# Supporting Information

**Bart et al. 10.1073/pnas.1208003109**

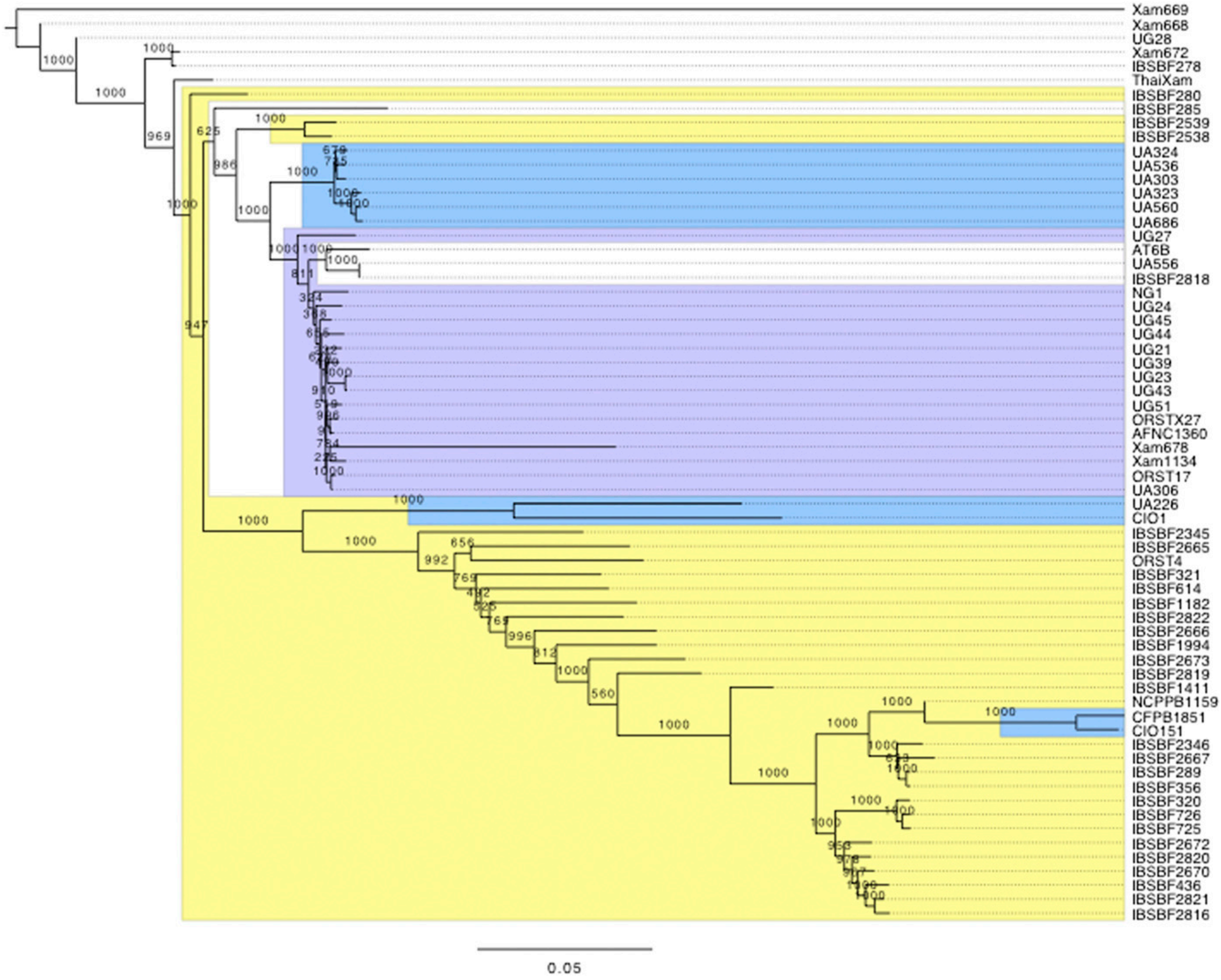


**Fig. S1.** Neighbor-joining analysis of *Xam* strains. Colors indicate geographic origins. Analysis is based on same data as in Fig. 2. CLC Genomics Workbench was used for reference-based assemblies and SNP identification for each *Xam* strain was compared with the *Xam668* genome (*Materials and Methods*). *Xam* strains cluster primarily by geographic region with a few outliers. Yellow, Brazil; blue, Colombia; purple, Africa; white, other/geographic outliers. Bootstrap (1,000 replicates) was performed as a test of significance. Tree is unrooted.

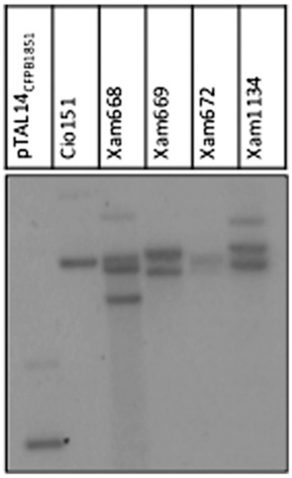


**Fig. S2.** *Xam* strains contain diverse transcription activator-like (TAL) effector-containing plasmid profiles. Plasmid DNA was isolated and run on a 0.75% agarose gel. Southern analysis was performed to assess the TAL effector sequence content of each plasmid.

