MPMI Vol. 25, No. 12, 2012, pp. 1523–1530. http://dx.doi.org/10.1094/MPMI-06-12-0148-TA. © 2012 The American Phytopathological Society

e-Xtra*

TECHNICAL ADVANCE

A Draft Genome Sequence of *Nicotiana benthamiana* to Enhance Molecular Plant-Microbe Biology Research

Aureliano Bombarely,[1] Hernan G. Rosli,[1] Julia Vrebalov,[1] Peter Moffett,[1,2] Lukas A. Mueller,[1] and Gregory B. Martin[1,3,4]

[1]Boyce Thompson Institute for Plant Research, Ithaca, NY 14853, USA; [2]Département de Biologie, Université de Sherbrooke, Sherbrooke, Quebec J1K 2R1, Canada; [3]Department of Plant Pathology and Plant-Microbe Biology, Cornell University, Ithaca, NY 14853, USA; [4]Genomics and Biotechnology Section, Department of Biological Sciences, Faculty of Science, King Abdulaziz University, P.O. Box 80203 Jeddah 21589, Saudi Arabia

Submitted 7 June 2012. Accepted 28 July 2012.

***Nicotiana benthamiana* is a widely used model plant species for the study of fundamental questions in molecular plant-microbe interactions and other areas of plant biology. This popularity derives from its well-characterized susceptibility to diverse pathogens and, especially, its amenability to virus-induced gene silencing and transient protein expression methods. Here, we report the generation of a 63-fold coverage draft genome sequence of *N. benthamiana* and its availability on the Sol Genomics Network for both BLAST searches and for downloading to local servers. The estimated genome size of *N. benthamiana* is 3 Gb (gigabases). The current assembly consists of approximately 141,000 scaffolds, spanning 2.6 Gb with 50% of the genome sequence contained within scaffolds >89 kilobases. Of the approximately 16,000 *N. benthamiana* unigenes available in GenBank, >90% are represented in the assembly. The usefulness of the sequence was demonstrated by the retrieval of *N. benthamiana* orthologs for 24 immunity-associated genes from other species including *Ago2*, *Ago7*, *Bak1*, *Bik1*, *Crt1*, *Fls2*, *Pto*, *Prf*, *Rar1*, and mitogen-activated protein kinases. The sequence will also be useful for comparative genomics in the Solanaceae family as shown here by the discovery of microsynteny between *N. benthamiana* and tomato in the region encompassing the *Pto* and *Prf* genes.**

Nicotiana benthamiana has become one of the most commonly used model plant species for research on molecular plant-microbe interactions as well as other areas of plant science (Goodin et al. 2008). The species was originally adopted by plant virologists because of its susceptibility to numerous viral pathogens (Christie and Crawford 1978). This feature led to its use for early studies of virus-induced gene silencing (VIGS) and the underlying mechanisms of RNA interference (Angell and Baulcombe 1997; Baulcombe 1999, 2004; Kumagai et al. 1995). VIGS is now a widely used method to knock down gene expression in *N. benthamiana* because it is simple to implement, is effective, and can be accomplished in about a month (Baulcombe 1999; Burch-Smith et al. 2004; Senthil-Kumar and Mysore 2011; Velasquez et al. 2009). The technique is used for both high-throughput screens to identify immunity-associated genes and for deeper investigation of specific genes implicated in plant-microbe interactions and other aspects of plant biology (Baulcombe 1999; Chakravarthy et al. 2010; del Pozo et al. 2004; Ekengren et al. 2003; Lu et al. 2003; Senthil-Kumar and Mysore 2011; Takahashi et al. 2007; Velasquez et al. 2009). The gene fragment used for VIGS can be designed to target a single gene or several genes at once, such as members of a gene family (Robertson 2004; Senthil-Kumar et al. 2007). However, the lack of a genome sequence for *N. benthamiana* frequently hinders the design of specific VIGS constructs.

N. benthamiana is also susceptible to attack by a large number of well-characterized bacterial, fungal, and oomycete pathogens, as well as certain aphids (Bos et al. 2010; de Jonge et al. 2012; Kamoun et al. 1998; Ramsey et al. 2007; Wei et al. 2007; Yoshino et al. 2012). As one example, the intensively studied tomato pathogen *Pseudomonas syringae* pv. *tomato* has been engineered to allow infection of *N. benthamiana*, permitting investigation of the molecular basis of pathogen-associated molecular pattern (PAMP)- and effector-triggered immunity and the mechanisms employed by type III effectors to promote bacterial virulence (Chakravarthy et al. 2009, 2010; Cunnac et al. 2011; Hann et al. 2007; Kvitko et al. 2009; Nguyen et al. 2010; Wei et al. 2007).

The large leaves of *N. benthamiana* and its susceptibility to pathogens have been exploited as a way to transiently express proteins, using either engineered viruses or syringe-infiltration of *Agrobacterium tumefaciens* (Chapman et al. 1992; Ma et al. 2012; Tang et al. 1996; Van der Hoorn et al. 2000; Wagner et al. 2004). These methods facilitate immunoprecipitation of proteins, visualization of protein subcellular localization, and the examination of post-translational modifications (Cournoyer et al. 2011; de Vries et al. 2006; Inze et al. 2012; Ishihama et al. 2011; Liu et al. 2010; Sacco et al. 2007). *N. benthamiana* has also been used to produce large numbers of proteins for spotting on microarrays (Popescu et al. 2009). Collectively then, the experimental methods available for *N. benthamiana* and its role as a host for many model pathogens have combined to make the species a powerful and increasingly widely used model system.

N. benthamiana is an herbaceous plant indigenous to Australia (Goodin et al. 2008). It is an allotetraploid from the Suaveolentes section, resulting from the hybridization of two

Corresponding author: G. B. Martin; Boyce Thompson Institute for Plant Research; Telephone: +1.607.254.1208; E-mail: gbm7@cornell.edu

*The *e*-Xtra logo stands for "electronic extra" and indicates that three supplementary figures are published online.

unknown progenitors, probably from the Noctiflorae or Sylvestris sections (Chase et al. 2003; Clarkson et al. 2004; Knapp et al. 2004). It is notable that commercial tobacco *N. tabacum* is also an allotetraploid and that one of its progenitors is closely related to *N. sylvestris* from the Sylvestris section (the other parent being closely related to *N. tomentosiformis*). The *N. benthamiana* genome consists of 19 chromosomes and is estimated to be 3 Gb (gigabases) (Goodin et al. 2008; Narayan 1987). As a member of the Solanaceae family, *N. benthamiana* is a close relative of tomato (*Solanum lycopersicum)* and potato *(Solanum tuberosum)*, two economically important species whose genomes have been sequenced recently (Tomato Genome Consortium 2012; Potato Genome Sequencing Consortium 2011). The origins of the *N. benthamiana* accessions used for research are generally not known, although one study indicates they are very similar and may have derived from a single source (Goodin et al. 2008).

