ACS Synthetic Biology
Cite This: ACS Synth. Biol. 2018, 7, 2365−2378
Research Article
pubs.acs.org/synthbio

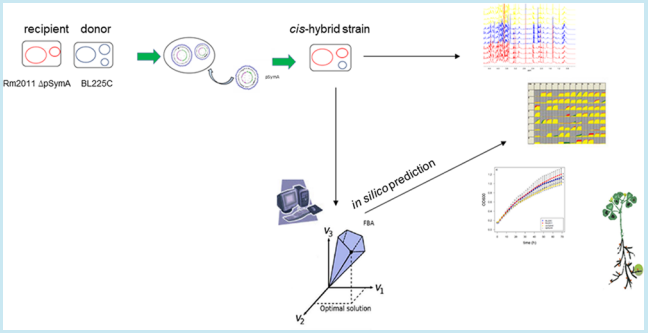
Creation and Characterization of a Genomically Hybrid Strain in the Nitrogen-Fixing Symbiotic Bacterium *Sinorhizobium meliloti*

Alice Checcucci,[†] George C. diCenzo,[†] Veronica Ghini,[‡] Marco Bazzicalupo,[†] Anke Becker,[§] Francesca Decorosi,[∥] Johannes Döhlemann,[§] Camilla Fagorzi,[†] Turlough M. Finan,[⊥] Marco Fondi,*[,†] Claudio Luchinat,[‡,#] Paola Turano,[‡,#] Tiziano Vignolini,[▽] Carlo Viti,[∥] and Alessio Mengoni*[,†]

[†]Department of Biology, University of Florence, 50019 Sesto Fiorentino, Italy
[‡]CERM & CIRMMP, University of Florence, 50019 Sesto Fiorentino, Italy
[§]LOEWE − Center for Synthetic Microbiology, 35043 Marburg, Germany
[∥]Department of Agri-food Production and Environmental Science, University of Florence, 50019 Florence, Italy
[⊥]Department of Biology, McMaster University, Hamilton, Ontario L8S 4L8, Canada
[#]CERM and Department of Chemistry, University of Florence, 50019 Sesto Fiorentino, Italy
[▽]European Laboratory for Non-Linear Spectroscopy, LENS, 50019 Sesto Fiorentino, Italy

Supporting Information

ABSTRACT: Many bacteria, often associated with eukaryotic hosts and of relevance for biotechnological applications, harbor a multipartite genome composed of more than one replicon. Biotechnologically relevant phenotypes are often encoded by genes residing on the secondary replicons. A synthetic biology approach to developing enhanced strains for biotechnological purposes could therefore involve merging pieces or entire replicons from multiple strains into a single genome. Here we report the creation of a genomic hybrid strain in a model multipartite genome species, the plant-symbiotic bacterium *Sinorhizobium meliloti*. We term this strain as *cis*-hybrid, since it is produced by genomic material coming from the same species' pangenome. In particular, we moved the secondary replicon pSymA (accounting for nearly 20% of total genome content) from a donor *S. meliloti* strain to an acceptor strain. The *cis*-hybrid strain was screened for a panel of complex phenotypes (carbon/nitrogen utilization phenotypes, intra- and extracellular metabolomes, symbiosis, and various microbiological tests). Additionally, metabolic network reconstruction and constraint-based modeling were employed for *in silico* prediction of metabolic flux reorganization. Phenotypes of the *cis*-hybrid strain were in good agreement with those of both parental strains. Interestingly, the symbiotic phenotype showed a marked cultivar-specific improvement with the *cis*-hybrid strains compared to both parental strains. These results provide a proof-of-principle for the feasibility of genome-wide replicon-based remodelling of bacterial strains for improved biotechnological applications in precision agriculture.



KEYWORDS: *replicon independence, genome coadaptation, experimental transplantation, accessory genome, Sinorhizobium meliloti*

Interest in large-scale genome modification and synthetic bacterial chromosome construction has strongly increased over the past decade (for instance, see the literature[1,2]) with a goal of engineering bacterial strains with new or improved traits. However, phenotypes are often the result of the coordinated function of many genes acting together in a defined genome architecture.[3] Hence, the ability to predict the phenotypic outcomes of large-scale genome modification requires a precise understanding of the genetic and regulatory interactions between each gene or gene product in the genome. As such, there is a need for integrated approaches, combining experimental evidence with computational-based methods, to interpret and potentially predict the outcomes of genome-wide DNA manipulations.

In this context, multipartite (or divided) genomes (*i.e.*, genomes possessing more than one informational molecule) are particularly interesting. The genome of bacteria with a multipartite structure is typically composed of a principal chromosome that encodes the core housekeeping and metabolic genes essential for cellular life, and one (or more) secondary replicons (termed chromids and megaplasmids). More than 10% of the presently sequenced bacterial genomes are characterized by the presence of a multipartite architecture.[4,5] The secondary replicons can account for up to half of the total genome size, and their level of integration into cellular

Received: April 10, 2018
Published: September 17, 2018

DOI: 10.1021/acssynbio.8b00158
ACS Synth. Biol. 2018, 7, 2365−2378

regulatory and metabolic networks is variable.[6−9] In some cases, strong replicon-centric transcriptional networks have been suggested.[10−14] The apparently functional modularity of secondary replicons is particularly attractive from both ecological and biotechnological viewpoints. Indeed, secondary replicons might act as plug-and-play functional modules, potentially allowing the recipient strain to obtain previously untapped genetic information.[15] This, in turn, might allow the emergence of novel phenotypic features leading, for example, to the colonization of a new ecological niche.[16] Moreover, such modularity paves the way for large-scale, genome-wide manipulations of bacterial strains with a multipartite genome structure, by synthetically merging complex biotechnologically important traits in the same strain.[17]

However, it remains unclear to what extent complex phenotypes can be directly transferred into a recipient strain, as secondary replicons are in part coadapted to the host genome, for example, through regulatory interaction and/or inter-replicon metabolic cross-talk.[10,18−21] In the past years, there have been examples of conjugal transfer of large plasmids in some bacterial species, including rhizobia and agrobacteria.[22−34] However, especially for the very large (>1 Mb) megaplasmids of rhizobia that have been transferred between species, the transfer was accompanied by only partial transmission of the plasmid phenotypic features. However, interspecies transfer events are conceptually distinct from the proposal to create *cis*-hybrid strains through intraspecies (interstrain) replicon transfers. We are aware of only one study that deeply examined the phenotypic consequences of replacing a large (>800 kb) native secondary replicon with a homologous replicon of closely related strains. In that study, the third replicon of *Burkholderia cepacia* complex strains was mobilized and the effects on various phenotypes including virulence was examined.[35] It was found that in some cases, phenotypes were dependent solely on the secondary replicon, whereas in other cases, the phenotypes depended on genetic/regulatory interactions with the other replicons.[35] However, additional studies are required to examine the generalizability and the controllability of those observations.

To further test the feasibility, the stability, and the predictability of secondary replicon shuffling on the phenotype(s) of the cell, here we have performed experimental and *in silico* replicon transplantation between two bacterial strains. We used the symbiotic nitrogen-fixing bacterium *Sinorhizobium meliloti* as a model, considering that it is characterized by a well-studied multireplicon genome structure.[16,36] Additionally, *S. meliloti* represents a highly valuable microorganism in agriculture, as its symbiosis with crops like alfalfa is estimated to be worth more than $70 million/year in the United States.[37] The genome of the mostly commonly studied *S. meliloti* strains (the two very closely related strains Rm1021 and Rm2011) are composed by a chromosome (~3.7 Mb), and by two secondary replicons: a chromid (~1.7 Mb, called pSymB) and a megaplasmid (~1.4 Mb, called pSymA,). The pSymA megaplasmid determines many of the biotechnologically relevant functions, as it harbors most of the key genes involved in plant symbiotic colonization and in nitrogen-fixation.[38] The *S. meliloti* large replicons have recently been proposed as scaffolds for novel shuttle vectors for synthetic biology.[39] Furthermore, previously performed genome reduction experiments have led to the complete removal of one or both of the two secondary replicons,[16,40] and an *in silico* genome-scale metabolic model has been reconstructed,[41] paving the way for massive genome-scale remodelling of *S. meliloti*.

Here, we constructed a hybrid strain containing the chromosome and the chromid of the laboratory *S. meliloti* Rm2011 strain with the pSymA replicon from the wild isolate *S. meliloti* BL225C. The genome of BL225C is 290 kbps larger than that of Rm2011;[36,42] moreover, up to 1,583 genes are differentially present in the genomes of these strains.[43,44] Furthermore, the BL225C strain has been shown to have several interesting biotechnological features, including plant growth promotion, and nodulation efficiency.[36,45] The majority of the genetic differences between Rm2011 and BL225C are associated with the homologous pSymA and pSINMEB01 homologous megaplasmids (as shown in Supplemental Figure S1); 836 of the 1,583 variable genes are located on this replicon.[43] We therefore expect that creating a hybrid strain between Rm2011 and BL225C, by moving the pSymA-equivalent from BL225C to the Rm2011 derivative lacking pSymA, will provide a good testing ground to examine (i) the feasibility of large replicon shuffling between strains, and (ii) the stability and predictability of the phenotypes linked to such replicons. We term this novel hybrid strain as *cis*-hybrid since it derives from *cis*-genic manipulation and contains genetic material from the pangenome pool of the same species (in contrast to a *trans*-genic strain that would contain genes from a distinct species). *cis*-Hybrid strains could be an important way to promote environmental-friendly and regulatory compliant biotechnology and synthetic biology in bacterial species of interest in agricultural and environmental microbiology.[17]

RESULTS AND DISCUSSION

We report here the creation of a *cis*-hybrid *S. meliloti* strain, where the symbiotic megaplasmid pSINMEB01, was transferred from the natural strain BL225C to the laboratory strain Rm2011. The pSINMEB01 megaplasmid is homologous to pSymA, and is ~1.6 Mb in size, accounting for nearly 23% of total genome. *In silico* metabolic reconstruction and a large set of phenotypic tests, including Nuclear Magnetic Resonance (NMR)-based metabolomic profiling, Phenotype Microarray, and symbiotic assays with different host plant cultivars have been performed.

Experimental Creation of a *cis*-Hybrid Strain. Starting with the derivative of the *S. meliloti* Rm2011 strain that lacks the pSymA replicon, herein referred to as ΔpSymA,[46] we produced a *cis*-hybrid strain that contains the Rm2011 chromosome and pSymB chromid, and the pSymA replicon from a genetically and phenotypically distinct *S. meliloti* strain, BL225C (all strains used in this work are listed in Table 1).[36,45] The *cis*-hybrid strain was produced through a series of conjugations as described in the Methods section. Briefly, a plasmid for overexpressing *rctB*[47] was transferred to BL225C; as RctB is a negative regulator of RctA,[47] which in turn is a negative regulator of the pSymA conjugal genes,[48,49] this step was necessary to promote pSymA transfer without mutating the replicon. Concurrently, a plasmid carrying an antibiotic resistance gene marker (gentamycin, plasmid pMP7605[50]) was transferred to the ΔpSymA strain to allow for the use of the gentamycin resistance marker in the selection of *cis*-hybrid transconjugants in the next step. Finally, a mating mixture of the ΔpSymA acceptor strain and the BL225C donor strain was prepared, and *cis*-hybrid transconjugants were isolated on a medium selective for the gain of the pSymA replicon (see Methods). Correct construction of the *cis*-hybrid strain was