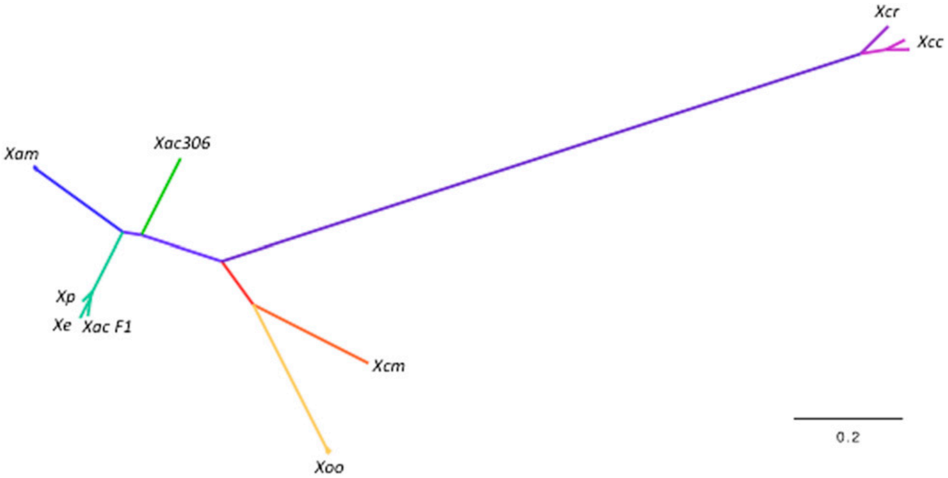
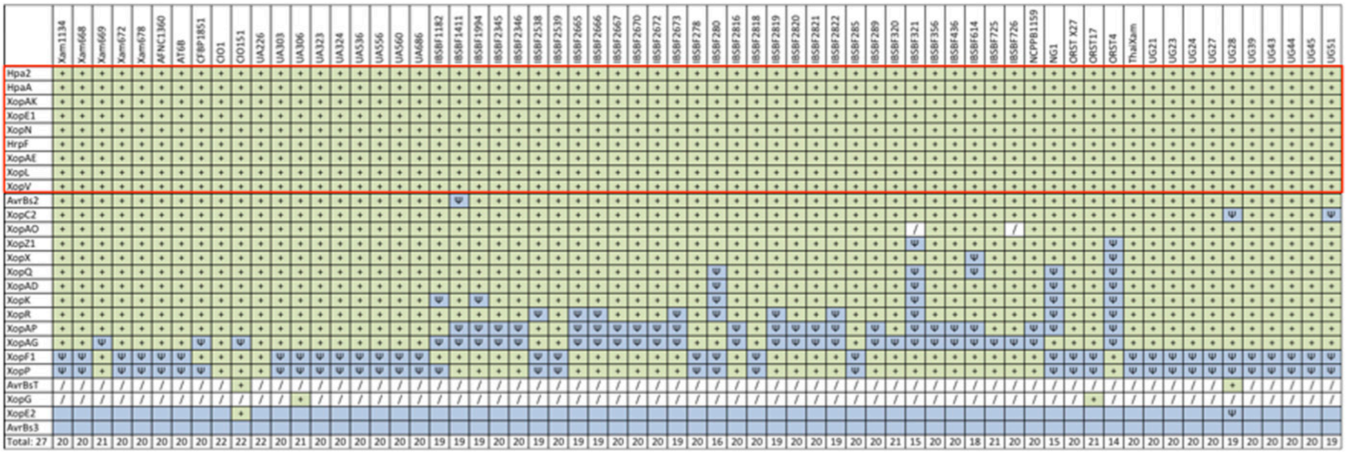


**Fig. S3.** Maximum-likelihood phylogeny of sequenced xanthomonads places *Xanthomonas axonopodis* pv. *manihotis* (*Xam*) among its closest relatives. Phylogeny is based on full-genome SNP analysis, 1,442,868 SNPs (*Materials and Methods*). *Xam*, strains NG1, Xam1134, Xam678, Xam668, Xam672, CIO151, and Xam669 cluster together; *Xp*, *Xanthomonas perforans* 91–118; *Xe*, *Xanthomonas euvesicatoria* 85–10; *Xac F1*, *X. axonopodis* pv. *citrumello F1*; *Xac306*, *X. a*. pv. *citri*; *Xoo*, *Xanthomonas oryzae* pv. *oryzae,* strains Pxo99A, 311018, and 10331 cluster together; *Xcm*, *Xanthomonas campestris* pv. *musacearum*; *Xcc*, *X. c.* pv. *campestris*, strains B100, 8004, and 33913 cluster together; *Xcr*, *X. c.* pv. *raphani* 756C.