Despite being used intensively in research for more than 20 years, there are very limited genomics resources available for *N. benthamiana*. Currently, GenBank holds about 58,000 expressed sequence tags and about 1,400 full-length mRNA sequences for this species. The Sol Genomics Network (SGN) website (Bombarely et al. 2011) provides approximately 16,000 unigenes derived from these data. In the future, it is expected that RNA-Seq will become available to annotate the genome and characterize the transcriptome of this plant species.

We have now generated and sequenced one paired-end (PE) and two mate-pair (MP) libraries of *N. benthamiana* genomic DNA, using Illumina HiSeq2000 technology, resulting in 63-fold coverage of the predicted genome. The assembled draft genome sequence was evaluated for the length of its contigs and scaffolds and for the presence of sequences corresponding to the approximately 16,000 available unigenes. To test its usefulness for plant-microbe interaction research, we searched the genome sequence for a subset of immunity-associated genes known from *Arabidopsis*, potato, tobacco, and tomato. We also demonstrated its utility for comparative genomics by identifying and characterizing microsynteny between *N. benthamiana* and tomato in the region encompassing the *Pto* and *Prf* genes. This draft sequence should enhance the use of *N. benthamiana* by enabling the identification of specific genes and their promoters, comparisons of genome organization in the Solanaceae spp., and the more efficient use of VIGS for gene characterization.

RESULTS

Construction of libraries for sequencing.

The *N. benthamiana* accession that was sequenced is referred to as Nb-1. As with most of the *N. benthamiana* accessions used for research purposes, the origin of this line is unknown (Goodin et al. 2008), although it has been in use in the Martin laboratory since the mid-1990s and is highly inbred. Most *N. benthamiana* accessions in use for research appear to be very similar and might be derived from a single source (Goodin et al. 2008). For this project, we isolated DNA from approximately 25 one-month-old seedlings and developed a PE library with insert size of about 400 bp and two MP libraries with insert sizes of about 2 or 5 kb (additional details below). The MP libraries are important for bridging genome regions with a high level of repetitive sequences, as the insert size in PE libraries is often too small to span such regions. A single run was performed on an Illumina HiSeq 2000 machine with five lanes of the PE library and one lane each of the MP libraries. The sequences derived from the three libraries represented a 63-fold coverage of the estimated *N. benthamiana* genome size (Table 1).

Genome sequence assembly and gene space evaluation.

N. benthamiana reads were preprocessed by using Fastq-mcf and Prinseq (discussed below). Approximately 13% of the reads were discarded during the preprocessing because of low-quality, short, or duplicated reads. The 0.4-kb PE library and the 2- and 5-kb MP libraries had redundancies of 10, 12, and 31%, respectively. Sequence assembly was optimized by using two different programs (ABySS [Simpson et al. 2009] and SOAPdenovo [R. Li et al. 2009]), and three different dataset combinations with kmer 63 to optimize the computational resources. Two assemblies (v0.3 by ABySS and v0.4 by

Table 1. *Nicotiana benthamiana* sequencing, assembly, and gene space evaluation statistics

Sequencing statistics	**Raw data**		**Processed data**	
Library	**Size**	**Coverage**	**Size**	**Coverage**
Paired end 0.4 Kb inserts	181.29 Gb	58 X	160.63 Gb	51 X
Mate pair 2 Kb inserts	46.59 Gb	15 X	39.99 Gb	13 X
Mate pair 5 Kb inserts	40.04 Gb	13 X	28.58 Gb	9 X
Total	267.92 Gb	86 X	229.2 Gb	63 X
Assembly statistics	**Contigs**		**Scaffolds**	
Total assembly size	2.46 Gb		2.63 Gb	
Total assembled sequences	461,463[a]		141,339	
Longest sequence length	208.21 Kb		615.59 Kb	
Average sequence length	5,336 bp		18,610 bp	
N90 index[b]	163,000 sequences		30,261 sequences	
N90 length	3,403 bp		23,201 bp	
N50 index	42,984 sequences		8,897 sequences	
N50 length	16,480 bp		89,778 bp	
Gene space statistics	**Length > 300 bp[c]**		**Coverage > 80%[d]**	
N. benthamiana unigenes	79%		93%	
Solanum lycopersicum gene models	61%		39%	

[a] A total of 485,798 contigs were originally present in the assembly but only contigs >200 bp (461,463) were included in the contig file and in this report. In terms of genome size the 24,335 contigs excluded add approximately 3.72 Mb to the total size.

[b] When ordering all contigs (or scaffolds) by size, the N50 or N90 index indicates the number of the longest sequences (contigs or scaffolds) that contain 50 or 90%, respectively, of the total assembled sequence. The N50 and N90 length indicate the length of the shortest sequence in the set of the largest contigs (or scaffolds) that contain 50 or 90%, respectively, of all the sequences in the assembly.

[c] The percentage of the *N. benthamiana* unigenes (16,024) or tomato gene models (34,739) for which we could identify, in the v0.4.2 draft sequence, a fragment of at least 300 bp.

[d] The percentage of *N. benthamiana* unigenes or tomato gene models that were found, in the v0.4.2 draft sequence, to be represented with a coverage of 80% or more.

SOAPdenovo) were chosen to close gaps by using the SOAP tool GapCloser (Li et al. 2010). The results were evaluated with assembly statistics parameters such as N50 and N90 index and length, total assembly size, and longest sequence for contigs and scaffolds (Table 1). Assembly v0.4.2 had the best assembly statistics and was chosen as the reference for the analysis presented below. A later assembly (v0.4.3) was also generated that filtered out a small number of putative bacterial and chloroplast sequences. All assemblies (v0.3, v0.4.2, and 0.4.3) are available for BLAST searches and downloading from the SGN website (Bombarely et al. 2011).

The representation of known genes in the genome sequence (gene space evaluation) was determined by counting the number of BLAT hits (Kent 2002) for two datasets, *N. benthamiana* unigenes from the SGN unigenes build (version 1) and coding sequence (CDS) regions from *S. lycopersicum* gene models (ITAG 2.3; Tomato Genome Consortium 2012). From these data, we determined the percentage of *N. benthamiana* transcripts and *S. lycopersicum* gene models represented in the genome sequence by a fragment of 300 bp or more (a useful size for VIGS constructs) (Table 1). We also determined the percentage of the *N. benthamiana* unigenes and *S. lycopersicum* gene models that are present with more than 80% coverage in the assembly v0.4.2 (Table 1). In many cases these represented full-length genes.

The possibility that variable regions such as homeologous genes (homologous genes that originated in the genomes of the two progenitors of *N. benthamiana*) or alleles had been collapsed during the assembly was evaluated through read remapping and single nucleotide polymorphism (SNP) calling. Considering only those regions with a minimum coverage of five reads (98.43% of the Niben.v0.4.2 assembly), 164,595 SNP were identified, representing a polymorphism rate of just 0.007%. The low level of SNP is consistent with Nb-1 being a highly inbred line.

Retrieval of immunity-associated gene sequences from the genome sequence.