DOI: 10.1021/acssynbio.8b00158
ACS Synth. Biol. 2018, 7, 2365−2378

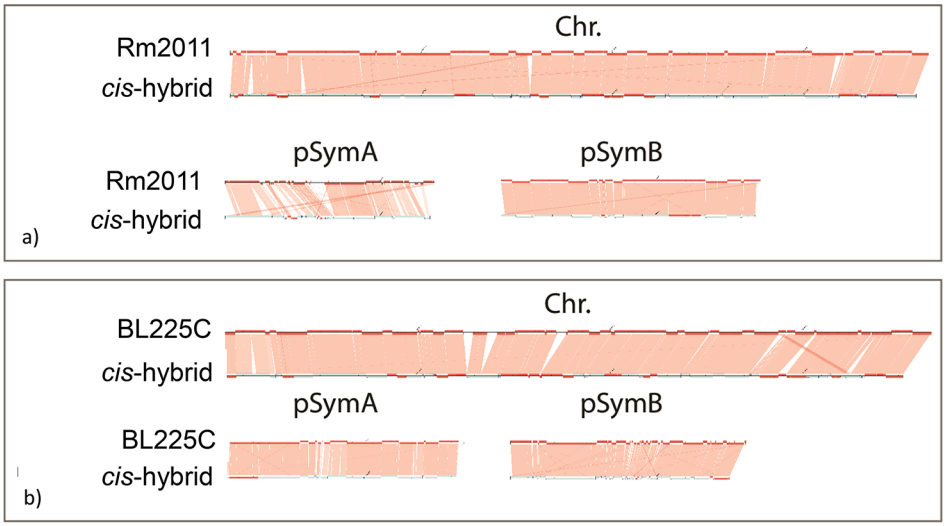
Table 1. Strains and Plasmids Used in This Study[a]

strain or plasmid	description	reference
Sinorhizobium meliloti		
Rm2011	Wild type SU47 derivative; Sm^R	Sallet *et al.*, 2013[42]
BL225C	Wild isolate from *Medicago sativa* in Lodi (Italy)	Galardini *et al.*, 2011[36]
RmP3498	Rm2011 ΔpSymAB+B with engA and tRNA into the chromosome; Sm^R Sp^R	diCenzo *et al.*, 2014[16]
BM 826	RmP3498 with pMp7605; Sm^R Sp^R Gm^R	This study
BM 806	BM 826 with pSymA from BL225C; Sm^R Sp^R Gm^R	This study
BM 848	BL225C with pTE3rctB; Tc^R	This study
Escherichia coli		
MT616	Helper strain carrying pRK600 that has the RK2 *tra* genes; Cm^R	Finan *et al.*, 1986[31]
Plasmids		
pMp7605	Broad host range vector constitutively expressing the *mCherry* gene; Gm^R	Lagendijk *et al.*, 2010[50]
pTE3rctB	Broad host range vector overexpressing the *Rhizobium etli rctB* gene; Tc^R	Nogales *et al.*, 2013[47]

[a]Code of strains and plasmid is reported. A succinct description of the main phenotypic features is shown; Sm^R, streptomycin resistance, Sp^R, spectynomcin resistance, Gm^R gentamycin resistance, Tc^R tetracyclin resistance, Cm^R, chloramphenicol resistance.

initially confirmed through PCR amplification on specific unique genes on the Rm2011 chromosome and pSymB, and the pSymA homologue of BL225C (pSINMEB01) (see Table S1). Subsequently, whole genome sequencing (Figure 1, Supplemental Figure S1) confirmed the complete transfer of pSymA, and the banding pattern observed in pulse-field gel electrophoresis (PFGE) (Supplemental Figure S2) was consistent with pSINMEB01 being present as an independent replicon (*i.e.*, not integrated into the chromosome or pSymB).

***In Silico* Metabolic Network Reconstruction.** In addition to the experimental creation of the *cis*-hybrid strain, we attempted to predict its metabolic outcomes by generating a new metabolic model which includes the genomic features present in the *cis*-hybrid strain. The curated iGD1575 reconstruction (herein referred to as the Rm2011 reconstruction) was used to represent metabolism of *S. meliloti* Rm2011;[46] although iGD1575 is based on the strain Rm1021, the genomic content of these strains is 99,9% identical, with the exception of numerous SNPs[42] that are not considered in the process of metabolic reconstruction. Next, our recently described pipeline[51] was used to build a representation of BL225C based on a draft reconstruction built with the KBase Web server and enhanced based on the iGD1575 model. An *in silico* representation of the *cis*-hybrid strain was then built by removing all pSymA genes (and dependent reactions) from the Rm2011 model, followed by the addition of all pSymA (pSINMEB01) genes (and associated reactions) from the BL225C model using our published pipeline.[51] Despite there being numerous (47 to 143 gene) differences in the gene content of the metabolic reconstructions, the Rm2011 model differed from the BL225C and the *cis*-hybrid models by no more than a half dozen reactions (Table 2). The low reaction variability between models may (i) reflects the difficulty in predicting the function of the *S. meliloti* variable gene content, (ii) suggests the presence of nonorthologous genes encoding proteins catalyzing the same reaction(s), and/or (iii) indicates that few metabolic features are dependent on the accessory gene set. Not surprisingly, given the near identical reaction content of the reconstructions, the outputs of flux balance analysis simulations for the different reconstructions



cis-hybrid	Rm2011	BL225C
Chromosome	100%	99.97%
pSymA	98.62%	100%
pSymB	100%	99.19%

c)

Figure 1. Confirmation of the genome structure of the *cis*-hybrid strain. (a) Comparison of *cis*-hybrid strain genome sequence with Rm2011 chromosome and pSymB, and with pSymA of the donor strain (BL225C) (contigs alignment performed with Contiguator). Apparent indels in the alignment of the *cis*-hybrid assembly with the Rm2011 genome are artifacts of the assembly, as revealed by directly mapping sequencing reads to the reference Rm2011 genome (Figure S1). (b) Percentage of identity of each replicon which composed the multipartite genome of *cis*-hybrid strain with those of the donor strains (Rm2011 and BL225C).

DOI: 10.1021/acssynbio.8b00158
ACS Synth. Biol. 2018, 7, 2365−2378

Table 2. Comparison of *S. meliloti* Metabolic Network Reconstructions[a]

	genes			reactions		
strain	BL225C	*cis*-hybrid	ΔpSymA	BL225C	*cis*-hybrid	ΔpSymA
Rm2011	1525/52/91	1551/26/76	1336/241/0	1821/6/6	1823/4/6	1755/72/0
BL225C	–	1598/18/29	1308/308/28	–	1827/0/2	1753/74/2
cis-hybrid	–	–	1336/291/0	–	–	1755/74/0

[a]The gene and reaction content of the four *S. meliloti* metabolic reconstructions used in this work are shown. For each cell, the values are a comparison of the strain indicated on the left with the strain indicated along the top. Three values are provided in each cell, and these correspond to the following. The first value is the number of genes or reactions in common between the models. The second value is the number of genes or reactions present in the reconstruction on the left but not in the one along the top. The third value is the number of genes or reactions present in the reconstruction along the top but not in the one on the left.

were nearly identical (data not shown); therefore, we do not describe these results further.

Metabolic Phenotypes and Profiles of the *cis*-Hybrid Strain. To confirm that the process of replicon swapping did not result in any unintended metabolic perturbations, the *cis*-hybrid strain and the parental strains were characterized using Phenotype MicroArray experiments to test metabolic capacity, and using ^{1}H nuclear magnetic resonance (NMR) to compare their metabolomic profiles.

Phenotype MicroArray experiments were performed to test the substrate utilization abilities (as presence of an oxidative metabolism, hence growth) of the *cis*-hybrid strain, as well as the parental and wild type strains, with 192 different carbon sources and 96 different nitrogen sources. Previous work has shown that the pSymA megaplasmid has little contribution to the metabolic capacity of *S. meliloti*.[16,40] Consequently, we expected a small number of differences in the *cis*-hybrid strain. A previous Phenotype MicroArray experiment identified only four carbon substrates (3-methyl glucose, D-gluconic acid, D-ribono-1,4-lactone, and β-D-allose) and one nitrogen source (cytosine) that appeared dependent on pSymA in Rm2011.[16] 3-methyl glucose was the only carbon source of these four that supported growth of Rm2011 in the current study. The contribution of pSymA to metabolism of BL225C has not been elucidated. Consistent with the past results, only minor changes in the substrate utilization abilities were observed following the introduction of the pSymA of BL225C (pSINMEB01) into the ΔpSymA strain (Figure 2, Table S2). Importantly, the ΔpSymA strain did not grow with 3-methyl glucose as a carbon source and cytosine as a nitrogen source. Both these metabolic abilities were restored in the *cis*-hybrid strain. These results confirm the expression of metabolic functions the pSymA homologue in the *cis*-hybrid strain.

Concerning the differences between the *cis*-hybrid strain and the two parental ones, Rm2011 and BL225C, the *cis*-hybrid strain showed a higher activity index (a metrics of substrate utilization[52]) with respect to both the parental strains on 22 out of 192 (11.4%) carbon sources tested, and a lower activity index in 29 of 192 (15.1%) carbon sources (Supplemental Table S2a). For the nitrogen sources utilization (Supplementary Table S2b), the *cis*-hybrid strain showed higher activity index values with respect to the parental strains in 6 out of 96 (6.2%) tested compounds, and a lower activity index in 27 of the 96 (28.1%) compounds. Most of the compounds on which the *cis*-hybrid strain displayed different activity index with respect to both parental strains (as for instance formic acid, propionic acids, pyruvic acid, L-alanine, L-serine *etc.*) are the same shown previously accounting for metabolic differences by *S. meliloti* wild-type strains, including Rm1021 and BL225C.[45]

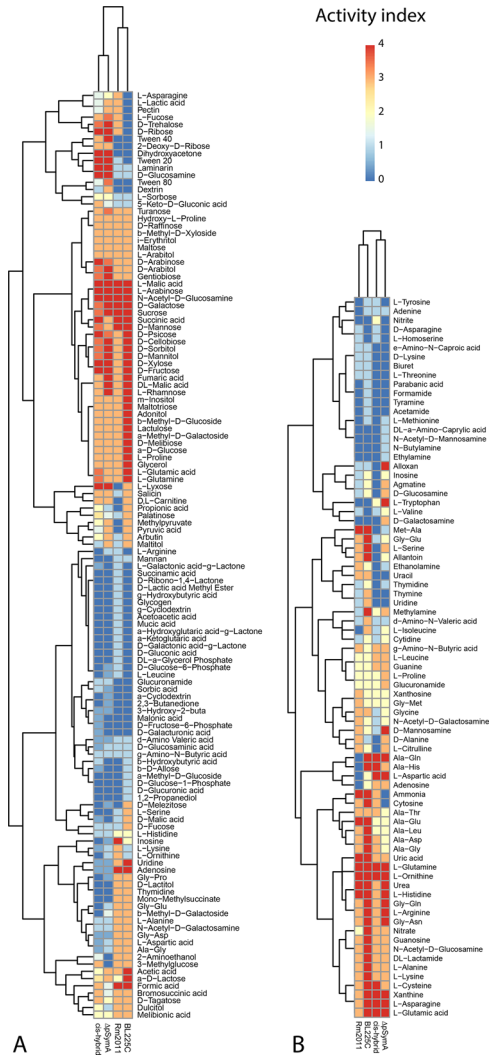


Figure 2. Metabolic phenotype of the *cis*-hybrid strain. Heatmap of Phenotype Microarray profiles of the growth on different carbon and nitrogen sources for Rm2011, BL225C, *cis*-hybrid and ΔpSymA strains. (a) Heatmap with Euclidean clustering; (b) values of pairwise Euclidean distances.

