MICROBIOLOGY

Chromosome architecture affects virulence and competitiveness in *Agrobacterium tumefaciens* C58

Ephraim Aliu[1,2,3], Keunsub Lee[1,2], Ram Sanath-Kumar[4]†, Christopher Youngstrom[1,2]†, Lander Geadelmann[3,5], Dibyajyoti Pramanik[1,2], Qing Ji[2], Xindan Wang[4], Kan Wang[1,2]*

Chromosome architecture plays a crucial role in bacterial adaptation, yet its direct impact remains unclear. Different bacterial species and even strains within the same species exhibit diverse chromosomal configurations, including a single circular or linear chromosome, two circular chromosomes, or a circular-linear combination. To investigate how these architectures shape bacterial behavior, we generated near-isogenic strains representing each configuration in *Agrobacterium tumefaciens* C58, an important soil bacterium widely used for plant genetic transformation. Strains with a single-chromosome architecture, whether linear or circular, exhibited faster growth, enhanced stress tolerance, and greater interstrain competitiveness. In contrast, bipartite chromosome strains showed higher virulence gene expression and enhanced transient plant transformation efficiency, suggesting a pathogenic adaptation. Whole-transcriptome analysis revealed architecture-dependent gene expression patterns, underscoring the profound impact of chromosome organization on *Agrobacterium* fitness and virulence. These findings highlight how chromosome structure influences bacterial adaptation and shapes evolutionary trajectories.

INTRODUCTION

Since the discovery of single circular and linear bacterial chromosomes (*1*, *2*), accumulated evidence indicates that prokaryotic genomes exhibit greater structural diversity than previously recognized (*3*–*5*). While most bacterial genomes consist of a single circular chromosome, alternative architectures, including multipartite genome structures with circular and linear chromosomes, have been observed across prokaryotic genomes (*3*–*5*). These alternative (altered) chromosome architectures can modulate gene expression, promote genome rearrangements, influence horizontal gene transfer, and alter mutation rates—all of which serve as substrates for bacterial adaptation (*6*–*9*). However, the direct impact of chromosome structure—specifically the division into multiple replicons or linearization—on bacterial fitness and adaptation remains poorly understood, largely due to the lack of ideal model systems in which genome architecture is the primary variable (*10*, *11*). In the case of *Agrobacterium*, an early evolutionary step involved chromid formation, followed by chromid linearization, which contributed to species diversification (*12*). These structural transitions may have facilitated genomic plasticity, enabling niche adaptation and long-term survival in host-associated environments (*12*). Therefore, understanding how such architectural changes affect bacterial fitness is key to uncovering the evolutionary implication of multipartite genomes.

The soil-borne, Gram-negative *Agrobacterium* species serve as an ideal model for studying the impact of structural variations in multipartite genomes, as they harbor multiple replicons in both circular and linear configurations (*12*). *Agrobacterium tumefaciens* naturally transfers its DNA into plant genomes, a capability that has been harnessed as a powerful tool for plant genetic transformation (*13*). Its genome comprises a primary circular chromosome (C1), a linear chromid-like replicon (C2), a non-oncogenic megaplasmid (pAt), and an oncogenic tumor-inducing plasmid (pTi). The C1 chromosome encodes core housekeeping functions and is highly conserved across strains (*12*), while the C2 replicon, now recognized as a chromid, is plasmid-derived but has acquired essential genes and chromosome-like stability (*14*). The pAt megaplasmid contributes accessory metabolic functions that enhance ecological fitness (*15*), while the tumor-inducing pTi plasmid harbors genes required for *Agrobacterium* pathogenesis (*13*). On pTi, virulence (*vir*) genes are essential for transfer DNA (T-DNA) processing and its transfer into plant cells. Once integrated into the host genome, plant hormone biosynthesis genes in the T-DNA are expressed, inducing neoplastic growth at the infected sites and resulting in the characteristic crown gall disease symptom (*13*). Also encoded in the T-DNA region are opine synthesis genes, which produce small molecular weight compounds that are used by *Agrobacterium* cells, enabling niche specialization within the tumor environments (*13*, *15*).

Beyond structural organization, genomic studies have uncovered evolutionary markers that distinguish *Agrobacterium* lineages. A key insight is the presence of the protelomerase gene *telA*, which is necessary to stabilize the linear chromid by forming closed hairpin telomeric ends (*16*) and serves as a defining feature for *Agrobacterium* (*12*, *15*, *17*). Strains harboring *telA* typically have a linear chromid, whereas those lacking *telA* retain circular configurations (*17*). One prominent example is the *A. tumefaciens* C58 strain, which was the first *Agrobacterium* genome to be sequenced and reported to harbor a linear chromid (*18*, *19*). The C58 strain has since become a foundational model in plant biotechnology due to its well-annotated genome and established role in plant genetic transformation (*18*–*20*).

The evolutionary drivers of chromid configuration remain an open question. Since the publication of the C58 genome in 2001 (*18*, *19*), more than 200 *Agrobacterium* genome assemblies have been published. Comparative analyses revealed that several *A. tumefaciens* strains also have a linear chromid and the *telA* gene (*15*, *17*). Ramírez-Bahena *et al.* (*17*) suggested that the persistence of a linear chromid might be attributed to the inherent difficulties in reverting it to circular configurations after linearization. Yet, the lack of linear chromids

[1]Department of Agronomy, Iowa State University, Ames, IA 50011, USA. [2]Crop Bioengineering Center, Iowa State University, Ames, IA 50011, USA. [3]Interdepartmental Plant Biology Major, Iowa State University, Ames, IA 50011, USA. [4]Department of Biology, Indiana University, Bloomington, IN 47405, USA. [5]Department of Genetics, Development and Cell Biology, Iowa State University, Ames, IA 50011, USA.
*Corresponding author. Email: kanwang@iastate.edu
†These authors contributed equally to this work.

in many *Agrobacterium* strains (*15*) suggests that factors other than structural constraint might affect chromid configuration. Consistent with this idea, multilocus sequence analysis of *Agrobacterium* strains with a linear chromid revealed that chromid linearity does not seem to alter the rate of genetic interchange (*11*).

Recently, it was discovered that some laboratory and natural *Agrobacterium* strains carry a large linear chromosome, which was generated by a fusion of a circular chromosome and a linear chromid (*21*, *22*). This type of chromosome-chromid fusion has been reported in other bacteria, including *Sinorhizobium meliloti* (*23*), *Vibrio cholerae* (*24*), and *Cupriavidus necator* (*25*), suggesting that homologous recombination-mediated chromosome-chromid fusion is common in bacteria. However, such fusions can also be mediated by transposable elements (*26*).

In this study, we used an efficient CRISPR RNA-guided genome engineering tool INTEGRATE (INsertion of Transposable Elements by Guide RNA–Assisted TargEting) (*27*–*29*) along with the Cre-*lox*P recombination system to generate and compare four near-isogenic *A. tumefaciens* C58 variants, enabling a comprehensive assessment of how chromosome architecture affects bacterial behavior. Our study demonstrates that chromosome architecture substantially influences *Agrobacterium* growth, interstrain competitiveness, stress tolerance, and virulence, directly affecting bacterial behavior and potentially shaping adaptation and evolution.

RESULTS

Natural fusion and synthetic circularization of *Agrobacterium* chromosomes

Using genome-wide high-throughput chromosome conformation capture (Hi-C) assays, Liao *et al.* (*21*) revealed that several laboratory strains of wild-type (WT) C58 (Fig. 1A) and their direct derivatives collected from sixteen different laboratories harbor a single large linear chromosome. The 4.9-Mb linear chromosome in C58F (Fig. 1C) spontaneously occurred through homologous recombination between two near identical genes, *Atu1814* on the circular chromosome (C1) and *Atu3364* on the linear chromid (C2) (*21*). Chromosome fusion events have also recently been reported in two natural *Agrobacterium* isolates, where recombination occurred between two ~6-kb regions containing identical ribosomal RNA (rRNA) and tRNA genes (*22*). The strain C58F (Fig. 1C) also carried an additional 194-kb deletion on the cryptic pAtC58 plasmid [pATΔ; (*21*, *30*)], which appears to be nonessential for *Agrobacterium* survival or virulence (*30*). Given that the major difference between the C58 WT and C58F strains is the chromosomal architecture—specifically, the presence of single or dual chromosomes with circular or linear configurations—the natural fusion variant strain C58F serves as an ideal model to investigate how chromosomal architecture influences *Agrobacterium* physiology, fitness, and interactions with plants.

To systematically assess the impact of chromosome architecture on the fitness and virulence of agrobacteria, we used the INTEGRATE system (*28*, *29*) alongside the Cre-*lox*P recombination to circularize the linear chromosomes in the C58 WT and C58F strains [see Materials and Methods and figs. S1A and S2 (A to D)]. By combining INTEGRATE with a site-specific Cre-*lox*P recombination system, we successfully circularized the C2 linear chromid in WT C58 and the single fused linear chromosome in C58F, respectively (fig. S1, B to E), resulting in a set of four test strains with different chromosome architecture as depicted in Fig. 1: C58 (WT, circular chromosome + linear chromid; Fig. 1A); C58 circularized (Cir, circular chromosome + circular chromid; Fig. 1, B and E), C58F (FL, fused linear chromosome; Fig. 1C), and C58F circularized (FCir, fused circular chromosome; Fig. 1, D and F).

Hi-C analysis ascertained the chromosomal architecture of all four strains (Fig. 2, A to D). Notably, the Cir strain exhibited circularization of the linear chromid (Fig. 2, B and F), while the FCir strain showed a single fused circular chromosome (Fig. 2, D and H). These structural changes were evidenced by the increased contract frequency at the circularization junctions [dark spots at the top left and bottom right corners, marked by red circles in Fig. 2 (B and D)]. Transposon insertion sequencing (Tn-seq) analysis further validated the circularization, as the protelomerase gene (*telA*)—essential for maintaining linear chromosome ends—became nonessential in both Cir and FCir strains (left in Fig. 2I and fig. S3). In addition, the *xerC* and *xerD* genes—the two site-specific recombinase genes important for maintaining the structural integrity of multiple replicons