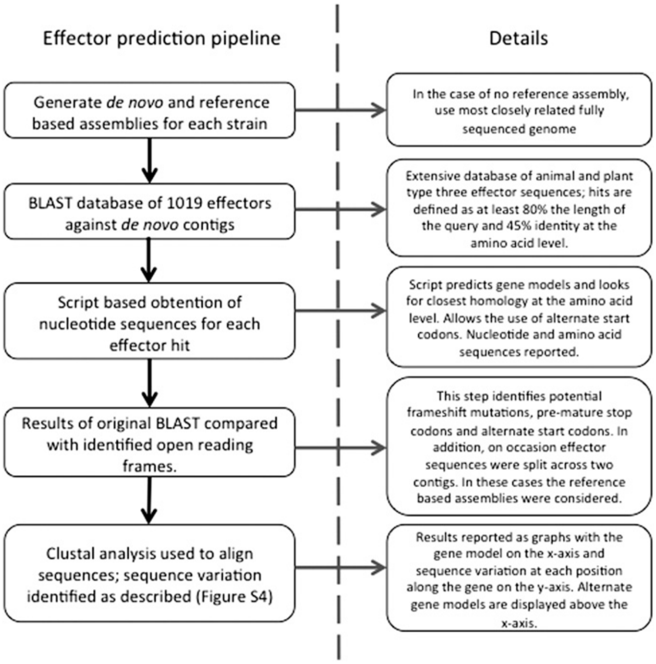


| | Xam1134 | Xam668 | Xam669 | Xam672 | Xam678 | AFNC1360 | AT68 | CFBP1851 | CIO1 | CIO151 | UA226 | UA303 | UA306 | UA323 | UA324 | UA536 | UA556 | UA560 | UA686 | IBSBF1182 | IBSBF1411 | IBSBF1994 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Hpa2 | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + |
| HpaA | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + |
| XopAK | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + |
| XopE1 | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + |
| XopN | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + |
| HrpF | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + |
| XopAE | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + |
| XopL | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + |
| XopV | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + |
| AvrBs2 | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | Ψ | + |
| XopC2 | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + |
| XopAO | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + |
| XopZ1 | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + |
| XopX | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + |
| XopQ | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + |
| XopAD | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + |
| XopK | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | Ψ | + | Ψ |
| XopR | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + |
| XopAP | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | Ψ | Ψ |
| XopAG | + | + | Ψ | + | + | + | + | Ψ | + | Ψ | + | + | + | + | + | + | + | + | + | Ψ | Ψ | Ψ |
| XopF1 | Ψ | Ψ | + | Ψ | Ψ | Ψ | Ψ | + | + | + | + | Ψ | Ψ | Ψ | Ψ | Ψ | Ψ | Ψ | Ψ | + | + | + |
| XopP | Ψ | Ψ | + | Ψ | Ψ | Ψ | Ψ | Ψ | + | + | + | Ψ | Ψ | Ψ | Ψ | Ψ | Ψ | Ψ | Ψ | Ψ | + | + |
| AvrBsT | / | / | / | / | / | / | / | / | / | + | / | / | / | / | / | / | / | / | / | / | / | / |
| XopG | / | / | / | / | / | / | / | / | / | / | / | / | + | / | / | / | / | / | / | / | / | / |
| XopE2 | | | | | | | | | | + | | | | | | | | | | | | |
| AvrBs3 | | | | | | | | | | | | | | | | | | | | | | |
| Total: 27 | 20 | 20 | 21 | 20 | 20 | 20 | 20 | 20 | 22 | 22 | 22 | 20 | 21 | 20 | 20 | 20 | 20 | 20 | 20 | 19 | 19 | 19 |