We further inspected if the changes in metabolic abilities were related to genetic differences in the megaplasmids. This analysis was performed with DuctApe dape module.[52] In 87 of 192 carbon sources, the *cis*- hybrid strain showed different metabolic activity with respect to the sole parental Rm2011, with whom it shares the chromosomal and chromid background. Among these, the *cis*-hybrid activity index is lower in 48 and higher in 39 of the 87 carbon sources. Only in few cases

DOI: 10.1021/acssynbio.8b00158
ACS Synth. Biol. 2018, 7, 2365−2378

the observed differences mirror the replicon-localization of the metabolic pathways. Indeed, a small number of genes involved in the metabolic pathways for the carbon sources are located on pSymA (for instance, SM2011_a0233 of Rm2011 and SinmeB_5530 of BL225C involved in D-trehalose metabolism, SM2011_a0796, SM2011_a1844, SM2011_a2213 of Rm2011 and SinmeB_5059, SinmeB_5059, SinmeB_6004 of BL225C involved in acetic acid metabolism, SM2011_a0878 of Rm2011 and SinmeB_6071 of BL225C involved in D-fructose-6-phosphate metabolism, SM2011_a0398 and SM2011_a0306 of Rm2011 and SinmeB_6485 and SinmeB_6542 of BL225C in histidine metabolism). Only few of such genes are present in one megaplasmid only (for instance, the Rm2011 aminotransferase related with L-alanine metabolism SM2011_a1495 and the ethanolamine ammonia-lyase subunit EutB (SinmeB_5765) involved in 2-aminoethanol metabolism, and the glucoamylase SinmeB_5936 of dextrin metabolism, present only in BL225C). The same trend was observed for the nitrogen sources: in 48 of the 96 tested, the *cis*- hybrid strain showed different metabolic activity with respect to the parental Rm2011 only, in 31 cases its activity index is lower and in 17 it is higher respect to the reference strain. Also in this case, only few genes involved in the nitrogen sources metabolism are present in one of megaplamids only (for instance, SinmeB_5853, involved in L-asparagine metabolism and SinmeB_5765 involved in agmatina metabolism, on pSymA homologue of BL225C). Interestingly, in some cases one gene present on pSymA of BL225C (SinmeB_6469) takes part to metabolic reactions which in Rm2011 request only genes on the chromosome (as in the case of guanosine, inosine and xanthosine). In this situation, the additional participation of a gene on BL225C pSymA-homologue (pSINMEB01), which is necessary in the metabolism of the strain, can determine an increase or a decrease of metabolism in a different chromosomal background. In conclusion, the DuctApe dape analysis showed that only part of the observed differences between the *cis*-hybrid strain and the parental ones can be related with genetics differences (as genes loss or acquisition) at the pSymA level. In agreement with such analysis is interesting to note that the substrate utilization abilities of the *cis*-hybrid strain were difficultly predictable with the metabolic modeling also. This is perhaps not surprising given that less than 1% of reactions differed between the *in silico* metabolic reconstructions of the strains. Consequently, we can speculate that the differences in Phenotype Microarray attributed of the *cis*-hybrid strain could be associated with regulatory differences, changes in gene expression levels and/or with alleles with different functionality, all differences which are not considered in the metabolic modeling.

To more in deep investigate the possible metabolic alterations following pSymA transplantation, a metabolomic analysis through NMR was performed. Using an untargeted approach, both cellular lysates and spent growth media were analyzed to identify the fingerprint of the endo- and exometabolomes of the two parental strains, the *cis*-hybrid strain, and the ΔpSymA recipient.

Principal component analysis (PCA) was used to generate an initial overview of the metabolome differences among the four strains (Figure 3a,b), followed by PCA-Canonical Analysis (CA) to obtain the best discrimination among the strains by maximizing the differences among their metabolomic profiles (Figure 3c,d). In both the PCA and PCA-CA score plots (Figure 3), the *cis*-hybrid strain clustered very close to both the ΔpSymA recipient strain and to the parental strain Rm2011, whereas the parental donor strain BL225C clustered separately. These results are consistent with previous data indicating that pSymA has little contribution to the metabolome,[7] proteome,[8] or transcriptome[53] of *S. meliloti* Rm2011 in laboratory conditions. Importantly, these results confirmed that the synthetic large-scale horizontal gene transfer performed here to produce the *cis*-hybrid strain did not result in a major perturbation of the cellular metabolism.

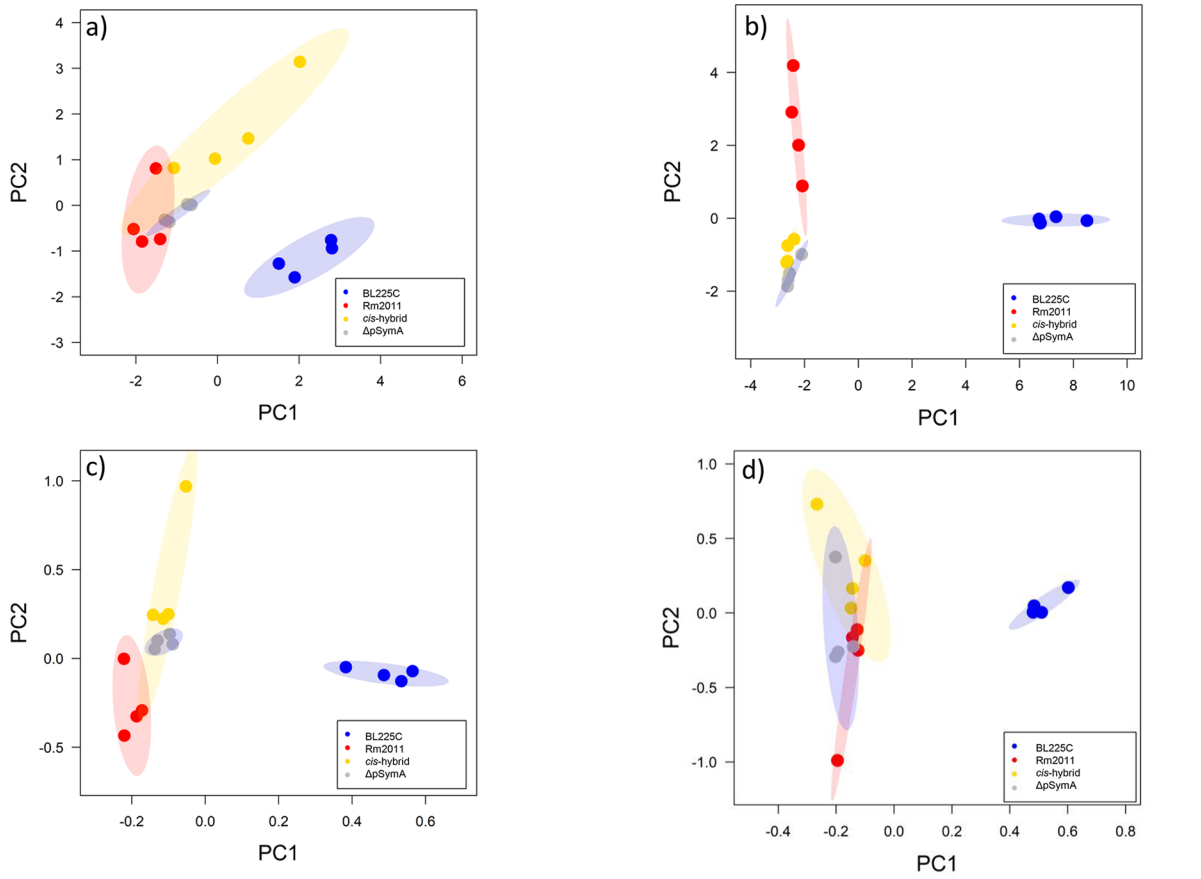
In addition to the multivariate analysis of the metabolic NMR fingerprints described above, the signals of 25 and 19 metabolites were unambiguously assigned and integrated in the ^{1}H NMR spectra of the cell lysates and growth media, respectively (Figure S3). The metabolites that are characterized by statistically significant differences in concentration levels in at least one strain with respect to the two other strains are indicated in Supplementary Table S3 and are also reported in Supplemental Figure S4. Validating the ability of this approach to detect metabolic differences between the strains, it was noted that the ΔpSymA strain exported cytosine unlike the wild type Rm2011 or the *cis*-hybrid strain, consistent with the inability of this strain to catabolize cytosine as shown by the Phenotype MicroArray data.

Overall, the data reported here are consistent with (i) at least some of the genes of the homologous pSymA being properly expressed, and their gene products function, and (ii) that the replicon swapping procedure did not result in any major, unintentional metabolic consequences.

Assessment of the Phenotypes of the *cis*-Hybrid Strain. *Growth Profiles in Synthetic Laboratory Media.* Growth profiles of the *cis*-hybrid and parental strains in complex (TY) and defined (M9-succinate) media are reported in Figure 4. In the complex TY medium (Figure 4a), the growth of the *cis*-hybrid strain was impaired relative to the recipient (ΔpSymA), and to the Rm2011 and BL225C parental strains. In other words, gain of the BL225C pSymA replicon by the ΔpSymA strain resulted in a decrease in the growth rate in TY medium. Although we cannot provide a definitive explanation for this phenomenon, it may be that the simultaneous gain of hundreds of new genes not integrated into cellular networks imposes a high metabolic cost to the cell, resulting in impaired fitness. In contrast, little to no difference was observed in the growth rate of any of the strains in the defined M9-succinate medium (Figure 4b). The lack of an observable growth impairment of the *cis*-hybrid strain in the M9-succinate medium may be due to the masked general decrease in growth rate of all strains in this medium. Moreover, the similarity of the growth profiles of all strains in the minimal medium suggest that, at least in artificial laboratory conditions, the main growth characteristics of these strains are primarily dependent on the core, not accessory, genome.

Growth Using Root Exudate as a Nutrient Source. Root exudates can be considered a proxy of the nutritional conditions of the plant rhizosphere and contain numerous compounds that can be used by bacteria as carbon and nitrogen sources.[54] We therefore evaluated the ability of the four strains to grow on M9 mineral medium supplemented with root exudates of *Medicago sativa*, a *S. meliloti* symbiotic partner. None of the strains were able to grow when the root exudate was used as the sole carbon source (data not shown); this was likely due to a very low amount of carbon sources contained in the root exudate unable to support the growth. In contrast, all strains can utilize the root exudate as the sole nitrogen source

DOI: 10.1021/acssynbio.8b00158
ACS Synth. Biol. 2018, 7, 2365−2378



	BL225C	Rm2011	*cis*-hybrid	ΔpSymA
BL225C	**1**	0	0	0
Rm2011	0	**1**	0	0
cis-hybrid	0	0	**1**	0
ΔpSymA	0	0.25	0	**0.75**

Discrimination accuracy 93.7%

	BL225C	Rm2011	*cis*-hybrid	ΔpSymA
BL225C	**1**	0	0	0
Rm2011	0	**0.75**	0.25	0
cis-hybrid	0	0	**0.75**	0.25
ΔpSymA	0	0	0	**1**

Discrimination accuracy 87.5%

Figure 3. ^{1}H nuclear magnetic resonance (NMR)-based metabolomic profiles of cellular lysates and growth media of Rm2011, BL225C, *cis*-hybrid and ΔpSymA strains. Score plot of PCA (a,b) and PCA-CA (c,d) analysis of cell lysates (a,c) growing media (b,d). The confusion matrices and the discrimination accuracy values for PCA-CA analysis are also reported. Ellipses in the score plots illustrate the 95% confidence level.

when provided succinate as a carbon source to the M9 medium, as confirmed in the control the nitrogen-free M9 medium (Supplementary Figure S5). Interestingly, different growth kinetics were observed (Figure 4c). In particular, BL225C displayed the highest growth among all four strains when grown with root exudates as a sole nitrogen source. Plating for viable colony forming units confirmed the differences in the final population densities (data not shown). As the robust growth of BL225C with root exudates did not transfer to the *cis*-hybrid strain, it is likely that this phenotype is primarily dependent on the chromosome and/or pSymB of BL225C, as was suggested by previous studies.[16,43,46,55] This observation would further suggest that the adaptation of the tested strains to growth in the rhizosphere occurred prior to the gain of pSymA and symbiotic abilities, consistent with recent work indicating that the majority of *S. meliloti* rhizosphere growth-promoting genes are chromosomally encoded.[56] Finally, considering that there are relatively few differences in the nitrogen metabolic capacity of Rm2011 and BL225C,[45] and that FBA (flux balance analysis) simulations for the metabolic model reconstructions were nearly identical (data not shown), we hypothesize that the growth differences observed between these strains are primarily related to regulatory differences, and less so to differences in metabolic genes.