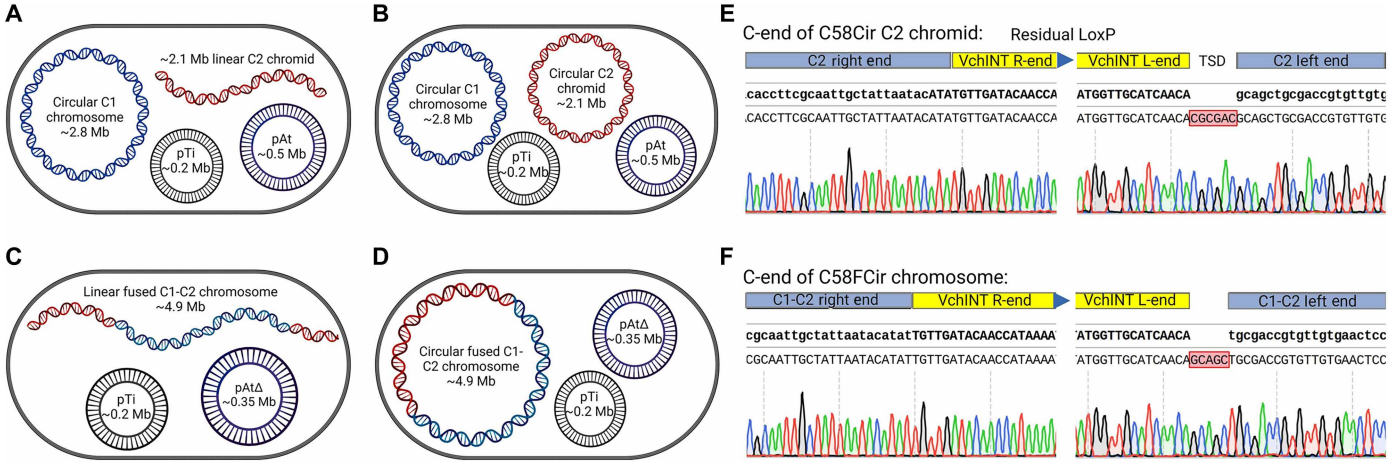


Fig. 1. Generation of isogenic C58 variants. (**A**) C58 (WT, WT *A. tumefaciens* C58, circular C1 chromosome + linear C2 chromid + pAt + pTi). (**B**) C58 circularized (Cir, C1 + circularized C2 chromid + pAt + pTi, generated in this work). (**C**) C58F (FL, a natural single-chromosome strain, linear fused C1-C2 + pAtΔ + pTi). (**D**) C58FCir (FCir, circularized C1-C2 + pAtΔ + pTi, generated in this work). Sanger sequencing chromatograms of the circularized chromosome ends (C-end) in Cir (**E**) and FCir (**F**).

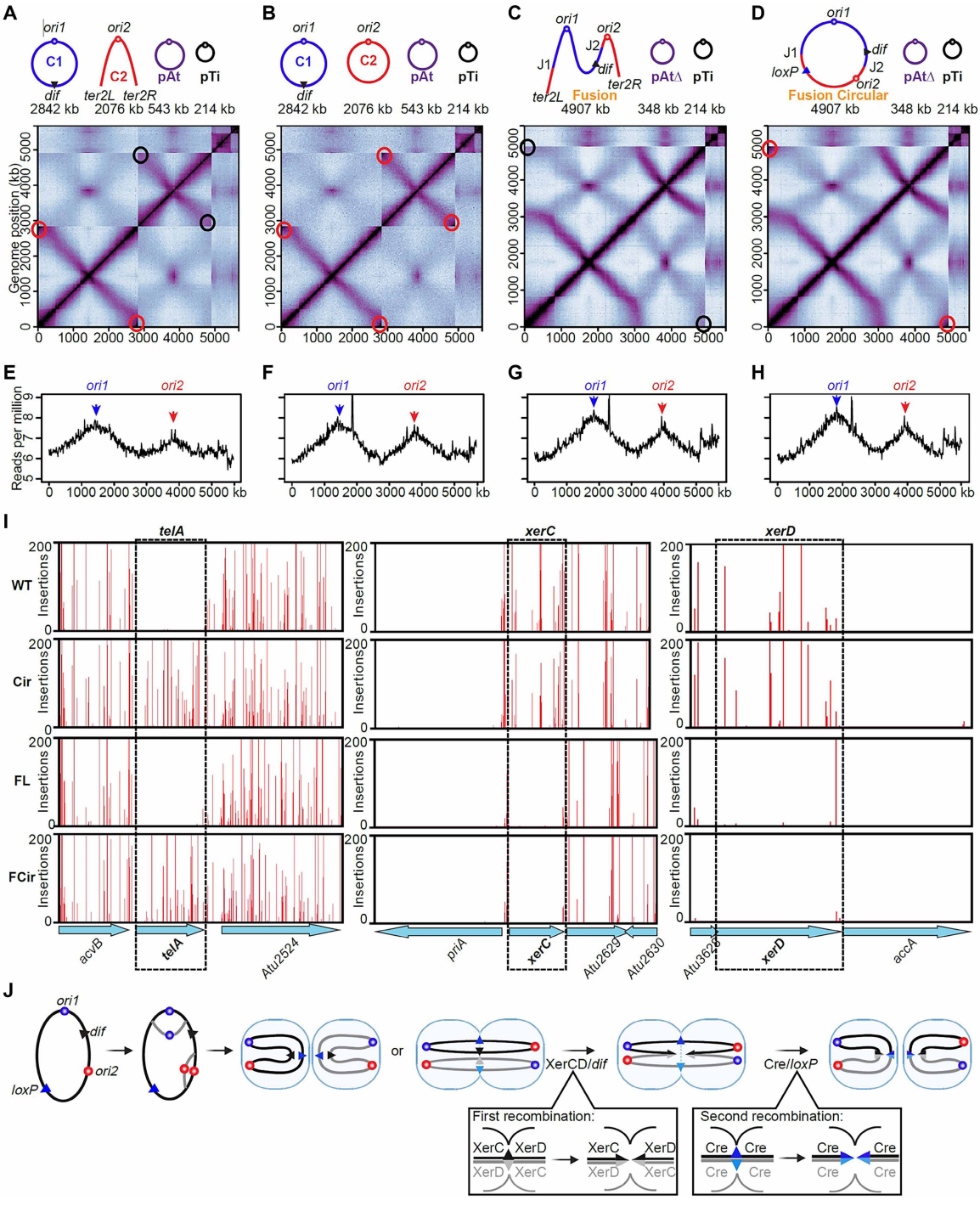


Fig. 2. Characterization of C58 variants. Normalized Hi-C contact maps for *Agrobacterium* chromosomal variants: (**A**) WT, (**B**) Cir, (**C**) FL, and (**D**) FCir. The *x* and *y* axes represent genomic coordinates in kilobases. Red circles denote the presence of circular chromosome ends; black circles indicate their absence. (**E** to **H**) Marker frequency analysis of exponentially growing cells for (E) WT, (F) Cir, (G) FL, and (H) FCir. The *y* axis represents normalized read depth (reads per million mapped reads) across genome positions on the *x* axis. *ori1* and *ori2*, the replication origin of C1 and C2, respectively. (**I**) Tn-seq plots displaying transposon insertion profiles at selected genomic loci. The *y* axis represents the number of sequencing reads at each insertion site, and the *x* axis denotes gene positions surrounding *telA* (left), *xerC* (middle), and *xerD* (right). Regions lacking Tn insertions suggest potential essentiality under the tested conditions. (**J**) Hypothetical schematic model of DNA replication and chromosome segregation in the FCir strain with a single fused circular chromosome. The model is based on the observed requirement for two active replication origins (*ori1* from the C1 chromosome, blue dots; *ori2* from the C2 chromosome, red dots). Black lines represent the parental chromosome, while gray lines indicate replicated DNA. During chromosome segregation, XerC/D recombinases mediate the first recombination at *dif* sites (black and gray triangles). A second recombination is catalyzed by Cre recombinase at *lox*P sites (blue and light blue triangles), facilitating complete resolution of chromosome dimers and proper segregation to daughter cells.

(*21*)—tolerated transposon insertions in WT and Cir strains but not in the fusion strains FL and FCir (Fig. 2I).

As reported by Liao *et al.* (*21*), the replication origins of both the circular and linear chromosomes, along with their partitioning loci, *parAB* and *repABC*, were essential in all four strains, indicating that they maintained the same replication and segregation program (figs. S4 and S5). The Cre recombinase used to circularize the fused chromosome became essential in the FCir strain but not in the Cir strain (Fig. 2J), as evidenced by the inability to cure the plasmid harboring the *cre* gene from FCir (fig. S2D). The original *cre* plasmid could only be removed when *cre* was supplied through an alternative source, such as a different plasmid (pEA384; table S1) or integrated into the FCir chromosome via the *telA* locus (fig. S6). These results suggest that both *XerCD/dif* and Cre-*lox*P systems are required to resolve the replicated circularized fused chromosome for proper segregation. As shown in Fig. 2J, FCir encounters a unique challenge during each cell cycle: Half of the cell population may inherit a chromosome that physically spans both daughter cells, obstructing cell division due to the presence of four unresolved DNA molecules (Fig. 2J). Two of these segments (Fig. 2J, black and gray triangles) can be resolved by *XerCD/dif*-mediated recombination (Fig. 2J, first recombination) (*21*, *22*), as in the FL strain. The remaining two DNA segments (Fig. 2J, blue and light blue triangles) require a second recombination event via the Cre-*lox*P system (Fig. 2J, second recombination). In this model, Cre-mediated resolution occurs between the residual *lox*P site left after chromosome circularization (Figs. 1F and 2D and fig. S1Aiii) and a second identical *lox*P site (Fig. 2J) produced during DNA replication, enabling proper chromosome segregation and preventing cell death.

Enhanced fitness and stress tolerance with single-chromosome architecture

We first tested how these four strains grow under different growth conditions to assess their fitness and stress tolerance. Using comparative growth assays, we determined the average doubling time in nutrient-rich yeast extract peptone (YEP) medium (Fig. 3A). Chromosome architecture significantly affected the growth rate, with single-chromosome strains (FL and FCir) exhibiting a 20 to 30% faster growth than WT and Cir strains (1.47 to 1.48 hours versus 1.82 to 2.08 hours; Fig. 3A and fig. S7A; paired *t* test, df = 3, $P < 0.05$). To examine whether this growth advantage persisted under nutrient-limiting conditions, we repeated the assay in *Agrobacterium* minimal (AB) medium (*31*), which more closely mimics the soil environment where *Agrobacterium* typically resides. In this setting, we compared the rate of change in optical density at 600 nm per hour (ΔOD_{600}/hour), measured at its maximum during the inflection point of each growth curve (*32*–*34*). Consistent with the YEP results, the single-chromosome strains exhibited superior growth in AB medium (Fig. 3B and fig. S7B), with the rate of change of 0.057 OD_{600}/hour for FL and 0.052 OD_{600}/hour for FCir, compared to 0.032 OD_{600}/hour for WT and 0.034 OD_{600}/hour for Cir. Statistical analysis using one-way analysis of variance (ANOVA) followed by Tukey's post hoc test (df = 3) revealed significant differences ($P < 0.0001$ for WT versus FL/FCir; $P < 0.001$ for Cir versus FL/FCir), indicating that the fitness advantage conferred by chromosome architecture is robust across distinct environmental conditions. Despite these differences in the growth rate, all strains remained fully viable (fig. S7, C and D).

Previously, it was reported that the 194-kb truncation on the pAtC58 plasmid resulted in slightly faster growth, likely due to a reduced metabolic burden (*30*). In agreement with this, we observed modest growth enhancement when we deleted the identical 194-kb segment on the pAtC58 in both WT and Cir strains using the INTEGRATE/Cre-*lox*P system (fig. S7E and table S1). In addition, growth curve analysis showed that single-chromosome strains exhibited a more rapid exponential growth phase than two-chromosome strains (fig. S7A). Thus, the faster growth of FL and FCir strains may be attributed to changes in chromosome architecture.