| | IBSBF2345 | IBSBF2346 | IBSBF2538 | IBSBF2539 | IBSBF2665 | IBSBF2666 | IBSBF2667 | IBSBF2670 | IBSBF2672 | IBSBF2673 | IBSBF278 | IBSBF280 | IBSBF2816 | IBSBF2818 | IBSBF2819 | IBSBF2820 | IBSBF2821 | IBSBF2822 | IBSBF285 | IBSBF289 | IBSBF320 | IBSBF321 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Hpa2 | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + |
| HpaA | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + |
| XopAK | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + |
| XopE1 | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + |
| XopN | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + |
| HrpF | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + |
| XopAE | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + |
| XopL | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + |
| XopV | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + |
| AvrBs2 | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + |
| XopC2 | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + |
| XopAO | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | / |
| XopZ1 | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | Ψ |
| XopX | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + |
| XopQ | + | + | + | + | + | + | + | + | + | + | + | Ψ | + | + | + | + | + | + | + | + | + | Ψ |
| XopAD | + | + | + | + | + | + | + | + | + | + | + | Ψ | + | + | + | + | + | + | + | + | + | Ψ |
| XopK | + | + | + | + | + | + | + | + | + | + | + | Ψ | + | + | + | + | + | + | + | + | + | Ψ |
| XopR | + | + | Ψ | + | Ψ | Ψ | + | + | + | Ψ | + | Ψ | + | + | Ψ | + | + | Ψ | + | + | + | Ψ |
| XopAP | Ψ | Ψ | + | + | Ψ | Ψ | Ψ | Ψ | Ψ | Ψ | + | + | Ψ | + | Ψ | Ψ | Ψ | Ψ | + | Ψ | + | Ψ |
| XopAG | Ψ | Ψ | + | + | Ψ | Ψ | Ψ | Ψ | Ψ | Ψ | + | + | Ψ | + | Ψ | Ψ | Ψ | Ψ | + | Ψ | Ψ | Ψ |
| XopF1 | + | + | Ψ | Ψ | + | + | + | + | + | + | Ψ | Ψ | + | Ψ | + | + | + | + | Ψ | + | + | + |
| XopP | + | + | Ψ | Ψ | + | + | + | + | + | + | Ψ | Ψ | + | Ψ | + | + | + | + | Ψ | + | + | + |
| AvrBsT | / | / | / | / | / | / | / | / | / | / | / | / | / | / | / | / | / | / | / | / | / | / |
| XopG | / | / | / | / | / | / | / | / | / | / | / | / | / | / | / | / | / | / | / | / | / | / |
| XopE2 | | | | | | | | | | | | | | | | | | | | | | |
| AvrBs3 | | | | | | | | | | | | | | | | | | | | | | |
| Total: 27 | 20 | 20 | 19 | 20 | 19 | 19 | 20 | 20 | 20 | 19 | 20 | 16 | 20 | 20 | 19 | 20 | 20 | 19 | 20 | 20 | 21 | 15 |

| | IBSBF356 | IBSBF436 | IBSBF614 | IBSBF725 | IBSBF726 | NCPPB1159 | NG1 | ORST X27 | ORST17 | ORST4 | ThaiXam | UG21 | UG23 | UG24 | UG27 | UG28 | UG39 | UG43 | UG44 | UG45 | UG51 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Hpa2 | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + |
| HpaA | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + |
| XopAK | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + |
| XopE1 | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + |
| XopN | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + |
| HrpF | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + |
| XopAE | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + |
| XopL | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + |
| XopV | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + |
| AvrBs2 | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + |
| XopC2 | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | Ψ | + | + | + | + | Ψ |
| XopAO | + | + | + | + | / | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + |
| XopZ1 | + | + | + | + | + | + | + | + | + | Ψ | + | + | + | + | + | + | + | + | + | + | + |
| XopX | + | + | Ψ | + | + | + | + | + | + | Ψ | + | + | + | + | + | + | + | + | + | + | + |
| XopQ | + | + | Ψ | + | + | + | Ψ | + | + | Ψ | + | + | + | + | + | + | + | + | + | + | + |
| XopAD | + | + | + | + | + | + | Ψ | + | + | Ψ | + | + | + | + | + | + | + | + | + | + | + |
| XopK | + | + | + | + | + | + | Ψ | + | + | Ψ | + | + | + | + | + | + | + | + | + | + | + |
| XopR | + | + | + | + | + | + | Ψ | + | + | Ψ | + | + | + | + | + | + | + | + | + | + | + |
| XopAP | Ψ | Ψ | Ψ | + | + | Ψ | Ψ | + | + | Ψ | + | + | + | + | + | + | + | + | + | + | + |
| XopAG | Ψ | Ψ | Ψ | Ψ | Ψ | Ψ | + | + | + | Ψ | + | + | + | + | + | + | + | + | + | + | + |
| XopF1 | + | + | + | + | + | + | Ψ | Ψ | Ψ | + | + | Ψ | Ψ | Ψ | Ψ | Ψ | Ψ | Ψ | Ψ | Ψ | Ψ |
| XopP | + | + | + | + | + | + | Ψ | Ψ | Ψ | + | + | Ψ | Ψ | Ψ | Ψ | Ψ | Ψ | Ψ | Ψ | Ψ | Ψ |
| AvrBsT | / | / | / | / | / | / | / | / | / | / | / | / | / | / | / | + | / | / | / | / | / |
| XopG | / | / | / | / | / | / | / | / | + | / | / | / | / | / | / | / | / | / | / | / | / |
| XopE2 | | | | | | | | | | | | | | | | Ψ | | | | | |
| AvrBs3 | | | | | | | | | | | | | | | | | | | | | |
| Total: 27 | 20 | 20 | 18 | 21 | 20 | 20 | 15 | 20 | 21 | 14 | 20 | 20 | 20 | 20 | 20 | 19 | 20 | 20 | 20 | 20 | 19 |

**Fig. S4.** Presence of effectors was determined from de novo and reference-based assemblies. XopE1 and XopE2 have high (~68%) identity. Sequences present in most genomes are more homologous to XopE1, although CIO151 also has a homolog to XopE1 and XopE2. Green, identified through BLAST, confirmed manually; blue, identified through BLAST, sequence polymorphism alters gene model (e.g., frameshift mutation, premature stop codon); white, not identified. Red box outlines conserved effectors with the same gene model as a characterized effector from the effector database. AvrBs3-like sequences were present in all strains of *Xam*; however, they were not completely assembled because of the repetitive central domain.