To assess the usefulness of the draft genome sequence, we searched it for orthologs of 24 genes that have roles in the immune response in various plant species (Table 2). These included genes encoding disease resistance proteins, pattern recognition receptors, signal transduction proteins, and others. If the gene was initially described in a species other than tomato (e.g., *Arabidopsis*, potato) we first identified the ortholog from the tomato genome sequence and, then, used that for BLAST analysis against the *N. benthamiana* scaffold database. In the case of *Pto*, which is not present in *S. lycopersicum*, the *S. pimpinellifolium* sequence was used (GenBank U02271). By using this approach, we successfully identified an *N. benthamiana* sequence corresponding to each of the targeted genes (Table 2). The sequences retrieved had an average DNA sequence identity to the tomato ortholog of 90%. In many cases, the retrieved *N. benthamiana* sequence covered 100% of the open reading frame of the tomato gene model. Based on this analysis, the *N. benthamiana* genome sequence should be useful for gene mining, for the design of VIGS constructs, and for assessing target and nontarget gene silencing.

Notably, except in the case of *Ago2* and *Mpk2*, each gene had at least two matching sequences in the *N. benthamiana* genome sequence, as is expected for an allotetraploid species in which the two homeologous genes have not been collapsed during assembly. In some instances, when it is known that two genes exist in tomato (e.g., *Pti1a* and *Pti1b* [G. Martin, *unpublished*]), we identified four *N. benthamiana* genes, indicating that duplication of this gene likely occurred in a progenitor of *N. benthamiana* and tomato. In the case of *Fls2*, we identified just two genes in *N. benthamiana*, although tomato also has two copies of this gene (Supplementary Fig. S1). It is possible that *Fls2* has undergone a more recent tandem duplication event in tomato, since the two genes are located adjacent to each other in the tomato genome (G. Martin, *unpublished*).

Use of the *N. benthamiana* draft genome sequence for comparative genomics.

The large size of many of the *N. benthamiana* scaffolds should facilitate discovery and characterization of microsynteny with tomato (and other Solanaceous species). For example, we found that each of the two *Prf* genes in *N. benthamiana* is located nearby a *Pto* homolog (scaffolds Niben.v0.4.2.Scf1223 and Niben.v0.4.2.Scf8803) (Fig. 1). In tomato, *Prf* is embedded in a small family of *Pto* homologs with *PtoB* (*Fen*) lying close to but in the opposite orientation to *Prf* (Pedley and Martin 2003). A *Pto*-like gene lies in this same orientation with respect to each of the *Prf* genes in *N. benthamiana*. One of these genes, *Pto1*, was found to be a pseudogene, while the other, *Pto3*, has an intact open reading frame (discussed below). The microsynteny in this region suggests this genome organization is of ancient origin, and phylogenetic analysis reveals some divergence of the *Pto* and *Prf* genes in *S. lycopersicum* (Heinz 1706 and VFNT Cherry), *S. pimpinellifolium*, and *N. benthamiana* (Supplementary Figs. S2 and S3).

DISCUSSION

We have generated a draft genome sequence for *N. benthamiana* that is of sufficient quality and coverage to enhance

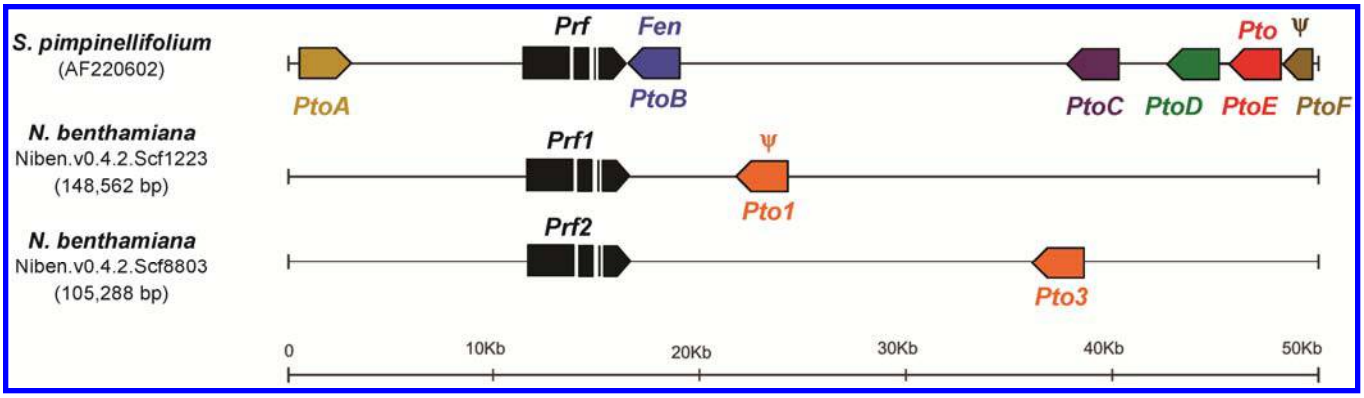

Fig. 1. Genome organization of the *Pto* and *Prf* region in *Solanum pimpinellifolium* (GenBank number AF220602) and *Nicotiana benthamiana* (scaffolds Niben.genome.v0.4.2.Scf1223 and Niben.genome.v0.4.2.Scf8803). The location and orientation of the *Pto* and *Prf* sequences are shown, with gaps in *Prf* indicating introns. *Pto1* and *PtoF* are pseudogenes (ψ).

the design and evaluation of constructs for VIGS, to facilitate retrieval of gene sequences, and for the study of genome organization and synteny between solanaceous species. The sequence is accessible for BLAST searches on the SGN and for downloading to a local server. The first assembly (v0.3) was made available in early February 2012, and the sequence has been downloaded more than 500 times since then, reflecting the widespread use of *N. benthamiana* as a model system for plant biology research. In the future, as the current assembly is further improved and as RNA-seq data become available, we plan to use the tomato gene models (Tomato Genome Consortium 2012) and RNA-seq data from *N. benthamiana* to annotate the gene space of the *N. benthamiana* genome sequence.

Table 2. BLAST hits for selected immunity-associated genes in the *Nicotiana benthamiana* genome