We then investigated whether the enhanced growth rate of single-chromosome strains conferred a fitness advantage during coculture with two-chromosome strains. Our simulations, based on the doubling times, predict that single-chromosome strains will outcompete two-chromosome strains over a 72-hour period (Fig. 3, C and D). Our competition assays confirmed that single-chromosome strains consistently outgrew two-chromosome strains in all combinations, ultimately dominating the population by the end of the assays (Fig. 3, E and F, and fig. S8, A and B). Consistent with the growth rate data, we observed no significant differences between WT and Cir (Fig. 3G and fig. S8C) or FL and FCir (Fig. 3H and fig. S8D) (Student's *t* test, $P > 0.05$), with both strains maintaining a 1:1 ratio throughout the assays. Similarly, when red fluorescent protein $(RFP)^+$ and RFP^- versions of the same strain were cocultured, they remained at a 1:1 ratio (fig. S8, E to H). These results suggest that the enhanced competitive advantage observed in single-chromosome variants over two-chromosome strains may be primarily attributable to their accelerated growth rate.

Considering the faster doubling times and competitive fitness advantages of the single-chromosome configurations, we investigated whether such advantages would persist under conditions that induce replication stress, such as reactive oxygen species (ROS) and ultraviolet (UV) irradiation. We assessed growth and survival in the presence of hydrogen peroxide that stimulates ROS generation. The single-chromosome strains (FL and FCir) showed smaller inhibition zones compared to the WT (fig. S9), suggesting an enhanced tolerance to ROS. However, it was the UV irradiation assay that showed a more profound difference: The single-chromosome strains consistently exhibited greater resilience to UV irradiation conditions (Fig. 3, I and J). Our data suggest that, despite requiring an additional recombination step to resolve chromosome segregation issues, the single-chromosome architecture in *A. tumefaciens* C58 confers a faster growth rate, competitive fitness advantages, and enhanced stress tolerance.

Reduced virulence with single-chromosome architecture

In nature, *A. tumefaciens* C58 strain exhibits a dual lifestyle, thriving as a free-living saprophyte or acting as a plant pathogen that induces crown gall tumors. This transition is regulated by gene expression changes in response to environmental cues (*35*). A key triggering agent produced by plants is phenolic compounds, such as acetosyringone (AS), which activate the virulence machinery in *Agrobacterium* (*36*). It is well established that soil-dwelling *Agrobacterium* strains harboring pTi plasmids become pathogenic upon detecting these phenolic signals typically released from wounded plant tissues, initiating the DNA transfer process that leads to tumor formation in host plants (*13*). To assess changes in the virulence of different chromosome architectures, we first performed a virulence (*vir*) gene induction assay. *Agrobacterium* regulates *vir* gene expression through the VirA/G two-component system (*37*). Under low pH conditions, AS is perceived by the transmembrane protein VirA (*38*), leading to the phosphorylation of VirG, which then activates downstream *vir*

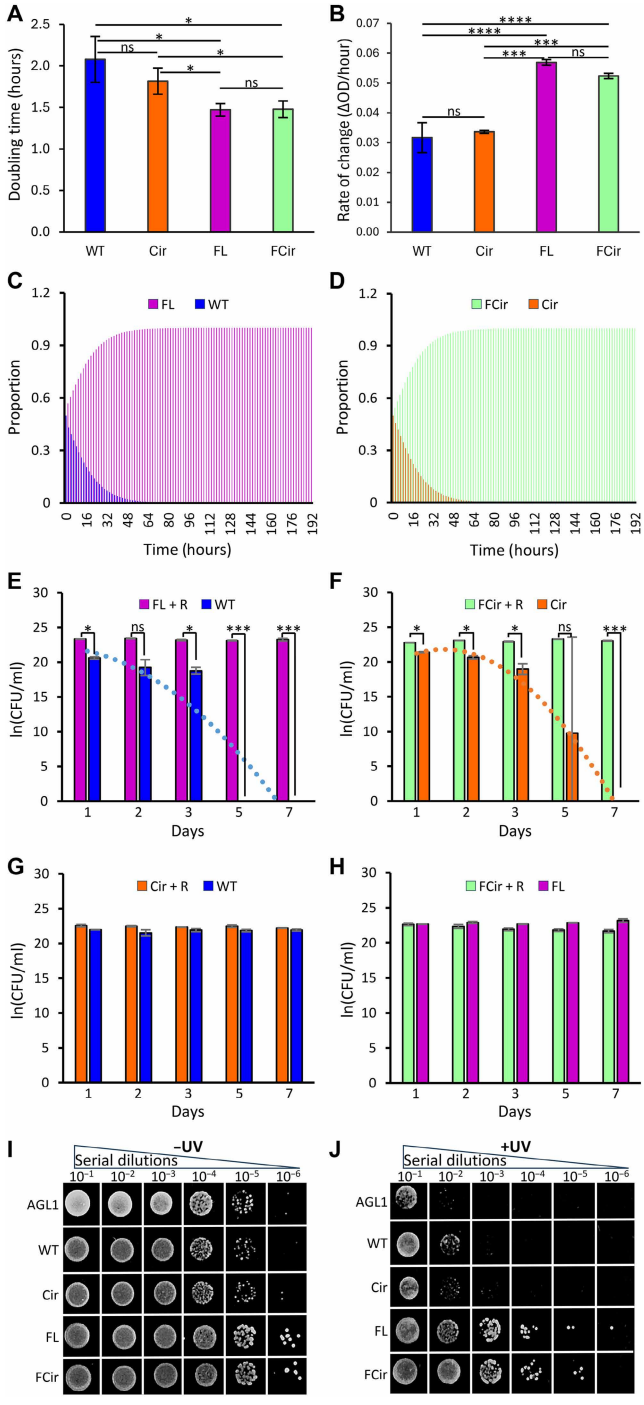


Fig. 3. Growth and stress tolerance of four *Agrobacterium* strains with different chromosome architectures. (**A**) Doubling times (hours) of single- and bipartite-chromosome C58 strains grown in YEP medium ($N = 12$ biological replicates per strain). (**B**) Rate of change in OD_{600} (ΔOD/hour) at the inflection point of growth curves from single- and bipartite-chromosome C58 strains grown in AB minimal medium ($N = 6$ biological replicates per strain; see fig. S7B for representative growth curve). Predicted changes in relative strain abundance over 7 days in pairwise coculture competitions, based on experimentally measured doubling times: (**C**) FL (linear fused chromosome) versus WT (circular C1 + linear C2) and (**D**) FCir (circularized fused chromosome) versus Cir (circular C1 + circularized C2). Strain competition experiments using RFP (*mCherry*)–tagged strains: (**E**) FL versus WT, (**F**) FCir versus Cir, (**G**) Cir versus WT, and (**H**) FCir versus FL. RFP-tagged strains are indicated "+R." Each bar represents a colony-forming units (CFU) count recorded on days 1, 2, 3, 5, and 7. The *y* axis shows the natural log-transformed CFU counts for each strain at days 1, 2, 3, 5, and 7. The *y* axis displays the natural log-transformed CFU values. Asterisks (*) indicate statistically significant differences between strains: in (A) and (E) to (H), by a two-tailed paired sample *t* test (***$P < 0.001$, and *$P \leq 0.05$; ns, not significant); in (B), by one-way ANOVA with Tukey's post hoc test (****$P < 0.0001$ and ***$P < 0.001$; no asterisk or not significant). Error bars represent the SD of the mean [(A) and (E) to (H)] or SEM (B). Survival of Agrobacterium strains in the absence (**I**) and presence (**J**) of ultraviolet (UV; 365 nm) irradiation. The *recA*-deficient AGL1 strain (table S1) was included as a DNA repair–deficient control.

genes (Fig. 4A) (*39*). We constructed a reporter plasmid pEA106c (table S1) containing a *mCherry* cassette under the control of the VirG-inducible *virB* promoter (PvirB) (*40*). This reporter plasmid was then introduced into all tested strains, and the fluorescence was measured with and without AS treatment to assess the *vir* gene induction (Fig. 4B).

Strains with two chromosomes (WT and Cir) showed significantly higher mCherry fluorescence than the single-chromosome variants (FL and FCir), indicating a reduced *vir* gene induction in the strains with fused chromosomes (Fig. 4C). In addition, we observed a progressive decrease in mCherry expression when the FCir strain was subcultured for 100 consecutive days (Fig. 4, D and E). For this assay, strains were grown daily in 5 ml of YEP liquid medium within 50-ml tubes, using a 1:1000 dilution in fresh YEP medium each day (24 hours). Every 10 days, bacterial cultures were collected, prepared as glycerol stocks, and stored at −80°C until further analysis. Following the 100-day subculture period, each stored bacterial culture was retrieved, made into competent cells, and transformed with the reporter plasmid pEA106c carrying the PvirB-*mCherry* cassette (Fig. 4B). The *vir* gene induction in AB medium was then quantified via mCherry fluorescence as described (see the "Fluorescence-based quantification of *vir* gene expression" section in Materials and Methods). As shown in Fig. 4 (D and E), FCir had a similar level of *mCherry* expression until day 50, after which it gradually declined, becoming undetectable by day 100. In contrast, the other three strains maintained a similar level of *mCherry* expression throughout the 100-day subcultures. Whole-genome sequencing (WGS) conducted after the 100th subculture confirmed that while WT, Cir, and FL maintained the Ti plasmid, the FCir strain lost pTiC58 (Fig. 4F). This suggests a potential trade-off between enhanced fitness and virulence. While the precise mechanism underlying this instability remains unclear, one possible contributing factor is the complex segregation dynamics associated with the FCir chromosome during cell division (Fig. 2J). The accelerated growth of the FCir strain, harboring a circularized fused chromosome, may further exacerbate this instability and contribute to the observed tendency to lose the Ti plasmid.

We further compared the virulence of the four *Agrobacterium* strains using *Kalanchoe* leaf tumorigenicity and *Nicotiana benthamiana* leaf infiltration assays. In the tumorigenesis assay conducted on 6- to 8-week-old *Kalanchoe* leaves, the two-chromosome strains produced larger tumors than the single-chromosome strains (Fig. 5A, compare "WT" and "Cir" with "FL" and "FCir"). Notably, the FCir strain produced the smallest tumors (Fig. 5A). Similarly, a quantitative transient transformation assay using the *RUBY* reporter (*41*, *42*) showed lower betalain content in *N. benthamiana* leaves infiltrated with single-chromosome strains compared to the WT strain (Fig. 5B and fig. S10), indicating reduced T-DNA delivery. A slight reduction in betalain content was also observed in *N. benthamiana* leaves infiltrated with the Cir strain compared to the WT strain (Fig. 5B; Student's *t* test, $P < 0.05$).