Generate *de novo* and reference based assemblies for each strain

In the case of no reference assembly, use most closely related fully sequenced genome

BLAST database of 1019 effectors against *de novo* contigs

Extensive database of animal and plant type three effector sequences; hits are defined as at least 80% the length of the query and 45% identity at the amino acid level.

Script based obtention of nucleotide sequences for each effector hit

Script predicts gene models and looks for closest homology at the amino acid level. Allows the use of alternate start codons. Nucleotide and amino acid sequences reported.

Results of original BLAST compared with identified open reading frames.

This step identifies potential frameshift mutations, pre-mature stop codons and alternate start codons. In addition, on occasion effector sequences were split across two contigs. In these cases the reference based assemblies were considered.

Clustal analysis used to align sequences; sequence variation identified as described (Figure S4)

Results reported as graphs with the gene model on the x-axis and sequence variation at each position along the gene on the y-axis. Alternate gene models are displayed above the x-axis.

**Fig. S5.** Cartoon summarizing our effector prediction and sequence variation analysis pipeline. Python scripts can be found at http://pypi.python.org/pypi/biotools/1.1.2.

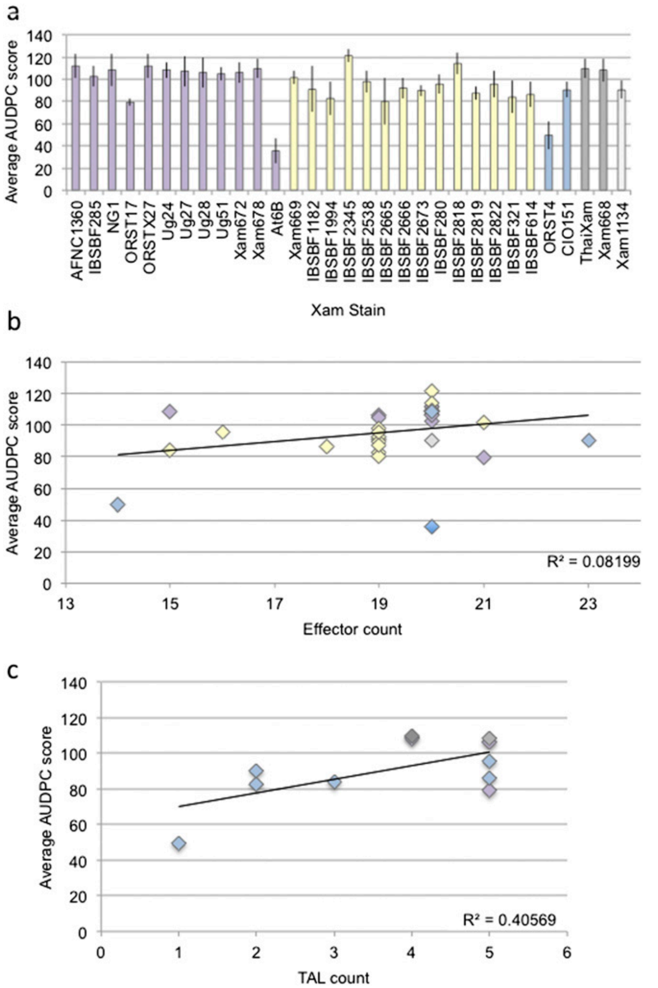


**Fig. S6.** (*A*) Area under the disease progression (AUDPC) measure of virulence does not show a significant trend between effector count and virulence level. Average disease symptoms *in planta* (measured by area under the disease progression curve) of *Xam* strains are grouped by origin (Africa, purple; Venezuela, orange; Brazil, yellow; Colombia, blue; Indonesia, gray; unknown, white); mean ± SD (*n* = 6) is displayed. (*B* and *C*) Effector and TAL count compared with disease symptoms *in planta* (measured by AUDPC) of tested *Xam* strains.