Gene[a]	GenBank accession no.	*S. lycopersicum* locus[b]	*N. benthamiana* v0.4.2 scaffold	Identity (%)[c]	Coverage (%)[d]
SlAdi3	AY849914	Solyc01g090240	Niben.v0.4.2.Scf35729	91	100
			Niben.v0.4.2.Ctg25452543	90	48
StAdr1	HQ906887	Solyc04g079420	Niben.v0.4.2.Scf36756	90	100
			Niben.v0.4.2.Scf41414	90	46
AtAgo2	NM_102866	Solyc02g069260	Niben.v0.4.2.Scf6644	86	86
AtAgo7	NM_105611	Solyc01g010970	Niben.v0.4.2.Scf4610	90	91
			Niben.v0.4.2.Scf36196	90	91
SlBak1	AK328403	Solyc10g047140	Niben.v0.4.2.Scf44412	94	100
			Niben.v0.4.2.Scf3411	94	95
			Niben.v0.4.2.Scf7356	89	95
			Niben.v0.4.2.Scf4596	88	95
AtBik1	NM_129522	Solyc10g084770	Niben.v0.4.2.Scf24800	91	100
			Niben.v0.4.2.Scf13273	90	95[e]
			Niben.v0.4.2.Scf29137	87	90
			Niben.v0.4.2.Scf9691	87	90[e]
SlBti9 (CERK1)	HM208130	Solyc07g049180	Niben.v0.4.2.Scf14676	88	100
			Niben.v0.4.2.Scf37790	87	100[e]
SlCoi1	NM_001247535	Solyc05g052620	Niben.v0.4.2.Scf18019	90	97
			Niben.v0.4.2.Scf15197	90	97
NbCrt1	GQ855284	Solyc02g084700	Niben.v0.4.2.Scf12767	91	96
			Niben.v0.4.2.Scf50466	91	97
SlEds1	AK247875	Solyc06g071280	Niben.v0.4.2.Scf1627	87	100[e]
			Niben.v0.4.2.Scf15551	87	100
AtFls2	NM_124003	Solyc02g070890	Niben.v0.4.2.Scf14580[f]	90	100
			Niben.v0.4.2.Scf5377[f]	90	100
SlMAPKKKα	AY500156	Solyc11g006000	Niben.v0.4.2.Scf1155	91	100
			Niben.v0.4.2.Scf19856	91	100
SlMAPKKKε	GU192457	Solyc11g033270	Niben.v0.4.2.Scf41747	93	100[e]
			Niben.v0.4.2.Scf47703	94	98
SlMKK1	AY691330	Solyc12g009020	Niben.v0.4.2.Scf2991	93	100
			Niben.v0.4.2.Scf10086	92	100
SlMKK2	AY691331	Solyc03g123800	Niben.v0.4.2.Scf6609	85	77
			Niben.v0.4.2.Scf12713	85	77
SlMPK1	AY261512	Solyc12g019460	Niben.v0.4.2.Scf60107	94	58
			Niben.v0.4.2.Scf3284	93	100
SlMPK2	NM_001247426	Solyc08g014420	Niben.v0.4.2.Scf36924	94	100
SlMPK3	NM_001247431	Solyc06g005170	Niben.v0.4.2.Scf16165	94	42
			Niben.v0.4.2.Scf55601	90	98
SlNpr1	AY640378	Solyc07g040690	Niben.v0.4.2.Scf20727	92	84[e]
			Niben.v0.4.2.Scf39188	91	96[e]
SlPti1	U28007	Solyc12g098980	Niben.v0.4.2.Scf13765	94	100
			Niben.v0.4.2.Scf5402	93	100
			Niben.v0.4.2.Scf61958	91	98
			Niben.v0.4.2.Scf25727	89	98
SlPto/Fen	U59318 (PtoB/Fen)	Solyc05g013290	Niben.v0.4.2.Scf1223[f]	86	100[e]
	U02271 (PtoE)	Not present	Niben.v0.4.2.Scf8803[f]	86	100
			Niben.v0.4.2.Scf23435[f]	86	100
			Niben.v0.4.2.Scf6911[f]	86	100
SlPrf	U65391	Solyc05g013280	Niben.v0.4.2.Scf1223[f]	88	100
			Niben.v0.4.2.Scf8803[f]	87	100
NtRar1	AF480487	Solyc11g072290	Niben.v0.4.2.Scf13235	90	96
			Niben.v0.4.2.Scf11152	90	98
SlZim1	NM_001247954	Solyc12g009220	Niben.v0.4.2.Scf1700	91	100
			Niben.v0.4.2.Scf44037	90	100

[a] Species gene from which the indicated Genbank accession gene derives: At = *Arabidopsis thaliana*, Sl = *Solanum lycopersicum*, St = *S. tuberosum*, Nb = *N. benthamiana*, Nt = *N. tabacum*.

[b] The homologous sequence in the tomato reference genome (Annotation ITAG2.3) for the GenBank accession sequence.

[c] The overall identity percentage between the gene model in the tomato reference genome and the *N. benthamiana* BLAST hit (Niben.genome.v0.4.2).

[d] The percentage of the tomato gene model length covered by the *N. benthamiana* BLAST hit.

[e] The predicted gene model for these sequences contains an early stop codon or an indel that changes the open reading frame.

[f] The *N. benthamiana* coding sequences were extracted from the scaffolds based on the alignments with the tomato proteins and renamed as: NibenFLS2.1 (Niben.v0.4.2.Scf14580.Ctg13:1896..6031), NibenFLS2.2 (Niben.v0.4.2.Scf5377.Ctg3:2505..6686), NibenPto1 (Niben.v0.4.2.Scf1223:77225..78148), NibenPto2 (Niben.v0.4.2.Scf6911:172166..173126), NibenPto3 (Niben.v0.4.2.Scf8803:24032..25234), NibenPto4 (Niben.v0.4.2.Scf23435:12062..13168), NibenPto5 (Niben.v0.4.2.Scf23435:1561..2765), NibenPrf1 (Niben.v0.4.2.Scf1223:83466..88954) and NibenPrf2 (Niben.v0.4.2.Scf8803:41183..47018).

Our analysis of the genome sequence and initial effort to retrieve orthologs of 24 immunity-associated *N. benthamiana* genes was informative for several reasons. First, it indicated we could identify a 300-bp or larger fragment for 80% of targeted *N. benthamiana* genes. This is a typical fragment size used for VIGS, and so the draft sequence should be useful for designing gene-silencing constructs. Second, our analysis revealed that the nucleotide sequence identity of the 24 tomato genes to their *N. benthamiana* orthologs was high (90%). With this level of conservation, it is very likely that a tomato gene fragment will successfully silence its ortholog in *N. benthamiana*. Our experience and that of others over the past 10 years of using tomato fragments for VIGS in *N. benthamiana* supports this prediction (Chakravarthy et al. 2010; Senthil-Kumar et al. 2007). Third, the retrieval, in many cases, of full-length genes means this draft sequence will facilitate the cloning of genes from *N. benthamiana* for functional characterization. Despite these encouraging results, it is difficult to infer an objective quality level for the draft genome sequence, due to the lack of a reference sequence, BAC-derived sequence data, or a genetic map. Nevertheless, although we have not yet directly assessed these purposes, it seems likely that the draft sequence will also enable retrieval of sufficient upstream sequence in many cases to allow characterization of promoter regulatory elements and may also be useful for gene identification subsequent to proteomics experiments.

The large average scaffold size of 18 kb (and an N50 length of approximately 90 kb) will also facilitate use of the sequence for comparisons of genome organization with other solanaceous species and, perhaps, lead to new hypotheses about the evolution of these species. For example, it previously had not been reported that *Prf* lies in close proximity to a *Pto*-like gene in any *Nicotiana* species. In fact, both of the *Prf* homeologs share this genome organization, suggesting that a *Pto*-like gene and *Prf* were closely linked before the interspecific hybridization event that created *N. benthamiana*.