The impact of chromid circularization was more pronounced in EHA105 (E5 versus E5Cir in Fig. 5B; Student's *t* test, $P < 0.001$), a widely used disarmed laboratory strain with C58 chromosomal background and a hypervirulent pTiBo542 plasmid (*43*). Circularization of the linear C2 chromid in EHA105 (E5Cir in Fig. 5B) significantly reduced transient transformation efficiency compared to the unmodified EHA105 strain (Student's *t* test, $P < 0.001$), suggesting the critical role of chromosome architecture in modulating the virulence across *A. tumefaciens* strains. Morton *et al.* (*30*) previously reported that a 194-kb truncation in pAtC58 enhanced virulence by alleviating the metabolic burden of maintaining the cryptic plasmid. Similarly, our laboratory-generated 194-kb deletion at the same location in pAtC58 resulted in a slight increase in *vir* gene expression (fig. S7F). Therefore, the observed reduction in *vir* gene expression in the single-chromosome strains may be attributable to changes in chromosome architecture rather than effects associated with the pAtC58 plasmid.

We also performed stable transformation experiments using EHA105 and EHA105Cir strains to infect maize B104 immature embryos. Both strains harbor a disarmed pTiBo542 plasmid (*43*) and carry additional copies of *vir* genes from pTiBo542 on a ternary plasmid pKL2299A (*44*). Unlike the transient transformation differences observed in *N. benthamiana* (Fig. 5B), no significant differences were observed between the two strains in maize stable transformation (fig. S11). This result suggests that the presence of the ternary *vir* helper plasmid compensates for any reduction in intrinsic virulence caused by chromid modification in the EHA105Cir strain.

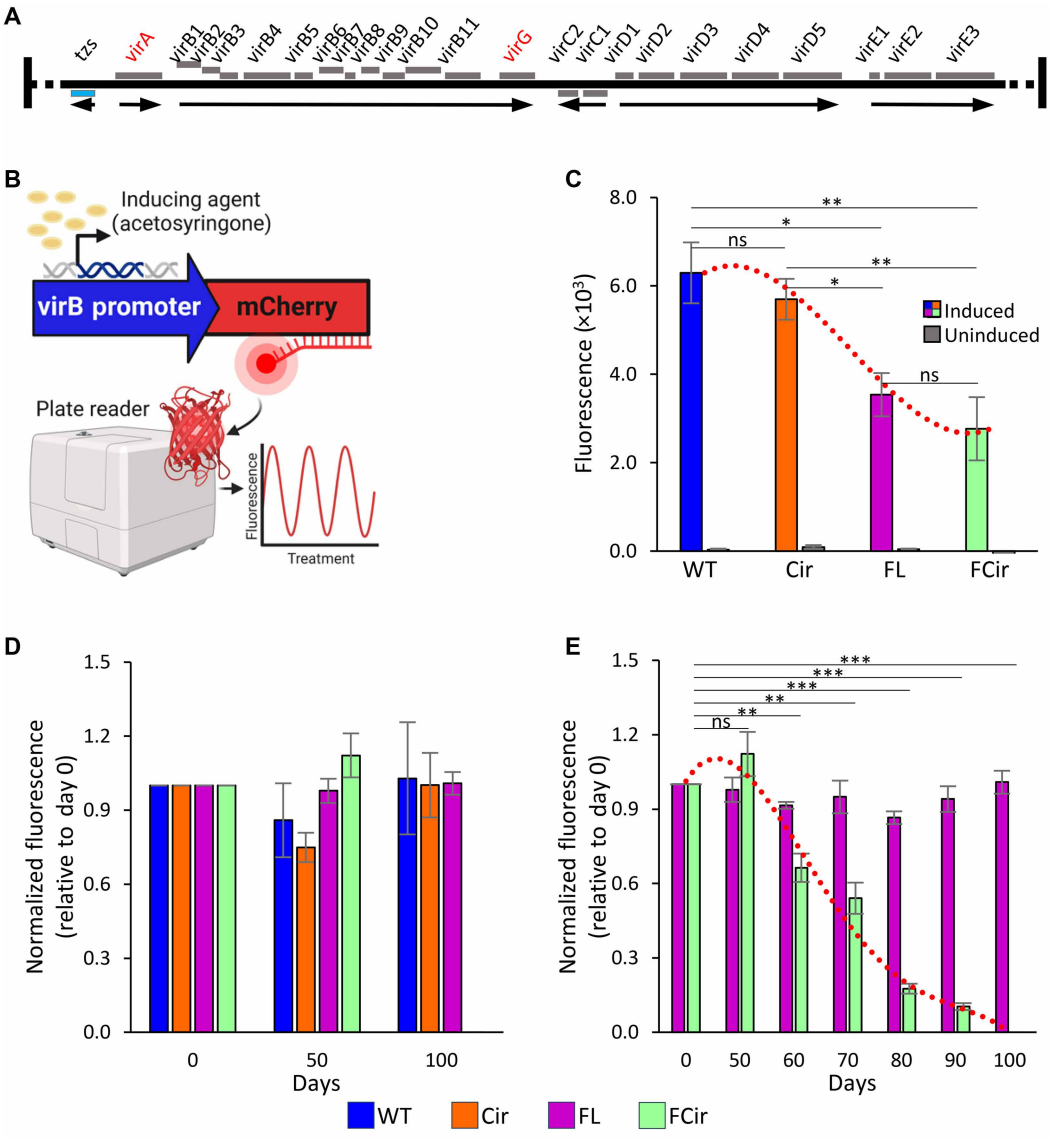
Together, our findings highlight the adaptive importance of chromosome architecture in *A. tumefaciens* strains. Among the variants tested, the WT chromosomal arrangement (WT C58), which consists of a circular chromosome and linear chromid, consistently exhibited the highest virulence. This suggests that the WT genome configuration optimizes Ti plasmid stability and regulatory dynamics that promote pathogenicity.

Distinct gene expression patterns associated with different chromosome architecture

We conducted a whole transcriptome sequencing analysis [RNA sequencing (RNA-seq)] to compare the overall gene expression patterns associated with the different chromosome architectures. A total of 24 RNA-seq samples were processed, comprising four strains, three biological replicates, and two experimental conditions, with or without the presence of AS (see Materials and Methods and table S2).

Normalized gene expression data showed that chromosome architecture had a significant impact on *Agrobacterium* virulence–related genes (Fig. 5, C to F; see the "RNA-seq data bioinformatic analysis" section in Materials and Methods). Consistent with the *in planta* virulence assay data (Fig. 5, A and B), WT strain had higher expressions than Cir, FL, and FCir strains for most genes related to *Agrobacterium* virulence, including the key *vir* gene regulator *virA*/*virG* two-component system (Fig. 5C), *virB* operon—containing the components of *Agrobacterium* type IV secretion system essential for the T-DNA delivery (Fig. 5D), chromosomally encoded *vir* genes *chvG/I* that senses changing pH (Fig. 5E) (*45*), and T-DNA processing genes *virC*, *virD*, and *virE* (Fig. 5F). Some individual *virB* genes showed minimal expression differences, with *virB8-11* remaining relatively unchanged across all strains. Meanwhile, *virB5* ($P < 0.05$), *virB6* ($P < 0.001$), and *virB7* ($P < 0.01$) exhibited moderately higher expressions in FL and FCir compared to WT and Cir strains. In contrast, *virB1-3* displayed significantly lower expression in FL and FCir relative to WT and Cir ($P < 0.0001$) (Fig. 5D). Overall, most *vir* gene expressions in these strains were consistently lower than the WT strain (Fig. 5, C to F). We further validated these findings by reverse transcription quantitative polymerase chain reaction (RT-qPCR) and observed similar patterns for the selected *vir* genes (fig. S12; see Materials and Methods).

Next, we assessed the impact of chromosome architecture on global gene expression by comparing differentially expressed genes



F

Strains	Replicons					Plasmidsaurus
	C1	C2	C1-C2 fusion	pAt	pTi	ID
WT	+	+	–	+	+	3GHFJW_1
Cir	+	+	–	+	+	46YNW4_1
FL (pAt Δ)	–	–	+	+	+	3GHFJW_2
FCir (pAt Δ)	–	–	+	+	–	46YNW4_2

Fig. 4. Virulence (*vir*) gene loci and the stability of Ti plasmid (pTiC58) of four *Agrobacterium* C58 strains with different chromosome architectures. (**A**) Schematic representation of the essential *vir* gene loci on the C58 Ti plasmid (pTiC58). The two-component regulatory system genes, *virA* and *virG*, are highlighted in red. (**B**) Schematic illustration of the quantification of AS-induced RFP (*mCherry*) gene expression driven by the VirG-inducible *virB* promoter using a fluorescent plate reader. (**C**) *mCherry* expression in WT, Cir, FL, and FCir strains, with or without AS induction. (**D**) Quantification of *mCherry* expression levels across all four strains. (**E**) Time course of *mCherry* expression in FL and FCir strains over 100 days of serial subculturing. Red dotted lines in (C) and (E) present third-order polynomial trendlines illustrating RFP induction over time. (**F**) Summary table of WGS results for all strains. C1, circular chromosome; C2, linear chromid; C1-C2 fusion, chromosome-chromid fusion; pAT, pAtC58 plasmid; pTi, pTiC58 plasmid; Plasmidsaurus ID, whole bacterial genome sequencing sample identifiers from Plasmidsaurus.