Biofilm Formation. Biofilm formation is a key factor in root colonization and plant invasion for many Proteobacteria.[57] In particular, optimal colonization of the roots by rhizobia can influence nodule formation efficiency and competitiveness.[58,59] The formation of biofilm is a complex phenomenon and is determined by a series of regulatory mechanism involving many genes.[59] Consequently, to evaluate if the large genomic manipulation of the *cis*-hybrid strain resulted in a possible impairment of this process, biofilm formation was measured for the *cis*-hybrid strain and the parental strains in two growth conditions (the complex TY medium and the defined Rhizobium defined medium (RDM) medium). Biofilm production (estimated as the total biofilm-to-biomass ratio) by the *cis*-hybrid was similar to the parental strains in TY medium, and was higher than the parental strains in minimal RDM medium ($p < 0.005$; Figure S6). This result supports the robustness of this phenotype over the genomic manipulation performed. The higher production of biofilm in RDM by the *cis*-hybrid may additionally suggest that some still unknown

DOI: 10.1021/acssynbio.8b00158
ACS Synth. Biol. 2018, 7, 2365−2378

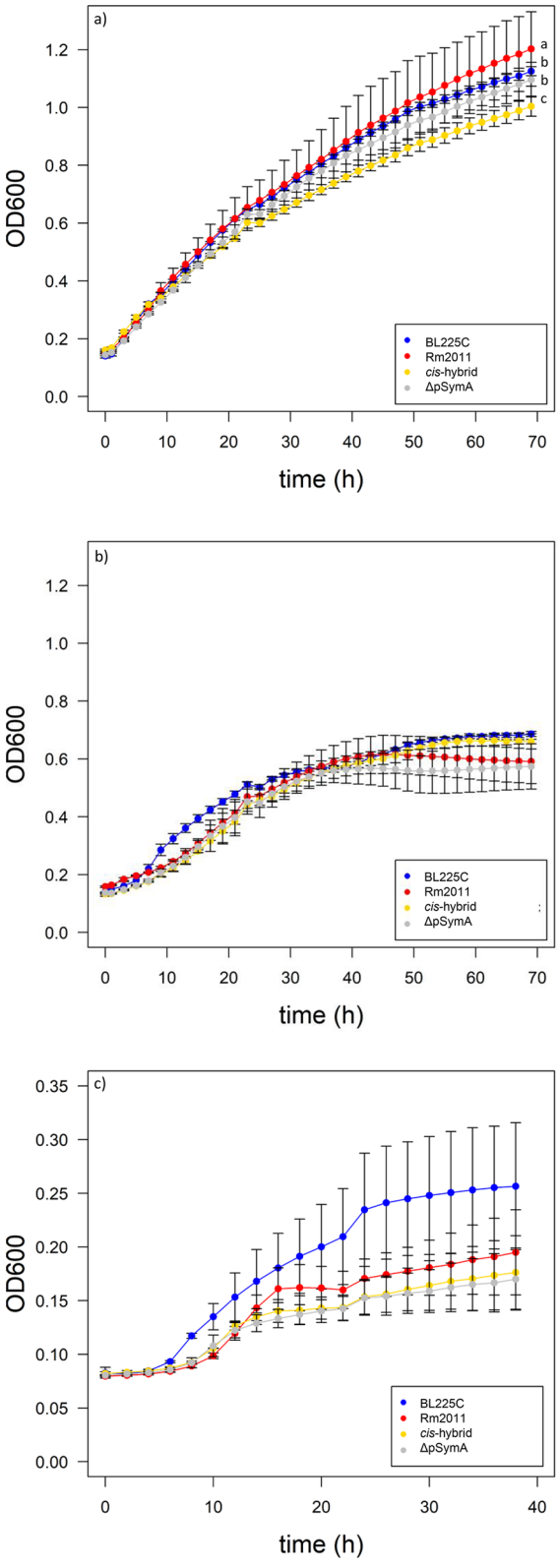


Figure 4. Growth phenotypes of the *cis*-hybrid strain. The growth of *S. meliloti* was examined in TY complex medium (a) M9 minimal medium (b) and (c) M9 + succinate and root exudates as sole N source. Data points represent averages from quadruplicate measurements. The letters on the curves represent the statistically significant differences among the strains growth ($p < 0.005$, Tukey *post hoc* contrasts).

regulatory mechanisms may have been modified. This difference between TY and RDM is also observed for the ΔpSymA recipient strain, which produced a higher ($p < 0.005$) level of biofilm with respect to the other three strains in the complex TY medium (Figure S6a), but a lower or similar level in the defined medium RDM (Figure S6b), as already described for the deletion of the common *nod* genes (harbored by pSymA) and of the entire pSymA.[60] The different phenotype observed on TY and RDM media further supports the suggestion that during growth in complex medium, there may be a still unknown pSymA-mediated negative regulation of biofilm formation.

Symbiotic Phenotypes of the *cis*-Hybrid Strain. Many of the key genes required for symbiotic abilities (*e.g.*, nodule formation and nitrogen-fixation) are present on pSymA of *S. meliloti* Rm2011[38,61] and the homologous megaplasmid pSINMEB01 of BL225C.[36] These replicons additionally contain nonessential genes that promote improved symbiotic abilities.[62] While many of the symbiotic genes are conserved between these strains, relevant differences between pSymA and pSINMEB01 are present. The 482 genes exclusive to pSINMEB01 encode symbiotic (*e.g.*, *nws*, *hemA* homologue, C P450[36]) and nonsymbiotic functions (*e.g.*, *acdS*).[63] For these reasons, this replicon swapping study was initiated in large part to evaluate whether swapping the symbiotic megaplasmid could promote differential symbiotic abilities.

To test the robustness of symbiotic abilities following replicon transplantation, *in vitro* symbiotic assays were performed on a panel of alfalfa cultivars, as alfalfa is the main host legume of *S. meliloti*[64] (Figure 5). In particular, the *cis*-hybrid strain and the two parental strains (Rm2011, BL225C) were tested in combination with six alfalfa cultivars (Table S4). These cultivars belong to the species *Medicago sativa*, *Medicago x varia*, and *Medicago falcata*, and they are representative of the variability of cultivars and germplasm mainly used as crops in Europe. Moreover, BL225C was originally isolated on the *M. sativa* cultivar "Lodi" at the CREA-FLC institute (Italy) during a long-course experiment.[65] The percentage of nodulated plants (Figure 5a), the number of nodules per plant (Figure 5b), the plant aerial part length (Figure 5c), the shoot dry weight (Figure 5d), and nodule colonization abilities (Supplemental Figure S7) were recorded using standard procedures.[45,66] Not surprisingly, for each strain there was high variability in the symbiotic phenotypes observed with the different cultivars. The symbiotic interaction is a multistep developmental process which involves a tight exchange of signals between the bacterium and the plant root at both rhizospheric and endophytic levels.[17,67] Earlier works have demonstrated strain and cultivar specificities in this process, which results in *S. meliloti* strains displaying differential symbiotic effectiveness with various plant genotypes.[65,68,69]

The *cis*-hybrid strain performed very poorly in symbiosis with some cultivars, such as in the cultivars "Prosementi", "Camporegio" and "Verbena". In particular, the number of nodules per plant and the length of the aerial part were lower (Figure 5b,c, Table S5) ($p < 0.005$), indicating that the pSymA and pSINMEB01 megaplasmids are not always interchangeable. This could reflect the importance of the genomic context of the symbiotic megaplasmid and hypothetically the importance of inter-replicon regulatory networks.[10,70,71] Alternatively, pSINMEB01 may lack important (but still unknown) symbiotic genes whose function may be replaced by chromosomal genes in BL225C but not by chromosomal genes in Rm2011.

DOI: 10.1021/acssynbio.8b00158
ACS Synth. Biol. 2018, 7, 2365−2378

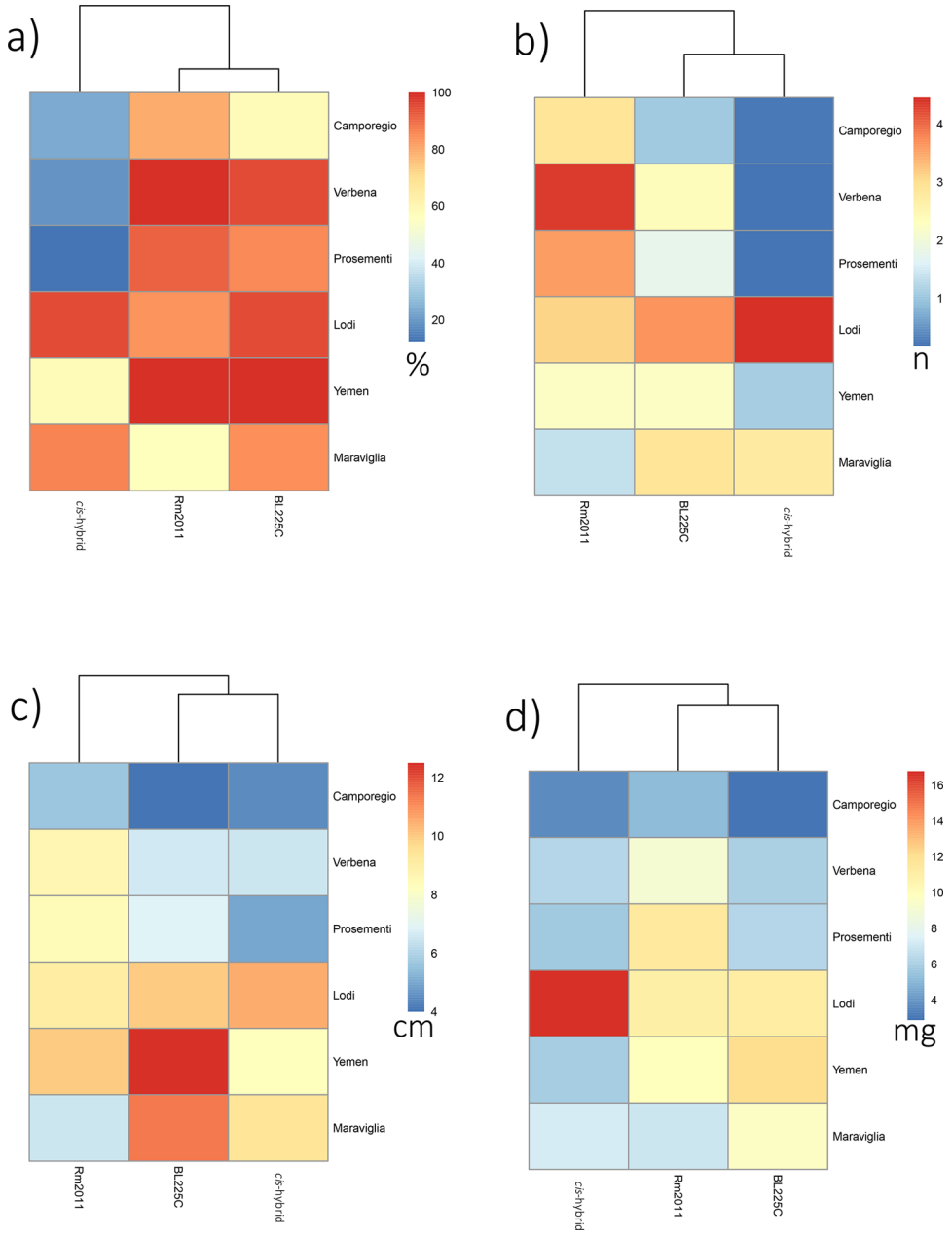


Figure 5. Symbiotic capabilities of the *cis*-hybrid strain. Heatmaps of symbiotic performances profiles for Rm2011, BL225C and cis-hybrid strains in a panel of six alfalfa cultivars; (a) percentage of nodulated plants, (b) number of nodules per plant, (c) plant aerial part length (cm) and (d) the shoot dry weight (mg).

Strikingly, the *cis*-hybrid strain displayed clearly improved symbiotic capabilities during symbiosis with the cultivar "Lodi" compared to both Rm2011 and BL225C (Figure 5). This was true for several key measures of symbiosis, including nodule number per plant, shoot dry weight, and length of the aerial part of the plant. The competitive abilities of the *cis*-hybrid strains, assessed through a coinoculation experiment, were similar to those of the two parental strains (Rm2011 and BL225C) (Figure S8).