In tomato, the Prf nucleotide-binding leucine-rich repeat protein acts in concert with the Pto and Fen kinases to recognize the effector proteins AvrPto and AvrPtoB, thereby conferring resistance to *Pseudomonas syringae* pv. *tomato*. *N. benthamiana* also recognizes AvrPto and certain forms of AvrPtoB in a Prf-dependent manner that has been predicted to involve a Fen- or Pto-like protein (Abramovitch et al. 2003; He et al. 2004; Yeam et al. 2010). The identification in *N. benthamiana* of four *Pto*-like genes with intact open reading frames will now enable testing of which of these genes may be involved in recognition of these type III effectors. The microsynteny that occurs in this region supports the idea that *Prf* and *Pto*-like genes have a long coevolutionary history due to the interaction of their proteins. We can also now surmise that the evolution of the five-member *Pto* gene family occurred subsequent to tomato–*N. benthamiana* speciation, likely in response to pathogen pressure on a progenitor of tomato.

In the future, the *N. benthamiana* draft genome sequence will be improved by the integration of additional genome sequence data as well as information from RNA-seq experiments. Since the release of the first assembly, we have been contacted by several colleagues offering such information, and we plan to incorporate these data and provide them to the community on the SGN website. Our lab is currently generating libraries for RNA-seq analysis from flowers, roots, stems, germinating seeds, and other tissues of *N. benthamiana*, as well as leaves exposed to diverse PAMPs, various strains of *P. syringae* pv. *tomato*, and other pathogens. These data, combined with that from other labs, will permit gene annotation and further enhance the use of *N. benthamiana* as a model for plant-microbe biology research and other areas of plant science.

MATERIALS AND METHODS

Isolation of high molecular weight DNA from *Nicotiana benthamiana*.

Approximately 25 seedlings of accession Nb-1 were grown in a greenhouse under standard conditions (22°C, 14 h light, 10 h dark). One-month-old seedlings were transferred to a dark growth chamber for 3 days to reduce their starch content, and young expanding leaves were collected for nuclei isolation as described in Zhang and associates (1995). Isolated nuclei were resuspended in 5 ml of 1× homogenization buffer (HB) (Zhang et al. 1995) and were lysed with 1 ml of 10% Sarkosyl. The DNA was treated with Proteinase K (10 µg/ml) and Ribonuclease A (10 µg/ml), followed by one phenol-chloroform (1:1) and one chloroform-octanol (24:1) extraction. A cesium chloride (CsCl) gradient was set up using 1 g of CsCl per 1 ml of 1× HB buffer and 1 µl/ml of 5 mg/ml ethidium bromide. The gradient was spun for 16 h in an ultracentrifuge (Optima TLX, Beckman, Irvine, CA, U.S.A.). The ethidium bromide–stained genomic DNA band was extracted with an 18G needle and was dialyzed for 24 h in TE buffer (10 mM Tris, 1 mM EDTA). Seeds of Nb-1 are available from G. Martin.

Sequencing of the *N. benthamiana* genome, sequence assembly, and sequence retrieval.

The PE and MP libraries were generated using standard protocols available online (Illumina chemistry version 2) and sequenced on an Illumina HiSeq 2000 unit at the Cornell-Weill Genomics Core Facility (New York).

N. benthamiana sequences were processed with Fastq-mcf ea-utils (ea-utils website) with a minimum quality score of 30 (Q30) and minimum read length of 50 bp (L50). Read duplication was analyzed using Prinseq (Schmieder and Edwards 2011). Duplicated reads were removed from the fastq files using a Perl script, FastqPairing. A total of 229 Gb of sequence (approximately 63× total genome coverage) was assembled using ABySS (Simpson et al. 2009) with a 63 kmer by using a 48 cores, 512 GB RAM Linux server (assembly version, Niben.v0.3). Subsequently, SOAPdenovo v.1.04 (R. Li et al. 2009) with a 63 kmer was used to assemble a subset of these sequences containing the maximum number of MP reads (approximately 49× total genome coverage), to optimize the server memory resources. The raw sequence reads (214 Gb) are available on a disk by request. The result was used as template for GapCloser v1.12 (Li et al. 2010), using, in this case, all the reads available after a sequence correction process (assembly version, Niben.v0.4.2; this is the version used in the analysis shown in Tables 1 and 2). The contig file for this version was created by splitting the scaffolds by gaps with five or more Ns. Contigs of less than 200 bp were removed from the file (24,335 sequences). Sequence statistics were calculated using a Perl script, FastaSeqStats.

We also performed a sequence homology search by BLAST with the *N. sylvestris* chloroplast sequence. Based on this screen, a total of 205 scaffolds with a hit length >90% of the scaffold length were removed. To detect bacterial contamination, we performed a BLAST search with the National Center for Biotechnology Information (NCBI) bacterial genome sequences dataset. From this screen, a total of 271 scaffolds with a hit length >200 bp were removed as possible contaminations. This filtered version is referred to as Niben.genome.v0.4.3. We have not removed scaffolds with sequences longer than 200 bp from any of the assemblies, as we believe they may contain sequences that could prove useful in the design of VIGS constructs (which may be as small as 300 bp). All assembly versions are available at the SGN *Nicotiana benthamiana* ftp site. The NCBI BioProject accession for the genome sequence is PRJNA170566.

Gene-space evaluation was performed by a sequence homology search using the 16,024 *N. benthamiana* SGN unigenes v.1 and the 34,739 *S. lycopersicum* CDS (Tomato Genome Consortium 2012) (ITAG.2.3). For this purpose, BLAT (Kent 2002) was used with the minimum identity set at 97% (for *N. benthamiana* sequences) and 60% (for tomato sequences). To accelerate the search a multithread Perl script was used, i.e., MultiThreadBlat. Hit filtering and counting were performed with simple combination of Linux commands (grep, awk, and wc). All Perl scripts are available at the GitHub repository. Read remapping was performed using BWA (Li and Durbin 2010), and the SNP calling was performed using Samtools (mpileup) and Bcftools (H. Li et al. 2009).

Retrieval and gene modeling for immunity-associated genes.

The search for immunity-associated genes in the *N. benthamiana* sequence assembly (v0.4.2) was performed with complete full-length CDS from the *S. lycopersicum* genome (ITAG2.3) (Tomato Genome Consortium 2012), with the exception of *Pto*, for which the corresponding *S. pimpinellifolium* sequence was used (GenBank U02271). The sequences were used as query in a BLASTN search (Altschul et al. 1997) with Niben.genome.v0.4.2.scaffolds as the database. For sequences with positive matches, the database sequence hit was retrieved with its surrounding 5,000 bp. A simple gene model was created by aligning these sequences with the corresponding tomato protein sequence, using Exonerate (genome2protein option) (Slater and Birney 2005).

Phylogenetic analysis of *Fls2*, *Pto*, and *Prf*.

MEGA v5.05 (Tamura et al. 2011) was used to perform the phylogenetic analysis of the *Fls2*, *Pto*, and *Prf* sequences. The sequences were aligned using Clustal W (Thompson et al. 1994) with the default parameters. A maximum likelihood tree was constructed for each sequence group assaying the robustness of the tree with 1,000 bootstrapping replicates. The Tamura-Nei model was used for the DNA sequence trees.