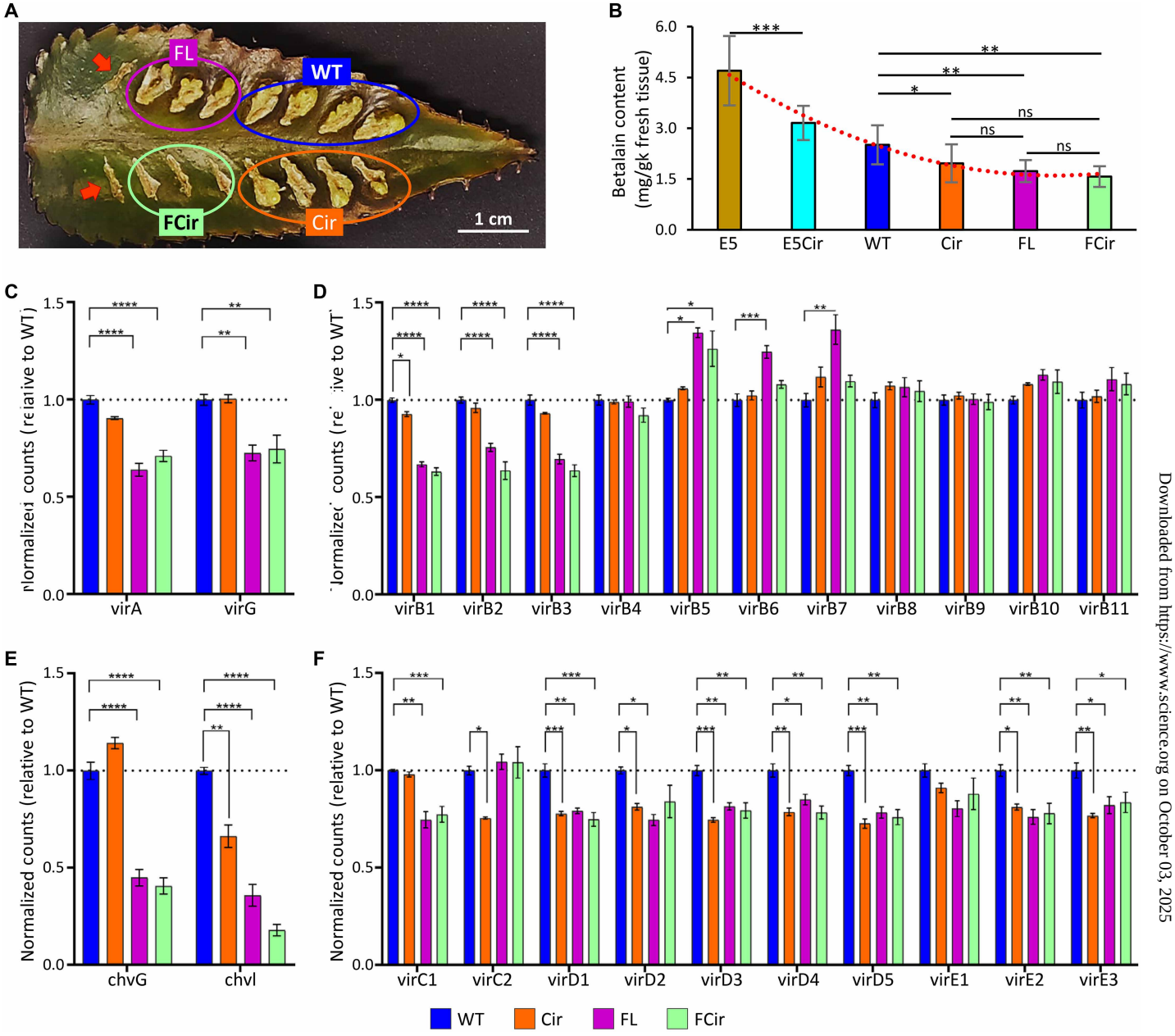


Fig. 5. Virulence (*vir*) gene expression and plant infection capacity of four *Agrobacterium* C58 strains with different chromosome architectures. (**A**) Tumorigenicity assay on *Kalanchoe* leaves for WT (blue), Cir (orange), FL (pink), and FCir (green) strains. Red arrows indicate mock inoculation with water. Eight independent inoculation experiments were conducted, each using two leaves from separate plants. (**B**) Quantitation of transient *RUBY* reporter expression in *N. benthamiana* leaves 2 days after agroinfiltration ($N = 10$ biological replicates per strain; see also fig. S11). Betalain content was used as a proxy for T-DNA transfer efficiency. (**C** to **F**) RNA-seq analysis of virulence-related gene expression across strains: (C) *virA* and *virG* (two-component regulators on pTi); (D) *virB* operon; (E) *chvG* and *chvI* (two-component system on chromosome C1); and (F) *virC*, *virD*, and *virE* genes. Asterisks (*) indicate statistical significant differences: in (B), by a two-tailed paired sample *t* test (***$P < 0.001$, **$P \leq 0.01$, and *$P \leq 0.05$; ns, not significant); in (C) to (F), by one-way ANOVA followed by Dunnett's multiple comparisons test [(C) to (F)] (****$P < 0.0001$, ***$P < 0.001$, **$P \leq 0.01$, *$P \leq 0.05$; no asterisk, not significant). RNA-seq was performed with three biological replicates per condition ($N = 3$). Error bars represent SD (B) or SEM [(C) to (F)].

(DEGs), performing principal components analysis (PCA), and identifying strain-specific expression patterns using weighted gene coexpression network analysis (WGCNA) (*46*). A gene was classified as differentially expressed if there was a minimum of twofold change in expression levels between the two conditions [AS or control solvent dimethyl sulfoxide (DMSO); false discovery rate–adjusted $P < 0.05$]. Overall, WT and Cir exhibited the most dynamic transcriptome, containing the highest number of DEGs, while FCir had the greatest impact on gene induction but contained the fewest DEGs (Fig. 6A and fig. S13A). The PCA grouped WT with Cir and FL with FCir (fig. S13B), suggesting that gene expression profiles are more strongly influenced by chromosome number (i.e., bipartite versus monopartite)

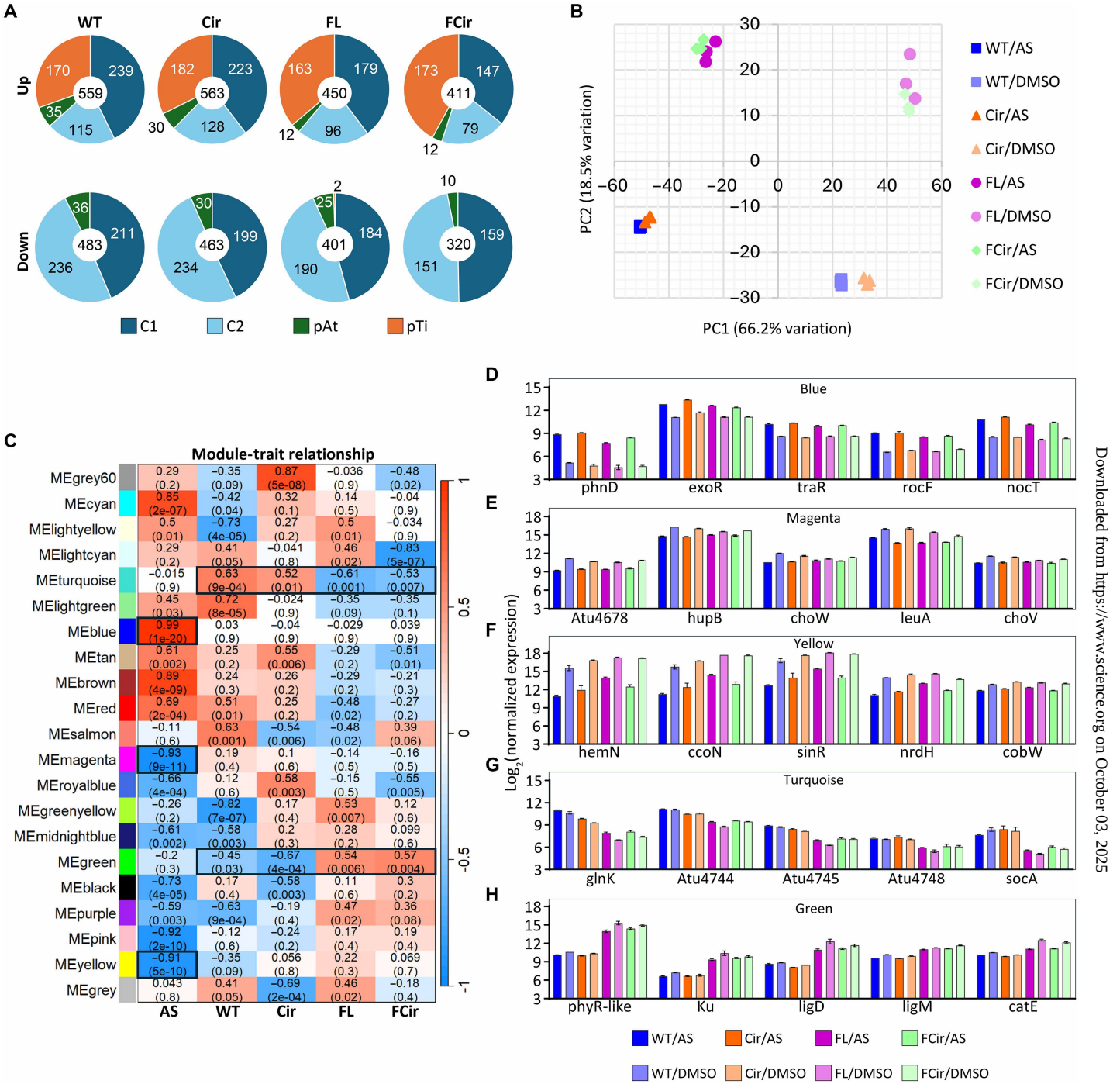


Fig. 6. Transcriptomic analysis of four *Agrobacterium* strains with different genome architectures. (**A**) Gene expression analysis under AS-induced conditions compared to the non-induced DMSO control. The number of DEGs (≥2-fold change, $P \leq 0.05$) is shown for each replicon (C1, C2, pAt, and pTi). Pie chart centers indicate the total number of up- and down-regulated genes. Up, AS-induced genes; down, AS-repressed genes; WT, wild-type C58; Cir, C58 with circularized C2 chromid; FL, C58 with a fused linear chromosome; FCir, C58F with a circularized fused chromosome. (**B**) PCA plot of transcriptomes from all strains and conditions. (**C**) WGCNA showing module-trait correlations with AS induction and chromosome architectures. Pearson correlation coefficients are shown at the top of each cell, with *P* values below. Five modules with significant associations (Pearson $r > 0.50$ or < -0.45, $P < 0.05$) are highlighted with black outlines and designated as modules of interest (MOIs). (**D** to **H**) Normalized expression levels of selected genes of interest from each MOI with high module membership (MM > 0.9, $P < 0.05$). Genes with higher expression under the AS condition relative to DMSO are considered as up-regulated, whereas those with higher expression in DMSO relative to AS are considered as down-regulated. Additional genes with MM > 0.9 (e.g., *vir* genes) are presented in fig. S17.

than by their shape (i.e., linear versus circular). Since the fused chromosome strains also carried a 194-kb deletion on pAt plasmid, we ran a separate PCA analysis after excluding the genes encoded within the deletion region. The second PCA analysis also produced identical groupings (Fig. 6B), confirming that it was the chromosome architecture that contributed to gene expression differences among the strains.

Most up-regulated genes were localized to the chromosomes and the Ti plasmid (pTi) (Fig. 6A, top, and fig. S14, A to D). Conversely, most down-regulated genes were located on the chromosomes, with very few genes showing notable modulation on the accessory plasmid (pAt) (Fig. 6A, bottom, and fig. S14, E to H). Analysis of DEGs across all four strains identified 270 genes consistently up-regulated, representing 33% of the total up-regulated genes in all replicons (fig. S14, A to D). Similarly, 161 genes were down-regulated, accounting for 21.5% of down-regulated genes within all replicons (fig. S14, E and F). Kyoto Encyclopedia of Genes and Genomes (KEGG)–based functional classification of DEGs (*47*) revealed that genes involved in metabolism, genetic information processing, signaling, cellular processes, and organismal processes were differentially expressed with some strain-specific patterns (fig. S15 and table S3).

The WGCNA identified a total of 21 modules with distinct coexpression patterns under different conditions (AS or control solvent DMSO) or across four strains, with module sizes (the number of genes included in a module) ranging from 70 to 699 genes (Fig. 6C and table S3). Among these, five modules of interest (MOIs) were selected on the basis of their correlation patterns. The blue, yellow, and magenta modules showed consistently weak associations across individual strains; however, the blue module was strongly positively correlated with the AS treatment, while the yellow and magenta modules were negatively correlated. The remaining two modules were selected for their contrasting expression patterns between strains: The turquoise module contains genes with higher expression in WT and Cir, whereas the green module includes genes more highly expressed in FL and FCir. These five MOIs are highlighted with boxes in Fig. 6C.