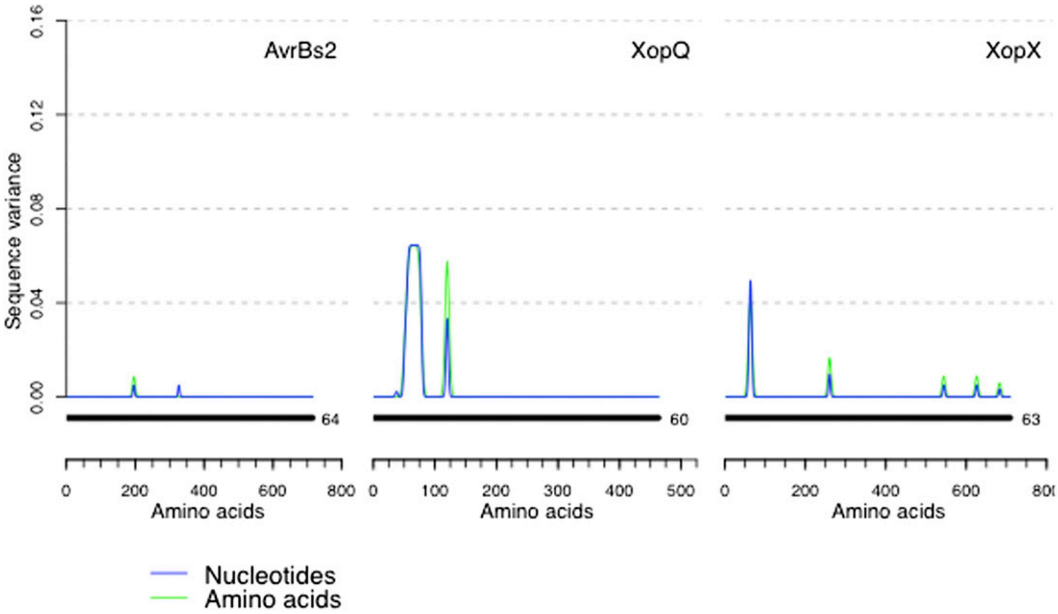


**Fig. S7.** Sequence variance of select *Xam* effectors. Sequence variance was calculated as described in *Materials and Methods*. Gene models are displayed immediately above the *x* axis. In some cases, potential alternate start codons are present 5′ to the predicted ATG. In these cases, the sequences have been excluded in these plots.

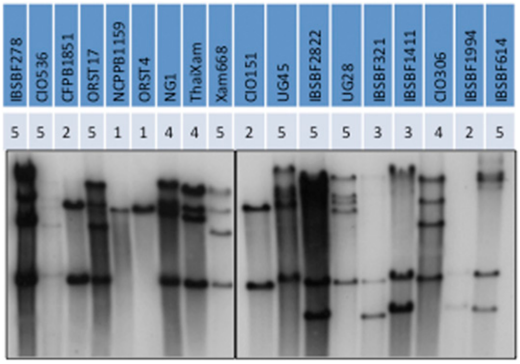


**Fig. S8.** Southern analysis of genomic DNA digested with EcoRI. All blots were probed with a $^{32}$P-labeled internal BamHI fragment of TAL 14. TAL effector copy number was compared with virulence level (Fig. 3). Faint bands were confirmed in repeated experiments.