Cloning and sequencing of *Pto1* and *Pto3*.

Based on assembly v0.4.2. *Pto1* had a deletion at position 280 that shifts the reading frame, leading to a truncated and probably nonfunctional protein. Sequencing of independent polymerase chain reaction (PCR) products using genomic DNA as the template confirmed the presence of this deletion (details below). The gene model for *Pto3*, based on assembly version 0.4.2, contained an 83-bp intron. This was unexpected, as an intron has never been observed in any *Pto* gene family member in tomato or its wild relatives (Jia et al. 1997; Riely et al. 2001) (G. Martin, *unpublished*). PCR products with identical sequences were amplified from cDNA or genomic DNA, indicating that the *Pto3* gene has an intact open reading frame and does not contain an intron (details below). This particular region of the genome sequence was found to have a low coverage, which might have caused this error in the assembly.

To characterize *Pto1* and *Pto3*, genomic DNA and RNA were isolated from fully expanded *N. benthamiana* leaves and cDNA was synthesized, using a SuperScript III kit (Invitrogen, Carlsbad, CA, U.S.A.). The following primers spanning internal portions of *Pto1* and *Pto3* were generated: Pto1F 5′-GGG AACAACTTCCATAAGTG-3′, Pto1Ra 5′-TGATCAAGCTCA AATACTCC-3′, Pto1Rb 5′-GTACCCACGAGTTCCAATCAC CG-3′, Pto3F 5′-CAGGAAATAGCCAGGAGCAG-3′, Pto3R 5′-CGTATGATGACGATCCACGA-3′, Pto3intF 5′-TGGAGA ATGGGAACCTCAAG-3′ and Pto3intR 5′-GATCAAGCTCA AGCCCTGTC-3′. PCR products from independent PCR reactions using GoTaq (Promega, Fitchburg, WI, U.S.A.) were cloned into pCR8/GW/TOPO TA vector (Invitrogen) and were sequenced.

ACKNOWLEDGMENTS

This work was supported by National Science Foundation grants IOS-1025642 (G. B. Martin, A. Collmer) and IOB-0343327 (P. Moffett) and by start-up funds (L. A. Mueller) and Innovation Funds (G. B. Martin) from the Boyce Thompson Institute for Plant Research. A. Bombarely assembled and analyzed the genome sequence and contributed to the writing; J. Vrebalov prepared the DNA; H. G. Rosli analyzed the genome sequence and contributed to the writing; P. Moffett, L. A. Mueller, and G. B. Martin designed and coordinated the research and wrote the paper.

LITERATURE CITED

Abramovitch, R. B., Kim, Y.-J., Chen, S., Dickman, M. B., and Martin, G. B. 2003. *Pseudomonas* type III effector AvrPtoB induces plant disease susceptibility by inhibition of host programmed cell death. EMBO (Eur. Mol. Biol. Organ.) J. 22:60-69.

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. 1997. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. Nucleic Acids Res. 25:3389-3402.

Angell, S. M., and Baulcombe, D. C. 1997. Consistent gene silencing in transgenic plants expressing a replicating potato virus X RNA. EMBO (Eur. Mol. Biol. Organ.) J. 16:3675-684.

Baulcombe, D. C. 1999. Fast forward genetics based on virus-induced gene silencing. Curr. Opin. Plant Biol. 2:109-113.

Baulcombe, D. 2004. RNA silencing in plants. Nature 431:356-63.

Bombarely, A., Menda, N., Tecle, I. Y., Buels, R. M., Strickler, S., Fischer-York, T., Pujar, A., Leto, J., Gosselin, J., and Mueller, L. A. 2011. The Sol Genomics Network (solgenomics.net): Growing tomatoes using Perl. Nucleic Acids Res. 39:D1149-1155.

Bos, J. I., Prince, D., Pitino, M., Maffei, M. E., Win, J., and Hogenhout, S. A. 2010. A functional genomics approach identifies candidate effectors from the aphid species *Myzus persicae* (green peach aphid). PLoS Genet. 6:e1001216. Published online.

Burch-Smith, T. M., Anderson, J. C., Martin, G. B., and Dinesh-Kumar, S. P. 2004. Applications and advantages of virus-induced gene silencing for gene function studies in plants. Plant J. 39:734-746.

Chakravarthy, S., Velasquez, A. C., and Martin, G. B. 2009. Assay for pathogen-associated molecular pattern (PAMP)-triggered immunity (PTI) in plants. JoVE 31:442. Published online.

Chakravarthy, S., Velasquez, A. C., Ekengren, S. K., Collmer, A., and Martin, G. B. 2010. Identification of *Nicotiana benthamiana* genes involved in pathogen-associated molecular pattern-triggered immunity. Mol. Plant Microbe Interact. 23:715-726.

Chapman, S., Kavanagh, T., and Baulcombe, D. 1992. Potato virus X as a vector for gene expression in plants. Plant J. 2:549-557.

Chase, M. W., Knapp, S., Cox, A. V., Clarkson, J. J., Butsko, Y., Joseph, J., Savolainen, V., and Parokonny, A. S. 2003. Molecular systematics, GISH and origin of hybrid taxa in Nicotiana (Solanaceae). Ann. Bot. 96:107-127.

Christie, S., and Crawford, W. 1978. Plant virus range of *Nicotiana benthamiana*. Plant Dis. Rep. 62:20-22.

Clarkson, J. J., Knapp, S., Garcia, V. F., Olmstead, R. G., Leitch, A. R., and Chase, M. W. 2004. Phylogenetic relationships in Nicotiana (Solanaceae) inferred from multiple plastid DNA regions. Mol. Phylogenet. Evol. 33:75-90.

Cournoyer, P., and Dinesh-Kumar, S. P. 2011. Studying NB-LRR immune receptor localization by agroinfiltration transient expression. Methods Mol. Biol. 712:1-8.

Cunnac, S., Chakravarthy, S., Kvitko, B. H., Russell, A. B., Martin, G. B., and Collmer, A. 2011. Genetic disassembly and combinatorial reassembly identify a minimal functional repertoire of type III effectors in *Pseudomonas syringae*. Proc. Natl. Acad. Sci. U.S.A. 108:2975-2980.

de Jonge, R., van Esse, H. P., Maruthachalam, K., Bolton, M. D., Santhanam, P., Saber, M. K., Zhang, Z., Usami, T., Lievens, B., Subbarao, K. V., and Thomma, B. P. 2012. Tomato immune receptor Ve1 recognizes effector of multiple fungal pathogens uncovered by genome and RNA sequencing. Proc. Natl. Acad. Sci. U.S.A. 109:5110-5115.

de Vries, J. S., Andriotis, V. M., Wu, A. J., and Rathjen, J. P. 2006. Tomato Pto encodes a functional N-myristoylation motif that is required for signal transduction in *Nicotiana benthamiana*. Plant J. 45:31-45.

del Pozo, O., Pedley, K. F., and Martin, G. B. 2004. MAPKKKα is a positive regulator of cell death associated with both plant immunity and disease. EMBO (Eur. Mol. Biol. Organ.) J. 23:3072-3082.