These data suggest the presence of nonlinear and genomic context dependent genic interactions in the establishment of symbiotic abilities. Such interactions may resemble (at the logic level) those present in some eukaryotic genomes that result in the so-called "hybrid vigor", *i.e.*, the tendency for hybrids to be superior to the parental genotypes.[72] However, since hybrid vigor is related to heterozygosis, in our case we may speculate that strain-by-strain variability of regulons,[10] as well as the metabolic redundancy of *S. meliloti* genome[5,73,74] (which could in some way mimic the presence of multiple alleles) could contribute to the increase in the observed symbiosis-related phenotypes.

Summing up, these data highlight the potential of a large-scale genome manipulation approach to obtain highly effective, and cultivar specific, rhizobial strains. This provides a rational basis for the use of similar approaches in the development of elite bioinoculants for use in precision agriculture.[17,75]

CONCLUSIONS

The work presented here provides a proof-of-principle for the feasibility of using a large-scale genome manipulation approach

DOI: 10.1021/acssynbio.8b00158
ACS Synth. Biol. 2018, 7, 2365−2378

that makes use of the species' pangenome (*i.e.*, the extended gene set present in a group of microbial strains belonging to the same species[76]) to produce daughter strains with improved biotechnologically relevant (*i.e.*, nitrogen fixing symbiosis) characteristics.[17] In the current work, the large-scale genome manipulation was based on the transplantation of the primary symbiotic megaplasmid of a bacterial multipartite genome, a genome organization commonly found in the rhizobia. Although an entire replicon accounting for more than 20% of the total genome content was replaced with a homologous replicon of a closely related strain resulting in the gain of 482 new genes (in addition to numerous SNPs) and the loss of 354 genes, most of the core metabolic phenotypes appeared largely resilient to modification with this approach. However, other phenotypes, particularly complex (*i.e.*, multigenic) phenotypes such as the symbiotic phenotypes, gave interesting features that support the validity of this approach to improve biotechnologically relevant properties.

METHODS

Microbiological and Genetic Methods. Strains and plasmids used in this study are described in Table 1. Conjugation between *E. coli* and *S. meliloti* were performed as described in the literature.[77] All growth media (LB, LBmc, TY, M9, RDM) and antibiotic concentrations were as previously described.[16,73,78]

***cis*-Hybrid Strain Construction.** First, a triparental mating between the wild type strain BL225C (the future donor), the helper strain *E. coli* MT616 (carrying pRK600 that has the RK2 *tra* genes),[31] and *E. coli* with the pTE3rctB vector (replicative plasmid overexpressing the *R. etli rctB* gene and carrying a tetracycline resistance marker)[47] was performed to create the BM848 (BL225C-rctB) strain. Second, a biparental mating between *S. meliloti* Rm3498 (ΔpSymA)[16] and an *E. coli* S17−1 strain carrying the pMp7605 vector (carrying a gentamycin resistance marker)[50] was performed to generate the strain BM826. Lastly, the *cis*-hybrid strain BM806 was created through a biparental mating between the strain BM848 (BL225C-rctB) as the donor and the strain BM826 (ΔpSymA + pMp7605) as the acceptor. Selection for the *cis*-hybrid transconjugant strain (which had the pSymA replicon of the donor strain) was performed on M9 medium containing 1 mM $MgSO_4$, 0.25 mM $CaCl_2$, 0.001 mg/mL biotin, 42 μM $CoCl_2$, 76 μM $FeCl_2$, 10 mM trigonelline, streptomycin, and gentamycin. Streptomycin and gentamycin were used to select for the recipient strain, while the presence of trigonelline as the sole carbon source selected for the gain of pSINMEB01, as the trigonelline catabolic genes are located on pSymA/pSINMEB01.[79]

Validation of the Transplanted Strain. Pulsed-Field Gel Electrophoresis (PFGE) was performed to verify the successful uptake of pSymA *via* restriction digestion of genomic DNA with *Pme*I. The applied PFGE protocol was modified from Herschleb et. al 2007[80] and Mavingui *et al.* 2002,[81] and a protocol from Sharon Long's research group (Stanford University, available at http://cmgm.stanford.edu/biology/long/files/protocols/Purification%20of%20S%20meliloti.pdf). *S. meliloti* cultures were grown to an OD_{600} of 1.0 in TY medium supplemented with suitable antibiotics and harvested by centrifugation (3000*g*, 15 min, 4 °C). All following steps were carried out either on ice or at 4 °C. Sedimented cells were washed with TE buffer (10 mM Tris-HCl, 1 mM EDTA) supplemented with 0.1% (w/v) *N*-Lauroylsarcosine, and a second time with TE buffer. Washed cell pellets were then resuspended in TE buffer and mixed (1:1) with 1.6% (w/v) low-melt agarose (50 °C), thereby resulting in a final concentration of ∼8×10^8 cells/mL. Two hundred μL of each suspension was casted into a moistened mold and gelatinized at 4 °C. The resulting agar plugs were subsequently incubated at 37 °C for 3 h in lysis buffer (6 mM Tris-HCl, 1 M NaCl, 100 mM EDTA, 0.5% (w/v) Brij-58, 0.2% (w/v) Sodium deoxycholate, 0.5% (w/v) *N*-Lauroylsarcosine) supplemented with 1.5 mg/mL lysozyme (SERVA Electrophoresis GmbH, Germany). Treated agar plugs were then washed in H_2O, followed by incubation at 50 °C for 48 h in Proteinase K buffer (100 mM EDTA, 10 mM Tris-HCl, 1% (w/v) *N*-Lauroylsarcosine, 0.2% (w/v) Sodium deoxycholate, pH 8.0) supplemented with 1 mg/mL Proteinase K (AppliChem GmbH, Germany). Finally, agar plugs were sequentially washed in four steps, 1 h per wash. After incubation in washing buffer (10 mM Tris-HCl, 50 mM EDTA), plugs were washed in washing buffer supplemented with 1 mM Phenylmethylsulfonyl fluoride, then in washing buffer, and finally in 0.1× concentrated washing buffer.

For restriction digestion with *Pme*I (New England Biolabs, USA), the prepared agar plugs were incubated in 1 mL of restriction enzyme buffer (1× concentrated) for 1 h with gentle agitation at room temperature. Then, the plugs were transferred into 300 μL of fresh enzyme buffer supplemented with *Pme*I (50 units per 100 μL agar plug). Restriction digestions were incubated overnight at 37 °C. After overnight incubation, agar plugs were washed in 1× washing buffer for 1 h. For PFGE analysis, 1/8th of each agar plug was used. PFGE was performed using the Rotaphor System 6.0 (Analytik Jena, Germany) following the manufacturer's instructions. Separation of DNA fragments was achieved using a 0.5% agarose gel (Pulse Field Certified Agarose, Bio-Rad, USA) and 0.5× TBE buffer (44.5 mM Tris-HCl, 44.5 mM boric acid, 1 mM EDTA). The following settings were applied. Step 1, 18 h, 130−100 V (logarithmic decrease), angle: 130°−110° (logarithmic decrease), interval: 50−175 s (logarithmic increase). Step 2, 18 h, 130−80 V (logarithmic decrease), angle: 110°, interval: 175−500 s (logarithmic increase). Step 3, 40 h, 80−50 V (logarithmic decrease), angle: 106°, interval: 500−2000 s (logarithmic increase). Buffer temperature was adjusted to 12 °C.

For whole genome sequencing, a Nextera XT DNA library was constructed[82] and sequenced using the Illumina MiSeq platform which generated 2,504,130 paired-end reads. After trimming, assembly was performed with SPAdes 3.9.0,[83] which produced 399 contigs. Contigs were aligned against the genomes of *S. meliloti* 2011 and BL225C with MeDuSa.[84] A further check with raw sequence reads was performed with QualiMap Software, using default parameters.[85] The assembly has been deposited to the GenBank database under the BioProject ID PRJNA434498.

Finally, several PCR primer pairs for amplification of unique genes of Rm2011 and BL225C (Supplementary Table S3), selected based on a comparative genome analysis with Roary,[86] were routinely used to ensure the correct identification of strains during all experiments.

Growth Curves. Growth curves were initiated by diluting overnight cultures to an OD_{600} of 0.1 in TY medium or in M9 medium supplemented with succinate as a carbon source. Incubation was performed in 150 μL volumes in a 96 well microtiter plate. The microplates were incubated without shaking at 30 °C and growth was measured with a microplate reader (Tecan Infinite 200 PRO, Tecan, Switzerland).

Growth with Root Exudate. The ability to colonize plant roots was tested using growth on root exudate as a

DOI: 10.1021/acssynbio.8b00158
ACS Synth. Biol. 2018, 7, 2365−2378

metabolic proxy for colonization. Root exudate were produced from seedlings of *M. sativa* (cv. Maraviglia), as previously described.[63] Strains were grown on TY plates, following which a single colony was resuspended in 0.9% NaCl solution to a final OD_{600} of 0.5 (1×10^9 CFU/mL). Then, each microplate well was inoculated with 75 μL of either M9 without a carbon source or a nitrogen-free M9 composition with succinate as a carbon source, 20 μL of root exudate, and 5 μL of the culture. The microplates were incubated without shaking' at 30 °C and the growth was measured on a microplate reader (Tecan Infinite 200 PRO, Tecan, Switzerland). At the end of the incubation period, aliquots from each well were diluted and viable titers of *S. meliloti* cells were estimated after incubation on TY plates at 30 °C.

Plant Symbiotic Assays. Symbiotic assays were performed in microcosm conditions in plastic pots containing a 1:1 mixture of sterile vermiculite and perlite, supplemented with 200 mL of Fahraeus N-free liquid plant growth medium[87] *S. meliloti* strains were grown in liquid TY medium at 30 °C for 48 h. Cultures were then washed three times in 0.9% NaCl solution and resuspended to an OD_{600} of 1.0. Approximately 1×10^7 cells were added to each pot, corresponding to $\sim 4 \times 10^4$ cells/cm^3. Washed cell-suspensions were then directly spread over the roots of one-week old seedlings that were directly germinated in the pots, and grown in a growth chamber maintained at 26 °C with a 16 h photoperiod (100 microeinstein/m^2/s) for 5 weeks. Nodule counts were performed after the 5 weeks, then the shoots dried at 50 °C for 7 days. Competition assays and the estimations of number of bacterial genome copies per nodule (determined with qPCR) were done as previously reported.[66] The alfalfa cultivars (*M. sativa*, *M. falcata*, *Medicago x varia*) used and their main features are reported in Supplemental Table S4.

Biofilm Assays. Strains were inoculated in 5 mL of tryptone-yeast extract (TY) medium and in 5 mL of RDM[78] and grown for 24 h with shaking. After growth, cultures were diluted to an OD_{600} of 0.02 in fresh TY and RDM medium, respectively, and 100 μL of the diluted culture was inoculated into a microtiter plate. The plates were incubated at 30 °C for 48 h, after which the OD_{600} was measured to determine the cell biomass. Each well was then stained with 20 μL of crystal violet solution for 10 min. The medium containing the planktonic cells was gently removed and the microtiter plate wells were washed three times with 200 μL of PBS (0.1 M, pH 7.4) buffer and allowed to dry for 15 min. The crystal violet in each well was then solubilized by adding 100 μL of 95% EtOH and incubating for 15 min at room temperature as described in.[88] The plate was then read at 560 nm using a microtiter plate reader (Tecan Infinite 200 PRO, Tecan, Switzerland).

Phenotype Microarray. Phenotype MicroArray experiments using Biolog plates PM1 (carbon sources), PM2A (carbon sources), and PM3 (nitrogen sources) were performed largely as described previously.[45] All bacterial strains used in this study (parental and *cis*-hybrid) are listed in Table 1. Data analysis was performed with DuctApe.[52] Activity index (AV) values were calculated following subtraction of the blank well from the experimental wells. Growth with each compound was evaluated with AV values from 0 (no growth) to 4 (maximal growth), after elbow test calculation (Table S3c,d).