**Table S1.** ***Xam*** **strain and assembly information**

| | | | | Assembly statistics* | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Strain | Year | Origin | Collector | N50 | Length | Contigs | Max | Mean | SD | Reads |
| AFNC1360 | 1998 | Benin or Nigeria | AFNC | 25 | 4,923,650 | 150 | 235,381 | 32,824 | 37,800 | 20,051,776 |
| AT6B | 2006 | Venezuela | M.S. | 22 | 4,946,093 | 163 | 237,170 | 30,344 | 39,662 | 19,151,609 |
| CFBP1851 | 1974 | Columbia | CFBP | 17 | 4,798,436 | 123 | 281,156 | 39,011 | 52,886 | 3,310,501 |
| CIO1 | 1995 | Columbia | CIO | 19 | 4,898,899 | 128 | 211,998 | 38,272 | 45,752 | 13,719,586 |
| CIO151 | 1995 | Columbia | CIO | 24 | 4,911,246 | 175 | 150,154 | 28,064 | 35,954 | 25,616,672 |
| UA226 | 2008 | Columbia | A.B. | 18 | 4,951,862 | 129 | 290,136 | 38,386 | 49,473 | 11,330,262 |
| UA303 | 2008 | Columbia | A.B. | 23 | 4,872,968 | 130 | 196,821 | 37,484 | 39,723 | 15,222,193 |
| UA306 | 2008 | Columbia | A.B. | 23 | 5,065,231 | 144 | 235,368 | 35,175 | 42,248 | 15,102,007 |
| UA323 | 2008 | Columbia | A.B. | 23 | 4,872,007 | 151 | 173,776 | 32,264 | 37,606 | 20,612,168 |
| UA324 | 2008 | Columbia | A.B. | 22 | 4,916,013 | 146 | 161,357 | 33,671 | 38,915 | 90,20,217 |
| UA536 | 2009 | Columbia | A.B. | 23 | 4,866,301 | 142 | 183,757 | 34,269 | 39,356 | 16,292,631 |
| UA556 | 2009 | Columbia | A.B. | 22 | 4,842,655 | 139 | 237,273 | 34,839 | 41,407 | 11,470,344 |
| UA560 | 2009 | Columbia | A.B. | 27 | 4,897,922 | 154 | 176,262 | 31,804 | 33,527 | 17,000,477 |
| UA686 | 2009 | Columbia | A.B. | 22 | 4,898,003 | 143 | 196,787 | 34,251 | 40,679 | 15,947,863 |
| IBSBF1182 | 1995 | Brazil | IBSBF | 128 | 4,773,692 | 648 | 46,493 | 7,366 | 6,968 | 2,149,549 |
| IBSBF1411 | 1998 | Brazil | IBSBF | 15 | 4,856,132 | 130 | 314,799 | 37,354 | 55,031 | 4,275,532 |
| IBSBF1994 | 2004 | Brazil | IBSBF | 137 | 4,774,518 | 677 | 45,564 | 7,052 | 6,508 | 2,232,128 |
| IBSBF2345 | 2006 | Brazil | IBSBF | 19 | 4,819,772 | 128 | 286,963 | 37,654 | 47,959 | 1,904,256 |
| IBSBF2346 | 2006 | Brazil | IBSBF | 18 | 4,874,593 | 122 | 257,932 | 39,955 | 48,696 | 17,994,941 |
| IBSBF2538 | 2007 | Brazil | IBSBF | 122 | 4,833,429 | 573 | 50,456 | 8,435 | 7,446 | 3,398,164 |
| IBSBF2539 | 2007 | Brazil | IBSBF | 24 | 4,952,612 | 138 | 196,970 | 35,888 | 39,356 | 15,572,597 |
| IBSBF2665 | 2009 | Brazil | IBSBF | 83 | 4,836,797 | 434 | 111,173 | 11,144 | 11,333 | 1,893,919 |
| IBSBF2666 | 2009 | Brazil | IBSBF | 49 | 4,842,277 | 311 | 123,056 | 15,570 | 18,537 | 2,797,254 |
| IBSBF2667 | 2009 | Brazil | IBSBF | 21 | 4,913,060 | 138 | 226,291 | 35,601 | 42,268 | 13,781,769 |
| IBSBF2670 | 2009 | Brazil | IBSBF | 17 | 4,902,902 | 139 | 28,0128 | 35,272 | 50,872 | 13,653,945 |
| IBSBF2672 | 2009 | Brazil | IBSBF | 18 | 4,856,510 | 131 | 257,571 | 37,072 | 48,420 | 8,192,559 |
| IBSBF2673 | 2009 | Brazil | IBSBF | 53 | 4,830,620 | 305 | 120,981 | 15,838 | 17,359 | 2,649,843 |
| IBSBF278 | 1965 | Brazil | IBSBF | 20 | 5,020,210 | 138 | 239,334 | 36,378 | 44,921 | 12,084,985 |
| IBSBF280 | 1973 | Brazil | IBSBF | 171 | 4,842,752 | 850 | 41,005 | 5,697 | 5,251 | 1,750,289 |
| IBSBF2816 | 2010 | Brazil | IBSBF | 19 | 4902619 | 145 | 256,960 | 33,811 | 46,242 | 13,153,736 |
| IBSBF2818 | 2010 | Brazil | IBSBF | 22 | 4851686 | 139 | 235,402 | 34,904 | 42,367 | 11,474,331 |
| IBSBF2819 | 2010 | Brazil | IBSBF | 104 | 4,852,810 | 557 | 59,730 | 8,712 | 8,583 | 2,479,861 |
| IBSBF2820 | 2010 | Brazil | IBSBF | 19 | 4,930,429 | 150 | 257,651 | 32,869 | 44,768 | 9,028,899 |
| IBSBF2821 | 2010 | Brazil | IBSBF | 21 | 4,906,958 | 161 | 257,592 | 30,478 | 41,909 | 13,077,491 |
| IBSBF2822 | 2010 | Brazil | IBSBF | 93 | 4,831,349 | 464 | 57,666 | 10,412 | 9,543 | 2,033,714 |
| IBSBF285 | 1973 | Nigeria | IBSBF | 21 | 4,972,047 | 147 | 186,156 | 33,823 | 40,851 | 12,946,844 |
| IBSBF289 | ? | Brazil | IBSBF | 23 | 4,847,529 | 155 | 226,288 | 31,274 | 38,049 | 15,106,454 |
| IBSBF320 | 1974 | Brazil | IBSBF | 17 | 4,803,726 | 121 | 257,673 | 39,700 | 49,436 | 14,285,701 |
| IBSBF321 | 1974 | Brazil | IBSBF | 183 | 4,593,056 | 888 | 35,503 | 5,172 | 4,646 | 1,869,048 |
| IBSBF356 | 1980 | Brazil | IBSBF | 18 | 4,847,079 | 130 | 257,951 | 37,285 | 49,019 | 14,264,479 |
| IBSBF436 | 1982 | Brazil | IBSBF | 19 | 4,903,327 | 149 | 257,580 | 32,908 | 44,502 | 13,185,959 |
| IBSBF614 | 1987 | Brazil | IBSBF | 136 | 4,730,788 | 706 | 51,491 | 6,700 | 6,516 | 2,138,932 |
| IBSBF725 | 1988 | Brazil | IBSBF | 19 | 4,886,517 | 141 | 257,945 | 34,656 | 45,893 | 14,295,676 |
| IBSBF726 | 1988 | Brazil | IBSBF | 21 | 4,832,905 | 140 | 257,968 | 34,520 | 43,377 | 17,714,917 |
| NCPPB1159 | 1941 | Brazil | NCPPB | 17 | 4,797,980 | 128 | 257,906 | 37,484 | 48,746 | 13,904,209 |
| NG1 | 2011 | Nigeria | L.K. | 160 | 4,737,721 | 785 | 38,252 | 6,035 | 5,419 | 3,934,608 |
| ORST17 | 1984 | Congo | ORST | 23 | 5,128,104 | 177 | 235,384 | 28,972 | 40,180 | 15,379,085 |
| ORST4 | 1974 | Columbia | ORST | 161 | 4,650,210 | 767 | 36,169 | 6,062 | 5,309 | 1,942,504 |
| ORST X27 | 1989 | Togo | ORST | 21 | 4,931,909 | 133 | 237,243 | 37,082 | 43,419 | 14,896,651 |
| ThaiXam | 1997 | Thailand | N.K. | 62 | 4,874,623 | 324 | 84,544 | 15,045 | 14,840 | 2,770,771 |
| UG21 | 2011 | Uganda | T.A. | 23 | 5,010,468 | 148 | 237,252 | 33,854 | 40,078 | 18,213,584 |
| UG23 | 2011 | Uganda | T.A. | 22 | 4,924,062 | 133 | 237,150 | 37,023 | 42,562 | 8,918,645 |
| UG24 | 2011 | Uganda | T.A. | 20 | 4,909,909 | 145 | 233,940 | 33,861 | 43,565 | 5,036,494 |
| UG27 | 2011 | Uganda | T.A. | 19 | 4,903,342 | 131 | 268,832 | 37,430 | 47,341 | 3,629,887 |
| UG28 | 2011 | Uganda | T.A. | 131 | 4,722,734 | 676 | 54,751 | 6,986 | 6,899 | 9,803,349 |
| UG39 | 2011 | Uganda | T.A. | 21 | 4,911,502 | 130 | 237,280 | 37,780 | 43,922 | 10,380,844 |
| UG43 | 2011 | Uganda | T.A. | 23 | 4,924,182 | 145 | 237,167 | 33,959 | 41,100 | 16,650,194 |
| UG44 | 2011 | Uganda | T.A. | 24 | 4,895,442 | 141 | 238,358 | 34,719 | 40,012 | 19,277,873 |
| UG45 | 2011 | Uganda | T.A. | 59 | 4,986,397 | 323 | 83,811 | 15,437 | 15,477 | 5,338,599 |
| UG51 | 2011 | Uganda | T.A. | 22 | 4,871,289 | 140 | 234,400 | 34,794 | 41,630 | 14,764,547 |
| Xam1134 | ? | ? | ? | 28 | 4,894,844 | 158 | 237,227 | 30,980 | 34,575 | 29,876,966 |