Ekengren, S. K., Liu, Y., Schiff, M., Dinesh-Kumar, S. P., and Martin, G. B. 2003. Two MAPK cascades, NPR1, and TGA transcription factors play a role in Pto-mediated disease resistance in tomato. Plant J. 36:905-917.

Goodin, M. M., Zaitlin, D., Naidu, R. A., and Lommel, S. A. 2008. *Nicotiana benthamiana*: Its history and future as a model for plant-pathogen interactions. Mol. Plant-Microbe Interact. 21:1015-1026.

Hann, D. R., and Rathjen, J. P. 2007. Early events in the pathogenicity of *Pseudomonas syringae* on *Nicotiana benthamiana*. Plant J. 49:607-618.

He, X., Anderson, J. C., Pozo Od, O., Gu, Y. Q., Tang, X., and Martin, G. B. 2004. Silencing of subfamily I of protein phosphatase 2A catalytic subunits results in activation of plant defense responses and localized cell death. Plant J. 38:563-577.

Inze, A., Vanderauwera, S., Hoeberichts, F. A., Vandorpe, M., Van Gaever, T., and Van Breusegem, F. 2012. A subcellular localization compendium of hydrogen peroxide-induced proteins. Plant Cell Environ. 35:308-320.

Ishihama, N., Yamada, R., Yoshioka, M., Katou, S., and Yoshioka, H. 2011. Phosphorylation of the *Nicotiana benthamiana* WRKY8 transcription factor by MAPK functions in the defense response. Plant Cell 23:1153-1170.

Jia, Y., Loh, Y.-T., Zhou, J., and Martin, G. B. 1997. Alleles of *Pto* and *Fen* occur in bacterial speck-susceptible and fenthion-insensitive tomato and encode active protein kinases. Plant Cell 9:61-73.

Kamoun, S., van West, P., Vleeshouwers, V. G., de Groot, K. E., and Govers, F. 1998. Resistance of *Nicotiana benthamiana* to *Phytophthora infestans* is mediated by the recognition of the elicitor protein INF1. Plant Cell 10:1413-1426.

Kent, W. J. 2002. BLAT--the BLAST-like alignment tool. Genome Res. 12:656-664.

Knapp, S., Bohs, L., Nee, M., and Spooner, D. M. 2004. Solanaceae—A model for linking genomics with biodiversity. Comp. Funct. Genomics 5:285-291.

Kumagai, M. H., Donson, J., Della-Cioppa, G., Harvey, D., Hanley, K., and Grill, L. K. 1995. Cytoplasmic inhibition of carotenoid biosynthesis with virus-derived RNA. Proc. Natl. Acad. Sci. U.S.A. 92:1679-1683.

Kvitko, B. H., Park, D. H., Velasquez, A. C., Wei, C. F., Russell, A. B., Martin, G. B., Schneider, D. J., and Collmer, A. 2009. Deletions in the repertoire of *Pseudomonas syringae* pv. *tomato* DC3000 type III secretion effector genes reveal functional overlap among effectors. PLoS Pathog 5:e1000388. Published online.

Li, H., and Durbin, R. 2010. Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics 26:589-595.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R. 2009. The sequence alignment/map format and SAMtools. Bioinformatics 25:2078-2079.

Li, R., Yu, C., Li, Y., Lam, T. W., Yiu, S. M., Kristiansen, K., and Wang, J. 2009. SOAP2: An improved ultrafast tool for short read alignment. Bioinformatics 25:1966-1967.

Li, R., Fan, W., Tian, G., Zhu196, H., He, L., Cai, J., Huang, Q., Cai, Q., Li, B., Bai, Y., Zhang, Z., Zhang, Y., Wang, W., Li, J., Wei, F., Li, H., Jian, M., Li, J., Zhang, Z., Nielsen, R., Li, D., Gu, W., Yang, Z., Xuan, Z., Ryder, O. A., Leung, F. C., Zhou, Y., Cao, J., Sun, X., Fu, Y., Fang, X., Guo, X., Wang, B., Hou, R., Shen, F., Mu, B., Ni, P., Lin, R., Qian, W., Wang, G., Yu, C., Nie, W., Wang, J., Wu, Z., Liang, H., Min, J., Wu, Q., Cheng, S., Ruan, J., Wang, M., Shi, Z., Wen, M., Liu, B., Ren, X., Zheng, H., Dong, D., Cook, K., Shan, G., Zhang, H., Kosiol, C., Xie, X., Lu, Z., Zheng, H., Li, Y., Steiner, C. C., Lam, T. T., Lin, S., Zhang, Q., Li, G., Tian, J., Gong, T., Liu, H., Zhang, D., Fang, L., Ye, C.,, Zhang, J., Hu, W., Xu, A., Ren, Y., Zhang, G., Bruford, M. W., Li, Q., Ma, L., Guo, Y., An, N., Hu, Y., Zheng, Y., Shi, Y., Li, Z., Liu, Q., Chen, Y., Zhao, J., Qu, N., Zhao. S., Tian, F., Wang, X., Wang, H., Xu, L., Liu, X., Vinar, T., Wang, Y., Lam, T. W., Yiu, S. M., Liu, S., Zhang, H., Li, D., Huang, Y., Wang, X., Yang, G., Jiang, Z., Wang, J., Qin, N., Li, L., Li, J., Bolund, L., Kristiansen, K., Wong, G. K., Olson, M., Zhang, X., Li, S., Yang, H., Wang, J., and Wang, J.2010. The sequence and de novo assembly of the giant panda genome. Nature 463:311-317.

Liu, L., Zhang, Y., Tang, S., Zhao, Q., Zhang, Z., Zhang, H., Dong, L., Guo, H., and Xie, Q. 2010. An efficient system to detect protein ubiquitination by agroinfiltration in *Nicotiana benthamiana*. Plant J. 61:893-903.

Lu, R., Malcuit, I., Moffett, P., Ruiz, M. T., Peart, J., Wu, A. J., Rathjen, J. P., Bendahmane, A., Day, L., and Baulcombe, D. C. 2003. High throughput virus-induced gene silencing implicates heat shock protein 90 in plant disease resistance. EMBO (Eur. Mol. Biol. Organ.) J. 22:5690-699.

Ma, L., Lukasik, E., Gawehns, F., and Takken, F. L. 2012. The use of agroinfiltration for transient expression of plant resistance and fungal effector proteins in *Nicotiana benthamiana* leaves. Methods Mol. Biol. 835:61-74.

Narayan, R. K. J. 1987. Nuclear DNA changes, genome differentiation and evolution in Nicotiana (Solanaceae). Plant System Evol. 157:161-180.