NMR Metabolomics of the Cell Lysates and Media. Overnight cultures were washed, resuspended, and diluted in 100 mL of fresh M9 medium (41 mM Na_2HPO_4, 22 mM KH_2PO_4, 8.6 mM NaCl, 18.7 mM NH_4Cl, 4.1 μM biotin, 42 nM $CoCl_2$, 1 mM $MgSO_4$, 0.25 mM $CaCl_2$, 38 μM $FeCl_3$, 5 μM thiamine-HCl, 10 mM succinate).[16] For cell lysates, when cultures reached an OD_{600} of 1, 50 mL of each culture was pelleted by centrifuging for 25 min at 15 000*g*. For the media, 1 mL of the supernatant of each culture was collected. For cell lysate analysis, each pellet was resuspended in 500 μL of PBS, and sonicated for 20 min with cycles of 1 s of activity and 9 s of rest (292.5 W, 13 mm tip), with contemporary cooling on ice. After cell lysis, the samples were centrifuged for 25 min at 4 °C at 8000*g*. For each strain, four independent experiments were performed. NMR samples were prepared in 5.00 mm NMR tubes (Bruker BioSpin) with 55 μL of a ^2H_2O solution containing 10 mM sodium trimethylsilyl[2,2,3,3-$^2H_4$] propionate (TMSP) and 500 μL of sample.

^{1}H NMR spectra were acquired for both the cell lysates and the growth media. High reproducibility between samples was seen (Supplemental Figure S3), as expected based on previous studies with eukaryotic cells.[89,90] NMR spectra were recorded using a Bruker 900 MHz spectrometer (Bruker BioSpin) equipped with a CP TCI ^{1}H/^{13}C/^{15}N probe. Before measurement, samples were kept for 5 min inside the NMR probe head for temperature equilibration at 300 K. ^{1}H NMR spectra were acquired with water peak suppression and a standard Carr−Purcell−Meiboom−Gill (CPMG) sequence (cpmgpr; Bruker BioSpin srl), using 192 or 256 scans (for cell lysates and growing media, respectively) over a spectral region of 18 kHz, 110 K points, an acquisition time of 3.07 s, and a relaxation delay of 4 s. This pulse sequence[91] was used to impose a T_2 filter that allows the selective observation of small molecular weight components in solutions containing macromolecules.

The raw data were multiplied by a 0.3 Hz exponential line broadening before applying Fourier transformation. Transformed spectra were automatically corrected for phase and baseline distortions and calibrated (chemical shift was referenced to the doublet of alanine at 1.48 ppm for cell lysates, and to the singlet of TMSP at 0.00 ppm for growth media) using TopSpin 3.5 (Bruker BioSpin srl). Multivariate and univariate analyses were performed on the obtained data using R software. For multivariate analysis, each spectrum in the region 10−0.2 ppm was segmented into 0.02 ppm chemical shift bins, and the corresponding spectral areas were integrated using the AMIX software (Bruker BioSpin). Binning is a way to reduce the number of total variables and to compensate for small shifts in the signals, making the analyses more robust and reproducible. The area of each bin was normalized to the total spectral area, calculated with exclusion of the water region (4.50−5.15 ppm), in order to correct the data for possible differences in the cell count of each of the NMR samples.

Unsupervised PCA was used to obtain a preliminary overview of the data (visualization in a reduced space, cluster detection, screening for outliers). CA was used in combination with PCA to increase supervised data reduction and classification. Accuracy, specificity, and sensitivity were estimated according to standard definitions. The global accuracy for classification was assessed by means of a leave-one-out cross-validation scheme. The metabolites, whose peaks in the spectra were well-defined and resolved, were assigned and their levels analyzed. The assignment procedure was performed using an internal NMR spectral library of pure organic compounds, public databases such as the *E. coli* Metabolome Database[92] storing reference NMR spectra of metabolites, and spiking NMR experiments.[93] The relative concentrations of the various metabolites were calculated by integrating the corresponding

DOI: 10.1021/acssynbio.8b00158
ACS Synth. Biol. 2018, 7, 2365−2378

signals in the spectra,[94] using a homemade program for signal deconvolution. The nonparametric Wilcoxon−Mann−Whitney test was used for the determination of the meaningful metabolites: a p-value of 0.05 was considered statistically significant. The molecule 1,4-dioxane was used as a standard to perform the quantitative NMR analysis with the aim of obtaining the absolute concentrations (μM) of the analyzed metabolites.

NMR data were uploaded on the MetaboLights database (www.ebi.ac.uk/metabolights) with the accession number MTBLS576.

Generation of the Metabolic Models. The manually curated iGD1575 reconstruction of *S. meliloti* Rm1021[41] was modified to expand the composition of the biomass reaction through the inclusion of an additional 31 compounds, including vitamins, coenzymes, and ions at trace concentrations as described elsewhere (Table S6).[74] Although iGD1575 is based on *S. meliloti* Rm1021, it is expected to accurately represent Rm2011 metabolism as these two strains are derived from the same field isolate (SU47) and have nearly identical gene contents;[42] while there are numerous SNPs between the strains, SNPs are not considered during the process of metabolic reconstruction.

All other metabolic models were constructed using our recently published protocols for template-assisted metabolic reconstruction and assembly of hybrid bacterial models.[51] Briefly, a draft metabolic reconstructions of *S. meliloti* BL225C was produced using the KBase Web server (www.kbase.us) with gap filling. The draft model was enhanced using the curated Rm1021 model as a template according to,[51] using orthologous gene sets between BL225C and Rm1021 produced with InParanoid.[95] Additionally, an appropriate "protein synthesis" reaction was manually added to the model. Finally, replicon transplantation between the BL225C model and the Rm1021 model was performed as described recently,[51] making use of the InParanoid generated orthology data and the information contained within each model. All metabolic reconstructions used in this work are provided in Supplementary File S1 in COBRA format within a MATLAB MAT-file. The enhancement and transplantation pipeline is available at https://github.com/TVignolini/replicon-swap.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acssynbio.8b00158.

Tables S1, S3−S6; Figures S1−S8 (PDF)

Table S2 (XLSX)

File S1: Supporting Data (ZIP)

■ AUTHOR INFORMATION

Corresponding Authors

*E-mail: marco.fondi@unifi.it.

*E-mail: alessio.mengoni@unifi.it.

ORCID

Anke Becker: 0000-0003-4561-9184

Marco Fondi: 0000-0001-9291-5467

Paola Turano: 0000-0002-7683-8614

Alessio Mengoni: 0000-0002-1265-8251

Author Contributions

A. Checcucci created the strains, performed microbiological analyses. G. diCenzo contributed in metabolic model creation and performed computation analyses on the metabolic modeling. V. Ghini, P. Turano, C. Luchinat performed NMR analyses and contributed in NMR spectra interpretation. V. Ghini contributed in preparing illustrations. A. Becker and J. Döhlemann contributed PFGE analysis and interpretation. T. Vignolini and M. Fondi contributed the first draft of metabolic model and preliminary computational simulations. G. diCenzo performed computational simulations. F. Decorosi and C. Viti contributed in Phenotype Microarray analysis and interpretation. A. Checcucci and C. Fagorzi contributed to *in vitro* symbiotic assays. M. Bazzicalupo and T. Finan provided data interpretation. A. Mengoni, M. Fondi, G. diCenzo, A. Checcucci conceived the work. A. Checcucci, A. Mengoni, V. Ghini, M. Fondi, G. diCenzo prepared the manuscript. All authors have read and approved the manuscript.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We are grateful to Dr. Carla Scotti (CREA-FLC, Lodi, Italy) for kindly providing seeds of *M. sativa* cultivars and to Gabriele Brazzini for assistance in symbiotic assays. NMR spectra were acquired and analyzed at CERM, Core Centre of Instruct-ERIC, an ESFRI Landmark, supported by national member subscriptions. This work was supported by the University of Florence, project "Dinamiche dell'evoluzione dei genomi batterici: l'evoluzione del genoma multipartito e la suddivisione in moduli funzionali", call "PROGETTI STRATEGICI DI ATENEO ANNO 2014" to AM. AC was supported by Fondazione Buzzati-Traverso. GCD was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) through a PDF fellowship, VG was supported by Fondazione Umberto Veronesi. Work in the TMF lab is supported by NSERC. AB and JD were supported by the German Research Foundation (TRR 174).

■ REFERENCES

(1) Lau, Y. H., Stirling, F., Kuo, J., Karrenbelt, M. A., Chan, Y. A., Riesselman, A., and Gibson, D. G. (2017) Large-scale recoding of a bacterial genome by iterative recombineering of synthetic DNA. *Nucleic Acids Res. 45*, 6971.

(2) Smith, H. O., Hutchison, C. A., Pfannkoch, C., and Venter, J. C. (2003) Generating a synthetic genome by whole genome assembly: φX174 bacteriophage from synthetic oligonucleotides. *Proc. Natl. Acad. Sci. U. S. A. 100* (26), 15440−15445.

(3) Burton, R. S., Rawson, P. D., and Edmands, S. (1999) Genetic architecture of physiological phenotypes: empirical evidence for coadapted gene complexes. *Am. Zool. 39* (2), 451−462.

(4) Harrison, P. W., Lower, R. P. J., Kim, N. K. D., and Young, J. P. W. (2010) Introducing the bacterial 'chromid': not a chromosome, not a plasmid. *Trends Microbiol. 18* (4), 141−148.

(5) diCenzo, G. C., and Finan, T. M. (2017) The Divided Bacterial Genome: Structure, Function, and Evolution. *Microbiol. Mol. Biol. Rev. 81* (3), x.

(6) Agnoli, K., Schwager, S., Uehlinger, S., Vergunst, A., Viteri, D. F., Nguyen, D. T., and Eberl, L. (2012) Exposing the third chromosome of *Burkholderia cepacia complex* strains as a virulence plasmid. *Mol. Microbiol. 83* (2), 362−378.

(7) Fei, F., Bowdish, D. M., McCarry, B. E., and Finan, T. M. (2016) Effects of synthetic large-scale genome reduction on metabolism and metabolic preferences in a nutritionally complex environment. *Metabolomics 12* (2), 23.

(8) Chen, H., Higgins, J., Oresnik, I. J., Hynes, M. F., Natera, S., Djordjevic, M. A., and Rolfe, B. G. (2000) Proteome analysis demonstrates complex replicon and luteolin interactions in pSymA-