**Table S1. Cont.**

| | | | | Assembly statistics* | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Strain | Year | Origin | Collector | N50 | Length | Contigs | Max | Mean | SD | Reads |
| Xam668 | 1978 | Indonesia | NZTCC | 29 | 4,954,860 | 195 | 276,859 | 25,409 | 33,003 | 12,733,850 |
| Xam669 | 1973 | Brazil | NZTCC | 21 | 4,820,297 | 137 | 201,435 | 35,184 | 40,806 | 15,193,524 |
| Xam672 | 1976 | Cameroon | NZTCC | 25 | 5,016,075 | 165 | 239,253 | 30,400 | 37,139 | 22,340,056 |
| Xam678 | 1973 | DR Congo | NZTCC | 47 | 4,894,142 | 56 | 104,760 | 19,117 | 19,014 | 7,734,076 |

A.B., Adriana Bernal; L.K., Lava Kumar; MS., María Angélica Santana; N.K., Nuttima Kositcharoenkul; T.A., Titus Alicai.

*N50, the smallest number of contigs needed to make up half of the assembly; Length, sum of all contig lengths (nucleotides); Max, maximum contig length; Mean, mean contig length; SD, SD from the mean contig length; Reads, total number of reads contributing to each assembly.

**Table S2. Illumina sequencing price breakdown**

| Item | Price per sample, $* |
|---|---|
| Library construction | 15 |
| Reagents | 70 |
| Sequencing | 35 |
| Total | 120 |

*Price assumes use of Apollo 324 automated library prep, IntegenX reagents, sequencing primers ordered through Integrated DNA Technology (PAGE purified), and 48 samples multiplexed in one lane of an Illumina HiSeq2000. Prices are current for University of California members as of 2010.

## Other Supporting Information Files

Dataset S1 (XLS)