Nguyen, H. P., Chakravarthy, S., Velásquez, A. C., McLane, H. L., Zeng, L., Nakayashiki, H., Park, D.-H., Collmer, A., and Martin, G. B. 2010. Methods to study PAMP-triggered immunity using tomato and *Nicotiana benthamiana*. Mol. Plant-Microbe Interact. 23:991-999.

Pedley, K. F., and Martin, G. B. 2003. Molecular basis of Pto-mediated resistance to bacterial speck disease in tomato. Annu. Rev. Phytopathol. 41:215-243.

Popescu, S. C., Popescu, G. V., Bachan, S., Zhang, Z., Gerstein, M., Snyder, M., and Dinesh-Kumar, S. P. 2009. MAPK target networks in *Arabidopsis thaliana* revealed using functional protein microarrays. Genes Dev. 23:80-92.

Potato Genome Sequencing Consortium. 2011. Genome sequence and analysis of the tuber crop potato. Nature 475:189-195.

Ramsey, J. S., Wilson, A. C., de Vos, M., Sun, Q., Tamborindeguy, C., Winfield, A., Malloch, G., Smith, D. M., Fenton, B., Gray, S. M., and Jander, G. 2007. Genomic resources for Myzus persicae: EST sequencing, SNP identification, and microarray design. BMC Genomics 8:423.

Riely, B. K., and Martin, G. B. 2001. Ancient origin of pathogen recognition specificity conferred by the tomato disease resistance gene Pto. Proc. Natl. Acad. Sci. U.S.A. 98:2059-2064.

Robertson, D. 2004. VIGS vectors for gene silencing: Many targets, many tools. Annu. Rev. Plant Physiol. Plant Mol. Biol. 55:495-519.

Sacco, M. A., Mansoor, S., and Moffett, P. 2007. A RanGAP protein physically interacts with the NB-LRR protein Rx, and is required for Rx-mediated viral resistance. Plant J. 52:82-93.

Schmieder, R., and Edwards, R. 2011. Quality control and preprocessing of metagenomic datasets. Bioinformatics 27:863-864.

Senthil-Kumar, M., and Mysore, K. S. 2011. New dimensions for VIGS in plant functional genomics. Trends Plant Sci. 16:656-665.

Senthil-Kumar, M., Hema, R., Anand, A., Kang, L., Udayakumar, M., and Mysore, K. S. 2007. A systematic study to determine the extent of gene silencing in *Nicotiana benthamiana* and other Solanaceae species when heterologous gene sequences are used for virus-induced gene silencing. New Phytol. 176:782-791.

Simpson, J. T., Wong, K., Jackman, S. D., Schein, J. E., Jones, S. J., and Birol, I. 2009. ABySS: A parallel assembler for short read sequence data. Genome Res. 19:1117-1123.

Slater, G. S., and Birney, E. 2005. Automated generation of heuristics for biological sequence comparison. BMC Bioinformatics 6:31.

Takahashi, Y., Nasir, K. H., Ito, A., Kanzaki, H., Matsumura, H., Saitoh, H., Fujisawa, S., Kamoun, S., and Terauchi, R. 2007. A high-throughput screen of cell-death-inducing factors in *Nicotiana benthamiana* identifies a novel MAPKK that mediates INF1-induced cell death signaling and non-host resistance to *Pseudomonas cichorii*. Plant J. 49:1030-1040.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. 2011. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. 28:2731-2739.

Tang, X., Frederick, R. D., Zhou, J., Halterman, D. A., Jia, Y., and Martin, G. B. 1996. Initiation of plant disease resistance by physical interaction of AvrPto and Pto kinase. Science 274:2060-2063.

Thompson, J. D., Higgins, D. G., and Gibson, T. J. 1994. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22:4673-4680.

Tomato Genome Consortium. 2012. The tomato genome sequence provides insights into fleshy fruit evolution. Nature 485:635-641.

Van der Hoorn, R. A. L., Laurent, F., Roth, R., and De Wit, P. J. G. M. 2000. Agroinfiltration is a versatile tool that facilitates comparative analyses of Avr9/Cf-9-induced and Avr4/Cf-4-induced necrosis. Mol. Plant-Microbe Interact. 13:439-446.

Velasquez, A. C., Chakravarthy, S., and Martin, G. B. 2009. Virus-induced gene silencing (VIGS) in *Nicotiana benthamiana* and tomato. JoVE 28: e1292. doi:10.3791/1292. Published online.

Wagner, B., Fuchs, H., Adhami, F., Ma, Y., Scheiner, O., and Breiteneder, H. 2004. Plant virus expression systems for transient production of recombinant allergens in *Nicotiana benthamiana*. Methods 32:227-234.

Wei, C. F., Kvitko, B. H., Shimizu, R., Crabill, E., Alfano, J. R., Lin, N. C., Martin, G. B., Huang, H. C., and Collmer, A. 2007. A *Pseudomonas syringae* pv. *tomato* DC3000 mutant lacking the type III effector HopQ1-1 is able to cause disease in the model plant *Nicotiana benthamiana*. Plant J. 51:32-46.

Yeam, I., Nguyen, H. P., and Martin, G. B. 2010. Phosphorylation of the *Pseudomonas syringae* effector AvrPto is required for FLS2/BAK1-independent virulence activity and recognition by tobacco. Plant J. 61:16-24.

Yoshino, K., Irieda, H., Sugimoto, F., Yoshioka, H., Okuno, T., and Takano, Y. 2012. Cell death of *Nicotiana benthamiana* is induced by secreted protein NIS1 of *Colletotrichum orbiculare* and is suppressed by a homologue of CgDN3. Mol. Plant-Microbe Interact. 25:625-636.
Zhang, H. B., Zhao, X. P., Ding, X. D., Paterson, A. D., and Wing, R. A. 1995. Preparation of megabase-size DNA from plant nuclei. Plant J. 7:175-184.

AUTHOR-RECOMMENDED INTERNET RESOURCES

BLAST site for *N. benthamiana* genome sequence: solgenomics.net/tools/blast/index.pl?db_id=196
ea-utils FASTQ processing utilities website: code.google.com/p/ea-utils/
Illumina Sequencing Technology website: www.illumina.com/technology/sequencing_technology.ilmn
NCBI BioProject for *N. benthamiana* genome sequencing page: www.ncbi.nlm.nih.gov/bioproject?term=PRJNA170566
Short Oligonucleotide Analysis Package (SOAP) database: soap.genomics.org.cn/about.html#resource2
SOL Genomics Network (SGN): solgenomics.net
SGN International Tomato Genome Sequencing Project: solgenomics.net/organism/Solanum_lycopersicum/genome
SGN *Nicotiana benthamiana* ftp site: ftp://ftp.solgenomics.net/benthamiana
SGN Unigene Build: *Nicotiana benthamiana* 1: solgenomics.net/search/unigene_build.pl?id=43
Solanaceae Genomics Resource: solanaceae.plantbiology.msu.edu
Tomato Functional database: http://ted.bti.cornell.edu
Tomato Genetics Resource Center: tgrc.ucdavis.edu
Weill Cornell Medical College website: weill.cornell.edu/research/research_support/core_facilities/genomics.html