DOI: 10.1021/acssynbio.8b00158
ACS Synth. Biol. 2018, 7, 2365−2378

cured derivatives of *Sinorhizobium meliloti* strain 2011. *Electrophoresis 21*, 3833−3842.
(9) González, V., Santamaría, R. I., Bustos, P., Hernández-González, I., Medrano-Soto, A., Moreno-Hagelsieb, G., Janga, S. C., Ramírez, M. A., Jiménez-Jacinto, V., and Collado-Vides, J. (2006) The partitioned Rhizobium etli genome: genetic and metabolic redundancy in seven interacting replicons. *Proc. Natl. Acad. Sci. U. S. A. 103* (10), 3834−3839.
(10) Galardini, M., Brilli, M., Spini, G., Rossi, M., Roncaglia, B., Bani, A., Chiancianesi, M., Moretto, M., Engelen, K., Bacci, G., Pini, F., Biondi, E. G., Bazzicalupo, M., and Mengoni, A. (2015) Evolution of Intra-specific Regulatory Networks in a Multipartite Bacterial Genome. *PLoS Comput. Biol. 11* (9), e1004478.
(11) Xu, Q., Dziejman, M., and Mekalanos, J. J. (2003) Exposing the third chromosome of Burkholderia cepacia complex strains as a virulence plasmid. *Proc. Natl. Acad. Sci. U. S. A. 100* (3), 1286−1291.
(12) Yoder-Himes, D. R., Konstantinidis, K. T., and Tiedje, J. M. (2010) Identification of potential therapeutic targets for *Burkholderia cenocepacia* by comparative transcriptomics. *PLoS One 5* (1), e8724.
(13) Ramachandran, V. K., East, A. K., Karunakaran, R., Downie, J. A., and Poole, P. S. (2011) Adaptation of *Rhizobium leguminosarum* to pea, alfalfa and sugar beet rhizospheres investigated by comparative transcriptomics. *Genome Biol. 12*, R106.
(14) López-Guerrero, M. G., Ormeño-Orrillo, E., Acosta, J. L., Mendoza-Vargas, A., Rogel, M. A., Ramírez, M. A., and Martínez-Romero, E. (2012) Rhizobial extrachromosomal replicon variability, stability and expression in natural niches. *Plasmid 68* (3), 149−158.
(15) Young, J. P. W. (2016) Bacteria Are Smartphones and Mobile Genes Are Apps. *Trends Microbiol. 24* (12), 931−932.
(16) diCenzo, G. C., MacLean, A. M., Milunovic, B., Golding, G. B., and T, F. (2014) Examination of Prokaryotic Multipartite Genome Evolution through Experimental Genome Reduction. *PLoS Genet. 10* (10), e1004742.
(17) Checcucci, A., diCenzo, G. C., Bazzicalupo, M., and Mengoni, A. (2017) Trade, diplomacy & warfare. The quest for elite rhizobia inoculant strains. *Front. Microbiol.*, DOI: 10.3389/fmicb.2017.02207.
(18) Ramachandran, R., Jha, J., Paulsson, J., and Chattoraj, D. (2017) Random versus cell cycle-regulated replication initiation in Bacteria: insights from studying *Vibrio cholerae* chromosome 2. *Microbiol. Mol. Biol. Rev.*, DOI: 10.1128/MMBR.00033-16.
(19) Villaseñor, T., Brom, S., Dávalos, A., Lozano, L., Romero, D., and Santos, AG-DL. (2011) Housekeeping genes essential for pantothenate biosynthesis are plasmid-encoded in *Rhizobium etli* and *Rhizobium leguminosarum*. *BMC Microbiol. 11*, 66.
(20) Pini, F., De Nisco, N. J., Ferri, L., Penterman, J., Fioravanti, A., Brilli, M., Mengoni, A., Bazzicalupo, M., Viollier, P. H., Walker, G. C., and Biondi, E. G. (2015) Cell Cycle Control by the Master Regulator CtrA in *Sinorhizobium meliloti*. *PLoS Genet. 11* (5), e1005232.
(21) Heidelberg, J. F., Eisen, J. A., Nelson, W. C., Clayton, R. A., Gwinn, M. L., Dodson, R. J., and Gill, S. R. (2000) DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*. *Nature 406*, 477−483.
(22) Rao, J. R., Fenton, M., and Jarvis, B. D. W. (1994) Symbiotic plasmid transfer in Rhizobium leguminosarum biovar trifolii and competition between the inoculant strain ICMP2163 and trans-conjugant soil bacteria. *Soil Biol. Biochem. 26* (3), 339−351.
(23) Novikova, N., and Safronova, V. (1992) Transconjugants of *Agrobacterium radiobacter* harbouring sym genes of Rhizobium galegae can form an effective symbiosis with *Medicago sativa*. *FEMS Microbiol. Lett. 93* (3), 261−268.
(24) Martínez, E., Palacios, R., and Sanchez, F. (1987) Nitrogen-fixing nodules induced by Agrobacterium tumefaciens harboring *Rhizobium phaseoli* plasmids. *J. Bacteriol. 169* (6), 2828−2834.
(25) Hooykaas, P. J. J., Snijdewint, F. G. M., and Schilperoort, R. A. (1982) Identification of the Sym plasmid of *Rhizobium leguminosarum* strain 1001 and its transfer to and expression in other rhizobia and *Agrobacterium tumefaciens*. *Plasmid 8* (1), 73−82.
(26) Abe, K., Miyake, K., Nakamura, M., Kojima, K., Ferri, S., Ikebukuro, K., and Sode, K. (2014) Engineering of a green-light inducible gene expression system in *Synechocystis* sp. PCC6803. Microbial. *Microb. Biotechnol. 7* (2), 177−183.
(27) Nakatsukasa, H. H., Uchiumi, T., Kucho, K. I., Suzuki, A., Higashi, S., and Abe, M. (2008) Transposon mediation allows a symbiotic plasmid of *Rhizobium leguminosarum* bv. trifolii to become a symbiosis island in *Agrobacterium* and *Rhizobium*. *J. Gen. Appl. Microbiol. 54*, 107−118.
(28) Wong, C. H., Pankhurst, C. E., Kondorosi, A., and Broughton, W. J. (1983) Morphology of root nodules and nodule-like structures formed by *Rhizobium* and *Agrobacterium* strains containing a *Rhizobium meliloti* megaplasmid. *J. Cell Biol. 97*, 787−794.
(29) Hirsch, A. M., Wilson, K. J., Jones, J. D., Bang, M., Walker, V. V., and Ausubel, F. M. (1984) Rhizobium meliloti nodulation genes allow *Agrobacterium tumefaciens* and *Escherichia coli* to form pseudonodules on alfalfa. *J. Bacteriol. 158*, 1133−1143.
(30) Truchet, G., Rosenberg, C., Vasse, J., Julliot, J. S., Camut, S., and Denarie, J. (1984) Transfer of Rhizobium meliloti pSym genes into *Agrobacterium tumefaciens*: host-specific nodulation by atypical infection. *J. Bacteriol. 157* (1), 134−142.
(31) Finan, Kunkel, De Vos, and Signer (1986) Second symbiotic megaplasmid in *Rhizobium meliloti* carrying exopolysaccharide and thiamine synthesis genes. *J. Bacteriol. 167* (1), 66−72.
(32) Hynes, M. F., Simon, R., Müller, P., Niehaus, K., Labes, M., and Pühler, A. (1986) The two megaplasmids of *Rhizobium meliloti* are involved in the effective nodulation of alfalfa. *Molecular and General Genetics MGG 202* (3), 356−362.
(33) Muro-Pastor, A. M., Kuritz, T., Flores, E., Herrero, A., and Wolk, C. P. (1994) Transfer of a genetic marker from a megaplasmid of *Anabaena* sp. strain PCC 7120 to a megaplasmid of a different *Anabaena* strain. *J. Bacteriol. 176* (4), 1093−1098.
(34) Romanchuk, A., Jones, C. D., Karkare, K., Moore, A., Smith, B. A., Jones, C., and Baltrus, D. A. (2014) Bigger is not always better: transmission and fitness burden of ∼ 1MB *Pseudomonas syringae* megaplasmid pMPPla107. *Plasmid 73*, 16−25.
(35) Agnoli, K., Freitag, R., Gomes, M. C., Jenul, C., Suppiger, A., Mannweiler, O., and Eberl, L. (2017) Use of Synthetic Hybrid Strains To Determine the Role of Replicon 3 in Virulence of the *Burkholderia cepacia* Complex. *Appl. Environ. Microbiol.*, DOI: 10.1128/AEM.00461-17.
(36) Galardini, M., Mengoni, A., Brilli, M., Pini, F., Fioravanti, A., Lucas, S., Lapidus, A., Cheng, J. F., Goodwin, L., Pitluck, S., Land, M., Hauser, L., Woyke, T., Mikhailova, N., Ivanova, N., Daligault, H., Bruce, D., Detter, C., Tapia, R., Han, C., Teshima, H., Mocali, S., Bazzicalupo, M., and EG, B. (2011) Exploring the symbiotic pangenome of the nitrogen-fixing bacterium *Sinorhizobium meliloti*. *BMC Genomics 12* (1), 235.
(37) Vance, C. P. (2001) Symbiotic nitrogen fixation and phosphorus acquisition. Plant nutrition in a world of declining renewable resources. *Plant Physiol. 127* (2), 390−397.
(38) Barnett, M. J., Fisher, R. F., Jones, T., Komp, C., Abola, A. P., Barloy-Hubler, F., Bowser, L., Capela, D., Galibert, F., Gouzy, J., Gurjal, M., Hong, A., Huizar, L., Hyman, R. W., Kahn, D., Kahn, M. L., Kalman, S., Keating, D. H., Palm, C., Peck, M. C., Surzycki, R., Wells, D. H., Yeh, K. C., Davis, R. W., Federspiel, N. A., and Long, S. R. (2001) Nucleotide sequence and predicted functions of the entire *Sinorhizobium meliloti* pSymA megaplasmid. *Proc. Natl. Acad. Sci. U. S. A. 98* (17), 9883−8.
(39) Döhlemann, J., Wagner, M., Happel, C., Carrillo, M., Sobetzko, P., Erb, T. J., and Becker, A. (2017) A family of single copy repABC-type shuttle vectors stably maintained in the alpha-proteobacterium *Sinorhizobium meliloti*. *ACS Synth. Biol. 6* (6), 968−984.
(40) Oresnik, Liu, Yost, and Hynes (2000) Megaplasmid pRme2011a of *Sinorhizobium meliloti* is not required for viability. *Journal of bacteriology 182* (12), 3582−3586.
(41) diCenzo, G. C., Zamani, M., Milunovic, B., and Finan, T. M. (2016) Genomic resources for identification of the minimal N2-fixing symbiotic genome. *Environ. Microbiol. 18* (8), 2534−2547.
(42) Sallet, E., Roux, B., Sauviac, L., Jardinaud, M. F. O., Carrère, S., Faraut, T., and Bruand, C. (2013) Next-generation annotation of