Representative genes illustrating the module's expression pattern from filtered genes (Fig. 6C and fig. S16, A to E) are shown in Fig. 6 (D to H). Module blue (Fig. 6D) represents genes highly induced by AS, including well-known *vir* genes required for plant infection and T-DNA transfer (e.g., *virB1*; Fig. 5D, figs. S16F and S17A, and table S3). Also included in this module are *phnD*, *rocF*, and *nocT*, which function in nutrient uptake from plant tumors (*48*–*50*), and *exoR* and *traR*, which play important roles in biofilm formation and Ti plasmid conjugation, respectively (*51*, *52*). In contrast, modules magenta (Fig. 6E, fig. S17B, and table S3) and yellow (Fig. 6F, figs. S16G and S17C, and table S3) contain genes whose expression was repressed by AS and includes genes known to be down-regulated in plant tumors (e.g., *hupB*, *choW*, *Atu4678*, *nrdH*, and *cobW*) (*53*–*57*).

Modules turquoise and green include genes with strain-specific expression patterns. In particular, module turquoise (Fig. 6G) represents genes that had higher expression in the strains with two chromosomes (WT and Cir) than those with a fused single chromosome (FL and FCir) and includes *Agrobacterium* genes that are known to be up-regulated in plant tumors and may function in host metabolite acquisition (e.g., *glnK* and *Atu4744-Atu4748*; fig. S17D and table S3) (*58*), possibly supporting pathogenic adaptation of the bipartite chromosome strains.

On the other hand, genes included in module green (Fig. 6H, fig. S17E, and table S3) had higher expression in the strains with fused single chromosomes (FL and FCir) than those with two chromosomes (WT and Cir). Representative members include stress response, DNA damage, and catalase genes (e.g., *phyR-like*, *ligD*, *Ku*, and *catE*, respectively) (*59*–*62*), consistent with our observations of enhanced stress tolerance of monopartite chromosome strains (Fig. 3, H and I, and fig. S18). *ligM*, a gene implicated in the degradation of phenolic compounds such as AS (*63*), was also included in the green module, potentially contributing to the lower virulence gene expression observed in the single chromosome strains.

Beyond the five major coexpression modules, RNA-seq data further revealed that the structural genes of the type VI secretion system, an important antibacterial weapon of *Agrobacterium* (*64*) and predominantly represented in the purple module, had a statistically significant up-regulation in the FL and FCir strains compared to the WT and Cir strains (fig. S18).

Together, these data clearly demonstrate that chromosome architecture drives distinct gene expression patterns. Strains with a fused single chromosome exhibited reduced expression of virulence-related genes but increased expression of stress response– and interbacterial competition–related genes. In contrast, the WT strain with two chromosomes showed higher *vir* gene expression, consistent with a more pathogenic phenotype.

DISCUSSION

The influence of chromosome architecture on bacterial adaptation and evolution remains a compelling but largely unresolved question. While diverse prokaryotic genome structures have been documented (*6*–*9*), their functional implications for bacterial physiology, fitness, and virulence are not well understood. Our study directly investigates how genomic organization affects bacterial behavior by integrating natural chromosomal variants with synthetic biology approaches in *A. tumefaciens* strain C58. By using precise genome engineering tools, we generated near-isogenic strains with distinct chromosome architectures, creating an ideal model system to assess the direct impact of genome structure on growth, stress tolerance, interstrain competition, and virulence.

Our findings highlight the profound effects of chromosome architecture on bacterial fitness and behavior. Single-chromosome strains exhibited enhanced replication efficiency, faster growth, and greater stress tolerance, suggesting that genome simplification through chromosome fusion improves overall bacterial adaptability as a saprobe. This observation aligns with the up-regulation of genes involved in DNA replication and metabolic pathways, which likely contribute to their physiological advantages. In addition, the increased competitiveness of single-chromosome strains in coculture suggests that genomic organization influences bacterial interactions, possibly through mechanisms such as the type VI secretion system (*64*).

Despite these fitness benefits, chromosome fusion and circularization led to a reduction in virulence in the single chromosome variants (FL and FCir). Certain *virB* genes within the operon appear to be selectively regulated across all isogenic variants (e.g., *virB1-3* and *virB5-7*), potentially due to posttranscriptional regulation or differences in RNA stability. Differential polycistronic mRNA decay is a conserved mechanism in bacteria that could drive selective regulation of genes encoded in an operon (*65*). In addition, posttranscriptional regulation is influenced by host-derived

factors, such as ribosome-inactivating proteins and antimicrobial peptides (*66*, *67*), and may modulate *Agrobacterium* C58 translation including *virB* transcripts, thereby influencing the type IV secretion system function. Hence, further studies dissecting the regulatory elements and their mechanism for the posttranscriptional regulation of *virB* operon will be essential to understanding these observations. Beyond this, the bipartite genome structure of C58 appears optimized for its pathogenic lifestyle, facilitating precise regulation of virulence genes. The progressive decline in virulence gene expression from WT to FCir (Fig. 5) suggests that genome simplification shifts the bacterium toward a more saprophytic state. This trade-off between fitness and virulence implies that chromosome architecture plays a crucial role in balancing adaptation to both free-living and host-associated environments.

The FCir strain exhibited a tendency to lose the Ti plasmid over successive generations, suggesting that chromosome architecture also influences the plasmid stability. Ti plasmid is known to be very stable in WT *Agrobacterium* strains without selection, due to intrinsic replication machinery (repABC) (*68*, *69*) and toxin-antitoxin system (*70*). Hence, it is normally very difficult to cure the Ti plasmid. While the precise mechanism underlying this instability remains unclear, it may be influenced by the complex segregation dynamics of FCir during cell division, as this strain has a single circular chromosome with two functional origins of replication. This finding supports the broader notion that bacterial genome organization is shaped by evolutionary pressures balancing fitness, virulence, and genetic stability (*5*). While the extent to which these findings apply to other bacterial species remains an open question, parallels observed across taxa suggest that chromosome organization may be a widespread determinant of bacterial evolution such as *S. meliloti* (*23*), *V. cholerae* (*24*), and *C. necator* (*25*).

Despite these insights, several questions remain. While we established a clear link between chromosomal architecture and bacterial behavior, the underlying molecular mechanisms require further investigation. For example, it remains unclear whether replication dynamics differ between the monopartite and bipartite chromosome strains despite the presence of two functional replication origins. In addition, while competition assays demonstrated fitness advantages in liquid culture, in planta studies will be essential to determine how chromosomal modifications affect plant colonization and tumor formation. Moreover, our study focused on a single *Agrobacterium* strain, raising the broader question of whether similar genome-architecture impacts exist in other bacteria with multipartite genomes.

Future research will therefore explore the molecular basis of the observed fitness differences, including replication timing, gene regulation, epigenetic modifications, and chromatin organization. In planta competition assays will be essential to assess the ecological success of *Agrobacterium* variants with different chromosome architectures. Overall, our findings demonstrate that chromosome architecture is not merely a structural feature but a key driver of bacterial physiology, fitness, and pathogenic potential. The engineered C58 variants in this study provide a valuable model for studying the evolutionary and functional consequences of genome organization. Understanding how these genomic configurations shape bacterial evolution and host interactions, especially in agricultural and ecological contexts, will be crucial. This work, therefore, deepens our understanding of microbial adaptation with broad relevance to basic science and biotechnology applications.

MATERIALS AND METHODS

Experimental design

This study investigates the functional importance of genome architecture in the biovar 1 *A. tumefaciens* strain C58, the most widely used chromosomal background in plant biotechnology. First, using a CRISPR RNA-guided integrase system, we engineered isogenic C58 variants with distinct chromosomal configurations: (i) WT C58 (circular chromosome + linear chromid), (ii) C58 circularized (Cir, circular chromosome + circular chromid), (iii) C58F (FL, fused linear chromosome), and (iv) C58F circularized (FCir, fused circular chromosome). Second, to assess the biological impact of these configurations on C58's dual lifestyles as a free-living saprobe and a plant pathogen, we conducted a series of physiological analyses, including growth dynamics, stress tolerance, interstrain competition, and plant infection efficiency. Third, we carried out detailed transcriptomic profiling to evaluate global gene expression differences among the genome variants and validate physiological findings. This three-step experimental design was intended to provide a comprehensive pilot framework of how chromosomal architecture may influence *Agrobacterium*'s ecological adaptability and host range potential. All experiments were conducted with appropriate replicates, and data acquisition was performed in a blind manner to minimize bias. No data were excluded from analyses.

Bacterial strains and growth conditions

A. tumefaciens strains used in this study are listed in table S1. When necessary, liquid media were supplemented with antibiotics at the following concentrations: kanamycin (50 mg/liter), spectinomycin (100 mg/liter), and gentamycin (50 mg/liter). *Escherichia coli* DH5α, used for plasmid cloning, was cultured at 37°C. *Agrobacterium* strains were cultured at 28°C, while AS induction treatments were carried out in a shaker incubator at laboratory ambient temperature (25°C). All cultures were incubated with shaking at 200 rpm. Cloning reactions were transformed into chemically competent *E. coli* DH5α cells, which were then cultured in Terrific Broth supplemented with the appropriate antibiotics. Plasmid DNA was extracted using the QIAprep Spin Miniprep Kit (QIAGEN, Hilden, Germany) or the Miniprep Kit from New England Biolabs (NEB, Ipswich, MA, USA), and constructs were validated by Sanger sequencing at the Iowa State University DNA facility (Ames, IA, USA).

Vector construction

All plasmid constructs and oligonucleotides used in this study are listed in tables S1 and S4, respectively.

Construction of single and multiplex targeting constructs

Derived from the type 1-F CRISPR-Cas system of *V. cholerae* (*27*, *28*), the INTEGRATE system uses a TniQ-Cascade complex and transposase proteins TnsABC to mediate sequence-specific cargo DNA integration (*28*). Target recognition is guided by a 5′-CC-3′ protospacer adjacent motif (PAM) and a 32-nt protospacer sequence, with cargo DNA integration occurring approximately 48 to 52 bp downstream of the protospacer sequence (*28*). The *V. cholerae* INTEGRATE *lox*P-based plasmid constructs (pEA258 and pEA297; table S1) were derived from the parental vector pKL2310 (*29*). *mCherry*-based INTEGRATE vectors (pEA286, pEA291A, and pEA302; table S1) were constructed using the pEA186 backbone (*29*), while all TtonB (*tonB* terminator)–based cargo vectors (pEA385 to pEA389) were generated using the pEA244 vector backbone (table S1). The

pVS1-based INTEGRATE vector pEA244 was developed by ligating a ~15-kb *Xho*I and *Pst*I fragment from pKL2310 to a 32-bp phosphorylated and annealed TtonB oligonucleotides duplex with compatible sticky ends. For spacer cloning across all vectors, complementary oligonucleotide pairs were phosphorylated, annealed, and used for ligation. All CRISPR RNAs (crRNAs) were designed with 32-nt spacers targeting genomic sites adjacent to a 5′-CC-3′ PAM sequence. Double- or triple-targeting CRISPR arrays were assembled by combining two or three annealed oligonucleotide duplexes with compatible sticky ends and ligating them into a *Bsa*I-digested vector backbone in a single ligation reaction.

