the root nodule actinobacterium *Frankia*. *BMC Genomics* **10** 468
- Chaia EE, Wall LG and Huss-Danell K 2010 Life in soil by the actinorhizal root nodule endophyte *Frankia*. A review. *Symbiosis* **51** 201–226
- Cournoyer B and Lavire C 1999 Analysis of *Frankia* evolution radiation using *glnII* sequences. *FEMS Microbiol. Lett.* **117** 29–34
- Ghodhbane-Gtari F, Beauchemin N, Bruce D, Chain P, Chen A, Walston Davenport K, Deshpande S, *et al.* 2013 Draft Genome sequence of *Frankia* sp. strain CN3, an atypical, non-infective (Nod-) ineffective (Fix-) isolate from *Coriaria nepalensis*. *Genome Announc.* **1** e00085-13
- Ghodhbane-Gtari F, Nouioui I, Chair M, Boudabous A and Gtari M 2010 16S-23S rRNA intergenic spacer region variability in the genus *Frankia*. *Microb. Ecol.* **60** 487–495
- Kurtz S, Phillippy A, Delcher AL, Smoot M, Shumway M, Antonescu C and Salzberg SL 2004 Versatile and open software for comparing large genomes. *Genome Biol.* **5** R12
- Markowitz VM, Korzeniewski F, Palaniappan K, Szeto E, Werner G, Padki A, Zhao XL, Dubchak I, *et al.* 2006 The integrated microbial genomes (IMG) system. *Nucleic Acids Res.* **34** D344–D348
- Mastrorunzio JE, Tisa LS, Normand P and Benson DR 2008 Comparative secretome analysis suggests low plant cell wall degrading capacity in *Frankia* symbionts. *BMC Genomics* **9** 47
- Mastrorunzio JE and Benson DR 2010 Wild nodules can be broken: proteomics of *Frankia* in field-collected root nodules. *Symbiosis* **50** 13–26
- Mastrorunzio JE, Huang Y and Benson DR 2009 Diminished exoproteome of *Frankia* spp. in culture and symbiosis. *Appl. Environ. Microbiol.* **75** 6721–6728
- Niemann J and Tisa LS 2008 Nitric Oxide and Oxygen Regulate Truncated Hemoglobin Gene Expression in *Frankia* Strain CcI3. *J. Bacteriol.* **190** 7864–7867
- Normand P, Lapierre P, Tisa LS, Gogarten JP, Alloisio N, Bagnarol E, Bassi CA, Berry AM, *et al.* 2007a Genome characteristics of facultatively symbiotic *Frankia* sp strains reflect host range and host plant biogeography. *Genome Res.* **17** 7–15
- Normand P, Orso S, Cournoyer B, Jeannin P, Chapelon C, Dawson JO, Evtushenko L and Mirsra AK 1996 Molecular phylogeny of the genus *Frankia* and related genera and emendation of the family *Frankiaceae*. *Int. J. Syst. Bacteriol.* **46** 1–9
- Normand P, Queiroux C, Tisa LS, Benson DR, Rouy Z, Cruveiller S and Medigue C 2007b Exploring the genomes of *Frankia*. *Physiol. Plantarum* **130** 331–343
- Nouioui I, Beauchemin N, Cantor MN, A. Chen A, Detter JC, Furnholm T, Ghodhbane-Gtari F, *et al.* 2013 Draft Genome sequence of *Frankia* sp. strain BMG5.12, a nitrogen-fixing actinobacterium isolated from Tunisian soils. *Genome Announc.* **1** e00468-13
- Nouioui I, Ghodhbane-Gtari F, Beauchemin NJ, Tisa LS, and Gtari M 2011 Phylogeny of members of the *Frankia* genus based on *gyrB*, *nifH* and *glnII* sequences. *Anton. van Leeuw.* **100** 579–587
- Perrine-Walker F, Doumas P, Lucas M, Vaissayre V, Beauchemin NJ, Band LR, Chopard J, *et al.* 2010 Auxin carriers localization drives

- auxin accumulation in plant cells infected by *Frankia* in *Casuarina glauca* actinorhizal nodules. *Plant Physiol.* **154**:1372–1380
- Persson T, Benson DR, Normand P, Vanden Heuvel B, Pujic P, Chertkov O, Teshima H, et al. 2011 Genome sequence of “*Candidatus Frankia datiscae*” Dg1, the uncultured microsymbiont from nitrogen-fixing root nodules of the Dicot *Datisca glomerata*. *J. Bacteriol.* **193** 7017–7018
- Popovici J, Comte G, Bagnarol E, Alloisio N, Fournier P, Bellvert F, Bertrand C and Fernandez MP 2010 Differential effects of rare specific flavonoids on compatible and incompatible strains in the *Myrica gale*-*Frankia* actinorhizal symbiosis. *Appl. Environ. Microbiol.* **76** 2451–2460
- Popovici J, Walker V, Bertrand C, Bellvert F, Fernandez MP and Comte G 2011 Strain specificity in the *Myricaceae*-*Frankia* symbiosis is correlated to plant root phenolics. *Funct. Plant Biol.* **38** 682–689.
- Sen A, Beauchemin N, Bruce D, Chain P, Chen A, Walston Davenport K, Deshpande S, Detter C, et al. 2013 Draft Genome sequence of *Frankia* sp. strain QA3, a nitrogen-fixing actinobacterium isolated from the root nodule of *Alnus nitida*. *Genome Announc.* **1** e00103-13
- Sen A, Sur S, Bothra AK, Benson DR, Normand P and Tisa LS 2008 The implication of life style on codon usage patterns and predicted highly expressed genes for three *Frankia* genomes. *Anton. van Leeuw. Int.* **93** 335–346
- Sen A, Thakur S, Bothra AK, Sur S and Tisa LS 2012 Identification of TTA codon containing genes in *Frankia* and exploration of the role of tRNA in regulating these genes. *Arch. Microbiol.* **194** 35–45
- Schwencke J and Carú M 2001 Advances in actinorhizal symbiosis: host plant-*Frankia* interactions, biology, and applications in arid land reclamation. A review. *Arid. Land Res. Manag.* **15** 285–327
- Udwary DW, Gontang EA, Jones AC, Jones CS, Schultz AW, Winter JM, Yang JY, Beauchemin N, et al. 2011 Significant natural product biosynthetic potential of actinorhizal symbionts of the genus *Frankia*, as revealed by comparative genomic and proteomic analyses. *Appl. Environ. Microbiol.* **77** 3617–3625
- Wall LG 2000 The actinorhizal symbiosis. *J. Plant Growth Regul.* **19** 167–182
- Wall L, Beauchemin N, Cantor MN, Chaia E, Chen A, Detter JC, Furnholm T, et al. 2013 Draft Genome sequence of *Frankia* sp. strain BCU110501, a nitrogen-fixing actinobacterium isolated from nodules of *Discaria trinevis*. *Genome Announc.* **1** e00503-13