DOI: 10.1021/acssynbio.8b00158
ACS Synth. Biol. 2018, 7, 2365−2378

prokaryotic genomes with EuGene-P: application to *Sinorhizobium meliloti* 2011. *DNA Res. 20* (4), 339−354.
(43) Galardini, M., Pini, F., Bazzicalupo, M., Biondi, E. G., and Mengoni, A. (2013) Replicon-Dependent Bacterial Genome Evolution: The Case of *Sinorhizobium meliloti*. *Genome Biol. Evol. 5* (3), 542−558.
(44) Giuntini, E., Mengoni, A., De Filippo, C., Cavalieri, D., Aubin-Horth, N., Landry, C. R., Becker, A., and Bazzicalupo, M. (2005) Large-scale genetic variation of the symbiosis-required megaplasmid pSymA revealed by comparative genomic analysis of *Sinorhizobium meliloti* natural strains. *BMC Genomics 6*, 158.
(45) Biondi, E. G., Tatti, E., Comparini, D., Giuntini, E., Mocali, S., Giovannetti, L., Bazzicalupo, M., Mengoni, A., and Viti, C. (2009) Metabolic capacity of *Sinorhizobium (Ensifer) meliloti* strains as determined by phenotype microarray analysis. *Appl. Environ. Microbiol. 75* (16), 5396−5404.
(46) diCenzo, Checcucci, Bazzicalupo, Mengoni, Viti, Dziewit, Finan, Galardini, and Fondi (2016) Metabolic modelling reveals the specialization of secondary replicons for niche adaptation in *Sinorhizobium meliloti*. *Nat. Commun. 7*, 12219.
(47) Nogales, J., Blanca-Ordóñez, H., Olivares, J., and Sanjuán, J. (2013) Conjugal transfer of the *Sinorhizobium meliloti* 1021 symbiotic plasmid is governed through the concerted action of one- and two-component signal transduction regulators. *Environ. Microbiol. 15*, 811−821.
(48) Pérez-Mendoza, D., Sepúlveda, E., Pando, V., Munoz, S., Nogales, J., Olivares, J., and Sanjuán, J. (2005) Identification of the rctA gene, which is required for repression of conjugative transfer of rhizobial symbiotic megaplasmids. *J. Bacteriol. 187* (21), 7341−7350.
(49) Blanca-Ordóñez, H., Oliva-García, J. J., Pérez-Mendoza, D., Soto, M. J., Olivares, J., Sanjuán, J., and Nogales, J. (2010) pSymA-dependent mobilization of the *Sinorhizobium meliloti* pSymB megaplasmid. *Journal of bacteriology 192* (23), 6309−6312.
(50) Lagendijk, E. L., Validov, S., Lamers, G. E., De Weert, S., and Bloemberg, G. V. (2010) Genetic tools for tagging Gram-negative bacteria with mCherry for visualization in vitro and in natural habitats, biofilm and pathogenicity studies. *FEMS Microbiol. Lett. 305* (1), 81−90.
(51) Vignolini, T., Mengoni, A., and Fondi, M. (2018) Template-assisted metabolic reconstruction and assembly of hybrid bacterial models. In *Metabolic Network Reconstruction and Modeling: Methods and Protocols*, Methods in Molecular Biology Series, pp 177−196, Humana Press, New York, NY.
(52) Galardini, Mengoni, Biondi, Semeraro, Florio, Bazzicalupo, Benedetti, and Mocali (2014) DuctApe: A suite for the analysis and correlation of genomic and OmniLog Phenotype Microarray data. *Genomics 103*, 1.
(53) diCenzo, G. C., Wellappili, D., Golding, G. B., and Finan, T. M. (2018) Inter-replicon Gene Flow Contributes to Transcriptional Integration in the *Sinorhizobium meliloti* Multipartite Genome. *G3: Genes, Genomes, Genet. 8*, g3−300405.
(54) Venturi, V., and Keel, C. (2016) Signaling in the rhizosphere. *Trends Plant Sci. 21* (3), 187−198.
(55) Finan, T. M., Weidner, S., Wong, K., Buhrmester, J., Chain, P., Vorholter, F. J., Hernandez-Lucas, I., Becker, A., Cowie, A., Gouzy, J., Golding, B., and Puhler, A. (2001) The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing endosymbiont *Sinorhizobium meliloti*. *Proc. Natl. Acad. Sci. U. S. A. 98* (17), 9889−94.
(56) Salas, M. E., Lozano, M. J., López, J. L., Draghi, W., Serrania, J., Torres Tejerizo, G. A., and Parisi, G. (2017) Specificity traits consistent with legume-rhizobia coevolution displayed by *Ensifer meliloti* rhizosphere colonization. *Environ. Microbiol. 19*, 3423.
(57) Ramey, B. E., Koutsoudis, M., von Bodman, S. B., and Fuqua, C. (2004) Biofilm formation in plant-microbe associations. *Curr. Opin. Microbiol. 7* (6), 602−609.
(58) Amaya-Gómez, C. V., H, A. M., and Soto, M. J. (2015) Biofilm formation assessment in *Sinorhizobium meliloti* reveals interlinked control with surface motility. *BMC Microbiol. 15*, 58.
(59) Calatrava-Morales, N., McIntosh, M., and Soto, M. J. (2018) Regulation Mediated by N-Acyl Homoserine Lactone Quorum Sensing Signals in the Rhizobium-Legume Symbiosis. *Genes 9* (5), 263.
(60) Fujishige, N. A., Rinauldi, L., Giordano, W., and Hirsch, A. M. (2006) Superficial liaisons: colonization of roots and abiotic surfaces by rhizobia. In *Biology of Plant−Microbe Interactions*, in Proceedings of the 12th International Congress on Molecular Plant−Microbe Interactions (Sánchez, F., Quinto, C., López-Lara, I. M., and Geiger, O., Eds) Vol. *5*, pp 292−299, ISMPMI Press, St. Paul, MN.
(61) Rosenberg, C., Boistard, P., Dénarié, J., and Casse-Delbart, F. (1981) Genes controlling early and late functions in symbiosis are located on a megaplasmid in *Rhizobium meliloti*. *Mol. Gen. Genet. 184* (2), 326−333.
(62) Pobigaylo, N., Szymczak, S., Nattkemper, T. W., and Becker, A. (2008) Identification of genes relevant to symbiosis and competitiveness in *Sinorhizobium meliloti* using signature-tagged mutants. *Mol. Plant-Microbe Interact. 21*, 219−231.
(63) Checcucci, A., Azzarello, E., Bazzicalupo, M., De Carlo, A., Emiliani, G., Mancuso, S., and Mengoni, A. (2017) Role and regulation of ACC deaminase gene in *Sinorhizobium meliloti*: is it a symbiotic, rhizospheric or endophytic gene? *Front. Genet.*, DOI: 10.3389/fgene.2017.00006.
(64) Sprent, J. I. (2001) *Nodulation in Legumes*, Royal Botanic Gardens, London.
(65) Carelli, M., Gnocchi, S., Fancelli, S., Mengoni, A., Paffetti, D., Scotti, C., and Bazzicalupo, M. (2000) Genetic diversity and dynamics of *Sinorhizobium meliloti* populations nodulating different alfalfa varieties in Italian soils. *Appl. Environ. Microbiol. 66*, 4785−4789.
(66) Checcucci, A., Azzarello, E., Bazzicalupo, M., Galardini, M., Lagomarsino, A., Mancuso, S., Marti, L., Marzano, M. C., Mocali, S., Squartini, A., Zanardo, M., and Mengoni, A. (2016) Mixed nodule infection in *Sinorhizobium meliloti - Medicago sativa* symbiosis suggest the presence of cheating behavior. *Front. Plant Sci. 7*, 835.
(67) Poole, P., Ramachandran, V., and Terpolilli, J. (2018) Rhizobia: from saprophytes to endosymbionts. *Nat. Rev. Microbiol. 16*, 291.
(68) Burghardt, L. T., Guhlin, J., Chun, C. L., Liu, J., Sadowsky, M. J., Stupar, R. M., and Tiffin, P. (2017) Transcriptomic basis of genome by genome variation in a legume-hizobia mutualism. *Mol. Ecol. 26*, 6122.
(69) Paffetti, D., Daguin, F., Fancelli, S., Gnocchi, S., Lippi, F., Scotti, C., and Bazzicalupo, M. (1998) Influence of plant genotype on the selection of nodulating Sinorhizobium meliloti strains by *Medicago sativa*. *Antonie van Leeuwenhoek 73* (1), 3−8.
(70) Bobik, C., Meilhoc, E., and Batut, J. (2006) FixJ: a major regulator of the oxygen limitation response and late symbiotic functions of *Sinorhizobium meliloti*. *J. Bacteriol. 188* (13), 4890−902.
(71) Barnett, M. J., Toman, C. J., Fisher, R. F., and Long, S. R. (2004) A dual-genome Symbiosis Chip for coordinate study of signal exchange and development in a prokaryote-host interaction. *Proc. Natl. Acad. Sci. U. S. A. 101* (47), 16636−41.
(72) Chen, Z. J. (2010) Molecular mechanisms of polyploidy and hybrid vigor. *Trends Plant Sci. 15* (2), 57−71.
(73) diCenzo, G. C., and F, T. M. (2015) Genetic redundancy is prevalent within the 6.7 Mb *Sinorhizobium meliloti* genome. *Mol. Genet. Genomics 290* (4), 1345−1356.
(74) diCenzo, G. C., Benedict, A. B., Fondi, M., Walker, G. C., Finan, T. M., Mengoni, A., and Griffitts, J. S. (2017) Robustness encoded across essential and accessory replicons in an ecologically versatile bacterium. *PLoS Genet. 14* (4), e1007357.
(75) Parnell, J. J., Berka, R., Young, H. A., Sturino, J. M., Kang, Y., Barnhart, D. M., and DiLeo, M. V. (2016) From the lab to the farm: an industrial perspective of plant beneficial microorganisms. *Front. Plant Sci. 7*, 1110.
(76) Medini, D., Donati, C., Tettelin, H., Masignani, V., and Rappuoli, R. (2005) The microbial pan-genome. *Curr. Opin. Genet. Dev. 15* (6), 589−594.
(77) Pini, F., Frage, B., Ferri, L., De Nisco, N. J., Mohapatra, S. S., Taddei, L., and Biondi, E. G. (2013) The DivJ, CbrA and PleC system

DOI: 10.1021/acssynbio.8b00158
ACS Synth. Biol. 2018, 7, 2365−2378

controls DivK phosphorylation and symbiosis in *Sinorhizobium meliloti*. *Mol. Microbiol. 90* (1), 54−71.
(78) Vincent, J. M. (1970) *A Manual for the Practical Study of the Root-Nodule Bacteria*, Blackwell Scientific Publications.
(79) Boivin, C., Barran, L. R., Malpica, C. A., and Rosenberg, C. (1991) (1991). Genetic analysis of a region of the *Rhizobium meliloti* pSym plasmid specifying catabolism of trigonelline, a secondary metabolite present in legumes. *J. Bacteriol. 173* (9), 2809−2817.
(80) Herschleb, J., Ananiev, G., and Schwartz, D. C. (2007) Pulsed-field gel electrophoresis. *Nat. Protoc. 2* (3), 677.
(81) Mavingui, P., Flores, M., Guo, X., Dávila, G., Perret, X., Broughton, W. J., and Palacios, R. (2002) Dynamics of genome architecture in *Rhizobium* sp. strain NGR234. *Journal of bacteriology 184* (1), 171−176.
(82) Adessi, A., Spini, G., Presta, L., Mengoni, A., Viti, C., Giovannetti, L., and De Philippis, R. (2016) Draft genome sequence and overview of the purple non sulfur bacterium *Rhodopseudomonas palustris* 42OL. *Stand. Genomic Sci. 11* (1), 24.
(83) Seemann, T. (2014) Prokka: rapid prokaryotic genome annotation. *Bioinformatics 30*, 2068.
(84) Bosi, E., Donati, B., Galardini, M., Brunetti, S., Sagot, M. F., Lió, P., Crescenzi, P., Fani, R., and Fondi, M. (2015) MeDuSa: a multi-draft based scaffolder. *Bioinformatics 31* (15), 2443−2451.
(85) García-Alcalde, F., Okonechnikov, K., Carbonell, J., Cruz, L. M., Götz, S., Tarazona, S., and Conesa, A. (2012) Qualimap: evaluating next-generation sequencing alignment data. *Bioinformatics 28* (20), 2678−2679.
(86) Page, A. J., Cummins, C. A., Hunt, M., Wong, V. K., Reuter, S., Holden, M. T., and Parkhill, J. (2015) Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics 31* (22), 3691−3693.
(87) Fahraeus, G. (1957) The infection of clover root hairs by nodule bacteria studied by a simple glass slide technique. *Microbiology 16* (2), 374−381.
(88) Rinaudi, L. V., and González, J. E. (2009) The low-molecular-weight fraction of exopolysaccharide II from *Sinorhizobium meliloti* is a crucial determinant of biofilm formation. *J. Bacteriol. 191*, 7216−7224.
(89) Bernacchioni, C., Ghini, V., Cencetti, F., Japtok, L., Donati, C., Bruni, P., and Turano, P. (2017) NMR metabolomics highlights sphingosine kinase-1 as a new molecular switch in the orchestration of aberrant metabolic phenotype in cancer cells. *Mol. Oncol. 11* (5), 517−533.
(90) Ghini, V., Di Nunzio, M., Tenori, L., Valli, V., Danesi, F., Capozzi, F., and Bordoni, A. (2017) Evidence of a DHA Signature in the Lipidome and Metabolome of Human Hepatocytes. *Int. J. Mol. Sci. 18* (2), 359.
(91) Carr, H. Y., and Purcell, E. M. (1954) Effects of diffusion on free precession in nuclear magnetic resonance experiments. *Phys. Rev. 94*, 630−638.
(92) Sajed, T., Marcu, A., Ramirez, M., Pon, A., Guo, A. C., Knox, C., and Wishart, D. S. (2016) ECMDB 2.0: A richer resource for understanding the biochemistry of E. coli. *Nucleic Acids Res. 44* (D1), D495−D501.
(93) Psychogios, N., Hau, D. D., Peng, J., et al. (2011) The human serum metabolome. *PLoS One 6*, e16957.
(94) Wishart, D. S. (2008) Quantitative metabolomics using NMR. *TrAC, Trends Anal. Chem. 27*, 228−237.
(95) Sonnhammer, E. L., and Östlund, G. (2015) InParanoid 8: orthology analysis between 273 proteomes, mostly eukaryotic. *Nucleic Acids Res. 43* (D1), D234−D239.

DOI: 10.1021/acssynbio.8b00158
ACS Synth. Biol. 2018, 7, 2365−2378