Construction of C58FCir-cre recombinase curing vector pEA384

C58FCir was resistant to sucrose-mediated *sacB* counter-selection for curing the Cre recombinase vector pKL2315 (*29*) following circularization. To evaluate whether Cre expression was essential for the viability of C58FCir, the construct pEA384 (table S1) was generated. To clone pEA384, pEA106 (*29*) and pKL2315 were first digested with *Kpn*I and *Hind*III. The 6268-bp backbone from pEA106 with a spectinomycin resistance gene was then ligated with the 1096-bp Cre recombinase gene fragment excised from pKL2315.

Construction of INTEGRATE Cre cargo vector pEA393

The pEA393 vector (table S1) harboring a Cre recombinase cargo was used to generate the C58FCir-cre strain. To clone the Cre recombinase as a transposon cargo, a 1232-bp Cre recombinase cassette was PCR-amplified using the primer pair 393-seq-F1/R1 (table S4). This cassette contained the constitutive J23107 promoter, the *cre* coding sequence, and the *rrnB* gene terminator (J23107-cre-TrrnB). This amplified fragment was then assembled into *Xho*I and *Pst*I-digested pKL2310 using Gibson assembly (*71*). PCR amplification was performed with Q5 High-Fidelity DNA polymerase (NEB) following the manufacturer's instructions.

Construction of pEA394

The pEA394 vector (table S1) contains partial fragments of *xerC*, *xerD*, *repB*, and *parB* genes and was used to assess the essentiality of these genes in C58FCir. For cloning, five primer pairs with overlapping sequences were used to PCR amplify the respective fragments: 394-seq-F1/R1 (pBBR1-Kanamycin backbone), 394-seq-F2/R2 (*xerC*), 394-seq-F3/R3 (*xerD*), 394-seq-F4/R4 (*repB*), and 394-seq-F5/R5 (*parB*) (table S4). The pSL1142 vector (*28*) served as the template for amplifying the pBBR1 origin of replication and kanamycin resistance cassette, while genomic DNA (gDNA) from *A. tumefaciens* C58 was used as the template for other gene fragments. The five PCR products were assembled using Gibson assembly to generate the final pEA394 vector.

Preparation of *Agrobacterium* electro-competent cells

Electro-competent *Agrobacterium* cells were prepared following a previously described protocol (*29*). Briefly, *Agrobacterium* cultures were grown overnight in YEP medium to an early stationary phase, diluted 1:100 into fresh medium, and incubated until reaching an OD_{600} of 0.4 to 0.6. Cells were then chilled on ice for 20 min and harvested by centrifugation at 4000*g* for 10 min at 4°C. The cell pellets were washed four times with prechilled sterile distilled water to remove residual electrolytes that could interfere with electroporation. After the final wash, cells were resuspended in 1 ml of prechilled 10% sterile glycerol and centrifuged at 4000*g* at 4°C for 5 min. The supernatant was removed, and the final cell pellets were resuspended in the desired volume of prechilled 10% sterile glycerol. Aliquots (40 μl) were flash-frozen in liquid nitrogen and stored at −80°C until use.

Introduction of INTEGRATE vectors into *A. tumefaciens*

The INTEGRATE vector DNAs used in this study were introduced into *Agrobacterium* by electroporation for use in the transposon integration. For each transformation, 50 to 100 ng of plasmid DNA was electroporated into 40 μl of electrocompetent cells using 2-mm cuvettes and a Bio-Rad Gene Pulser (Bio-Rad Laboratories, Hercules, CA, USA). Following electroporation, cells were recovered in 0.5 ml of Super Optimal Broth with Catabolite repression medium or Lysogeny Broth (LB) medium and incubated at 28°C with shaking (200 rpm) for 2 hours. Recovered cells were plated on YEP agar with the appropriate antibiotics and incubated at 28°C for 48 hours to obtain parent colonies. Colony purification was then carried out as described by Aliu *et al.* (*29*). Briefly, all parent colonies were pooled, resuspended in YEP broth, and incubated overnight at 28°C with shaking. After 12 to 16 hours of growth, the culture was serially diluted 10^6-fold, and 100 μl of the final dilution was plated on YEP agar with appropriate antibiotics and incubated at 28°C for 48 hours.

PCR-based detection of clonally integrated colonies

Colonies resulting from the transposition assays were screened by colony PCR using genome-specific primers flanking the target integration site to confirm the integration of the transposon cargo into the *A. tumefaciens* genome. PCR reactions were performed in 12 μl of reaction volumes using either Taq or Q5 High-Fidelity DNA polymerase (NEB) master mix (2×), with 0.5 μM of each primer and 4 μl of *A. tumefaciens* cells resuspended in UltraPure water as the template. Additional genome- and cargo-specific primers were used to determine the orientation of transposon integration [target site adjacent to the right end of the transposon (T-RL), or target site adjacent to the left end of the transposon (T-LR); table S4]. Thermocycling was conducted according to a previously established protocol (*29*), with initial denaturation at 98°C for 30 s; followed by 30 cycles of 98°C for 10 s, primer-specific annealing temperature for 30 s, and 72°C for 30 s; and a final extension at 72°C for 5 min. PCR amplicons were resolved by electrophoresis on 1 to 1.3% agarose gels stained with RedSafe (Bulldog Bio, Portsmouth, NH, USA). Colonies yielding the expected amplicon sizes were further analyzed via Sanger sequencing after purification using the QIAquick PCR purification kit (QIAGEN).

INTEGRATE-Cre-*lox*P–mediated circularization of linear replicons

INTEGRATE-Cre-*lox*P–mediated circularization was performed in two steps. In the first step, a *lox*P-containing transposon cargo was inserted into preselected target sites using the INTEGRATE system. A 34-bp *lox*P sequence was cloned into the transposon cargo, and two crRNAs targeting the subtelomeric regions of the linear replicon were incorporated into the CRISPR array of the INTEGRATE vector (pKL2310), which is based on the pVS1 backbone. The resulting construct, pEA258, was electroporated into *Agrobacterium* strains EHA105, C58, and C58F. Following transformation, colonies were purified as described in the "Introduction of INTEGRATE vectors into *A. tumefaciens*" section. After incubation for 2 days at 28°C, colony PCR was used to identify clones with successful site-specific integration. Colonies with dual-targeted insertion events were further confirmed by Sanger sequencing.

In the second step, electro-competent cells of the dual *lox*P insertion mutants (with insertions in the same orientation, either T-RL or T-LR) were prepared and electroporated with a second pVS1-based

plasmid, pKL2315. This vector carried a kanamycin resistance gene and constitutively expressed the Cre recombinase under the control of the J23107 promoter. The kanamycin marker allowed for the selection of transformants and simultaneous counter-selection against the original INTEGRATE vector, which conferred spectinomycin resistance. To identify *Agrobacterium* colonies that successfully circularized their linear replicon, outward PCR was performed using a forward primer from target site 2 and a reverse primer from target site 1, as illustrated in fig. S1. Negative control PCRs using gDNAs from corresponding WT strains and target-specific primers (table S4) were included to confirm specificity. Dual *lox*P insertion sites were rechecked to confirm circularization, and precise recombination events at the telomeric ends were confirmed via Sanger sequencing. WGS analysis, described below, was also performed for final confirmation of chromosomal circularization.

Agrobacterium WGS analysis

For WGS, approximately 5×10^9 exponentially growing *Agrobacterium* cells cultured in YEP broth were harvested by centrifugation, and the supernatant was discarded. The cells were washed with 1 ml of phosphate-buffered saline, pelleted again, and resuspended in 0.5 ml of Zymo 1× DNA/RNA Shield (Zymo, Irvine, CA, USA). Samples were then submitted to Plasmidsaurus (Eugene, Oregon, USA) for sequencing using Oxford Nanopore Long-Reads sequencing technology (Oxford Nanopore Technologies, Oxford, UK). The resulting reads were mapped to the *A. tumefaciens* C58 reference genome [National Center for Biotechnology Information (NCBI) genome assembly ASM9202v1] and analyzed using Geneious R11 (www.geneious.com) and SnapGene software (www.snapgene.com).

Determination of chromosomal stability in circularized mutants

To assess chromosomal stability of the circularized mutant strain, a single colony was streaked onto fresh YEP agar plates through 20 consecutive passages. Outward PCR was conducted after every fifth passage to verify the retention of the circularized chromosomal junctions. In addition, a 100-day successive subculture experiment was performed, during which outward PCR was carried out every 10th subculture. To further confirm chromosomal stability, selected PCR products were subjected to Sanger sequencing, and WGS was conducted on cells collected at the end of the experiment.

Growth and cell viability assay

For growth and cell viability assays, each strain was inoculated in 3.5 ml of YEP medium in 50-ml conical tubes and incubated in a shaking incubator (28°C, 200 rpm) for 12 to 18 hours to generate seed cultures. These seed cultures were then diluted into either YEP broth or *Agrobacterium* minimal (AB sucrose) medium (pH 7.0; 31) to an OD_{600} of 0.01 (YEP) or 0.02 (AB sucrose) in 250-ml flasks. Batch cultures were grown under the same incubation conditions, and OD_{600} was measured every 3 hours for 24 hours in the YEP assay or periodically over 72 hours during the AB sucrose assay. A 10-fold dilution was applied to cultures after 12 hours to ensure accurate OD_{600} readings. Growth rates for YEP medium ($N = 12$ biological replicates per strain) were calculated on the basis of changes in cell density following the methods from Aliu *et al.* (*29*, *43*). Statistically significant differences between strains were determined by a two-tailed paired sample *t* test.

For growth in AB minimal medium, the Gompertz model (*32*) was fit to the 0- to 49-hour growth data for each replicate ($N = 6$ biological replicates per strain) using GraphPad Prism v10.4.1 (GraphPad Software, Boston, MA, USA). The maximum rate of change in OD_{600} was determined at the inflection point of each growth curve with the expression bk/*e* (*33*, *34*), which was rewritten as $(Y_M - Y_0)(K)/e$. A one-way ANOVA followed by Tukey's multiple comparisons test was performed using GraphPad Prism v10.4.1. To determine viable cell counts, 100 μl of serially diluted cultures was plated onto YEP agar and incubated for 24 to 48 hours, and colony-forming units (CFU per milliliter) were calculated.

Interbacterial competition assay

Agrobacterium strain competition assays were carried out by integrating a RFP (*mCherry*) expression cassette into the genome of competing strains using the INTEGRATE vectors pEA291A and pEA302 (table S1), as previously described (*29*). To control for any potential metabolic burden associated with *mCherry* expression, two experimental setups were used in which the fluorescent marker was swapped between strains in half of the competition assays. Competing strains were mixed in a 1:1 ratio at an initial OD_{600} of 0.01 and grown overnight in YEP medium. This process was repeated over seven serial passages. At passages 1, 2, 3, 5, and 7, cultures were plated to estimate the population of the competing strains via their respective CFU, tracking the competition dynamics. Relative fitness was calculated by plotting the natural logarithm of the ratio of the two competing strains over the number of generations. Statistical significance was determined using a two-tailed Student's *t* test.

Fluorescence-based quantification of *vir* gene expression

To quantify *vir* gene expression, *Agrobacterium* strains were transformed with a reporter plasmid (pEA106 or pEA106c; table S1) containing *mCherry* under the control of an inducible promoter (PvirB) from the *virB* operon. Because EHA105 harbors the Bo542 Ti plasmid (pTiBo542), pEA106 with the corresponding PvirB promoter was introduced into EHA105 and its circularized mutant. In contrast, pEA106c, which contains the PvirB promoter from the C58 Ti plasmid (pTiC58), was used for the C58 and C58F strains and their respective mutants.

To prepare culture for *vir* induction, a single colony from each *Agrobacterium* strain was inoculated into 3.5 ml of liquid YEP medium in a 50-ml conical tube and grown at 28°C with shaking at 200 rpm until reaching early stationary phase. Cultures were harvested by centrifugation and resuspended in AB induction medium (pH 5.5) (*31*) to an OD_{600} of 0.2. For *vir* gene induction, AS was added to the resuspended cultures at a final concentration of 200 μM, while the control culture received an equivalent volume of DMSO. The cultures were then incubated at 25°C with shaking at 200 rpm for 21 hours. After incubation, fluorescence was measured using a fluorometer equipped with a ROX/Texas Red filter set (excitation at 585 nm, emission at 610 nm) to quantify *mCherry* expression.

High-throughput chromosome conformation capture (Hi-C)

The Hi-C procedure was performed as previously described (*21*, *72*). *A. tumefaciens* strains were cultured to exponential phase (~0.6 OD_{600}) in ammonium tryptone glucose NaCl (ATGN) broth at 30°C (*30*). Cells were cross-linked with 3% formaldehyde for 30 min at room temperature (22°C), and the reaction was quenched with 125 mM

glycine for 5 min. Cross-linked cells were pelleted and resuspended to an OD_{600} of 1.2 and then lysed with ready-lyse lysozyme, followed by treatment with 0.5% SDS. Chromatin was solubilized and digested with *Hind*III for 2 hours at 37°C. Digested DNA ends were filled in with Klenow DNA polymerase and Biotin-14–deoxyadenosine triphosphate, deoxyguanosine triphosphate, deoxycytidine triphosphate, and deoxythymidine triphosphate. The biotinylated fragments were ligated with T4 DNA ligase at 16°C for ~20 hours. Cross-link was then reversed at 65°C for ~20 hours in the presence of EDTA, proteinase K, and 0.5% SDS. DNA was purified by two rounds of phenol/chloroform/isoamyl alcohol (PCI; 25:24:1, v/v/v) extraction, followed by ethanol precipitation and resuspension in 20 μl of 0.1× tris EDTA buffer (pH 8.0). To remove biotin from nonligated DNA ends, samples were treated with T4 DNA polymerase at 20°C for 4 hours, followed by PCI extraction. DNA was then sheared for 12 min at 20% amplitude using a Qsonica Q800R2 water bath sonicator (Qsonica, Newtown, CT, USA). Libraries were prepared using the NEBNext UltraII DNA Library Prep Kit (NEB), and biotinylated DNA fragments were captured with 10 μl of streptavidin beads. The bead-bound DNA was PCR-amplified in a 50-μl reaction for 14 cycles. PCR products were purified using Ampure beads. Sequencing was performed on an Illumina NextSeq2000 platform (Illumina, San Diego, CA, USA) at the Indiana University Center for Genomics and Bioinformatics (Bloomington, IN, USA). Paired-end sequencing reads were mapped to genome sequences for each strain. Reads in which both reads were uniquely aligned to the genome were processed and sorted by *Hind*III restriction fragments and binned into 10-kb intervals using the HiC-lib pipeline. Downstream analyses and data visualizations were conducted in R.

WGS and marker frequency analysis

Cells were grown in ATGN broth at 30°C to exponential phase (OD_{600} = 0.6; ~5×10^9 CFU/ml) (*30*). Five milliliters of culture was collected and pelleted by centrifugation at 5000*g* for 10 min. gDNA was extracted using QIAGEN DNeasy Blood and Tissue Kit (QIAGEN), followed by sonication with a Qsonica Q800R2 water bath sonicator at 20% continuous amplitude for 6 min. The sheared gDNA was then used to prepare sequencing libraries with the NEBNext UltraII Kit and sequenced on an Illumina NextSeq2000 platform. Sequencing reads were mapped to the *A. tumefaciens* C58 reference genome (NCBI GCA_000092025.1) using CLC Genomics Workbench (QIAGEN). Read counts were normalized to the total number of mapped reads. Plotting and analysis were performed using R.

Transposon insertion sequencing

Tn-seq was performed as previously described (*21*). Briefly, the Mariner transposon-based plasmid pTND2823 (a gift from T. Dalia and A. Dalia, Indiana University) was transformed into the diaminopimelic acid auxotroph *E. coli* strain MFDpir, which was used to conjugate with *Agrobacterium* strains (*73*). *A. tumefaciens* conjugants were plated onto 10 large (150 mm in diameter) LB agar plates supplemented with kanamycin (300 μg/ml), at a density of approximately 1×10^6 colonies per plate. Plates were incubated at 30°C for 3 days, after which bacterial lawns were scraped and pooled into a single-cell suspension. Cells equivalent to 5 OD_{600} units were collected for gDNA extraction using the QIAGEN DNeasy Blood and Tissue Kit, following the manufacturer's protocol.

Three micrograms of gDNA was digested with *Mme*I for 90 min and treated with calf intestinal alkaline phosphatase for 60 min at 37°C. DNA was then purified by phenol-chloroform extraction, precipitated using ethanol, and resuspended in 15 μl of double-distilled water (ddH_2O). The digested end was ligated to an annealed adapter using T4 DNA ligase (*74*) and incubated at 16°C for ~16 hours. Adapter-ligated DNA was PCR-amplified using primers complementary to both the adapter and the transposon inverted repeat sequence. PCR products were gel purified using the Monarch DNA Gel Extraction Kit (NEB) and sequenced using the Illumina NextSeq2000 platform at the Indiana University Center for Genomics and Bioinformatics. Sequencing results were visualized using Artemis (table S5; https://sanger.ac.uk/tool/artemis/).

RT-qPCR of *Agrobacterium* virulence (*vir*) genes

Agrobacterium culture preparation

Frozen *Agrobacterium* glycerol stock cultures were streaked onto YEP agar plates and incubated at 28°C. A single colony was picked and inoculated into 3.5 ml of liquid YEP medium in a 50-ml conical tube and then cultured at 28°C with shaking at 200 rpm until reaching early stationary phase. Strains were subsequently induced with AS as described above in the "Fluorescence-based quantification of *vir* gene expression" section. Following induction, the bacterial cultures were incubated at 25°C with shaking at 200 rpm for an additional 21 hours.

Total RNA extraction and cDNA synthesis

A 2.5-ml sample of the induced culture was treated with QIAGEN's RNA protection reagent, following the manufacturer's protocol. Total RNA was extracted using the QIAGEN RNeasy Mini kit, which included an on-column deoxyribonuclease (DNase) treatment to eliminate any gDNA contamination. RNA integrity was verified by agarose gel electrophoresis, and purity and concentration were determined by measuring the absorbance at 260 and 280 nm using a spectrophotometer. Samples with residual gDNA contamination were further treated with the Turbo DNA-free kit (Ambion, Austin, TX, USA), followed by cleanup using the Monarch RNA purification kit (NEB). DNase-treated RNA was either stored at −80°C or immediately used for first-strand cDNA synthesis.

For cDNA synthesis, 1 μg of high-quality DNase-treated RNA was used in a 20-μl reaction with random hexamers, using the ProtoScript First Strand cDNA Synthesis Kit (NEB), according to the manufacturer's instructions. Before performing real-time quantitative PCR (qPCR), the resulting cDNA was diluted 1:10 with nuclease-free water, and 2 μl of the diluted cDNA was used as the template for each RT-qPCR reaction.

Real-time qPCR assays

Real-time qPCR assays were performed using a STRATAGENE Mx3005P system (Agilent Technologies, Santa Clara, CA, USA) and the Luna Universal qPCR SYBR Green Master Mix (NEB), following the manufacturer's protocol. Each 20-μl reaction contained 2 μl of diluted cDNA as the template, 250 nM of each primer, and 1× SYBR Green qPCR Master Mix, with nuclease-free water used to adjust the final volume. RT-qPCR reactions were run in triplicate for each cDNA sample, and the entire experiment was repeated four times. The thermal cycling conditions were as follows: an initial incubation at 95°C for 60 s for Uracil-DNA glycosylase activation, followed by 40 cycles of denaturation at 95°C for 15 s, annealing at 54°C for 30 s, and extension at 60°C for 30 s. A melting curve analysis was performed at the end of each run to confirm the specificity of amplification and absence of primer-dimer artifacts. Data were analyzed using the Mx3005P qPCR software, and cycle threshold (CT) values

were extracted for each reaction. The average CT value of the triplicates was used for subsequent statistical analyses.

Target genes and verification of PCR amplification efficiency

The expression levels of several candidate genes, including *virA*, *virB1*, *virC1*, *virD2*, *virE2*, *virG*, *chvG*, and *chvI*, were assessed by RT-qPCR. Primers for these target genes (table S4) were designed using the IDT PrimerQuest Tool (Integrated DNA Technology Inc., Iowa City, IA, USA) or Primer3 (*75*), based on the full coding sequences of the genes. Each primer was approximately 20 nucleotides in length with a melting temperature of 60°C, and the resulting amplicons ranged from 95 and 105 bp. Before running the RT-qPCR, primer specificity was validated by conventional PCR followed by agarose gel electrophoresis and Sanger sequencing.

To determine PCR amplification efficiency, standard curves were constructed for each target gene using 10-fold serial dilutions of gDNA from the *Agrobacterium* C58 strain. Four-point standard curves were generated by plotting the average CT values against the logarithm of the initial DNA quantity. Amplification efficiency (E) for each primer pair was calculated using the formula: $E\,(\%) = [10^{(-1/\text{slope})} - 1] \times 100$, where the slope was derived from the linear regression of the standard curve. Primer efficiencies, ranging from 98 to 109% (table S6), were used in subsequent statistical analyses.

Data analysis

RT-qPCR data were normalized to the expression of the *rpoD*, which encodes the RNA polymerase sigma factor and is commonly used as an endogenous control in RT-qPCR experiments (*76*). Relative transcript abundance of the target gene was calculated using the $2^{-\Delta\Delta Ct}$ method (*77*), with normalization to the WT C58 strain. Statistical significance of expression differences was evaluated using Student's t test, with $P \leq 0.05$ considered statistically significant.

RNA-seq library preparation

Agrobacterium cultures and RNA extraction procedures were described in the "RT-qPCR of *Agrobacterium* virulence (*vir*) genes" section. Three biological replicates of each strain were prepared and harvested across independent experimental shifts to ensure robust reproducibility and minimize batch-specific variation. After RNA extraction and quantification, an aliquot from each batch was run on an agarose gel to assess RNA integrity. Samples were then stored at −80°C until they were shipped to Zymo Research for RNA-seq analysis. RNA-seq libraries were constructed from 500 ng of total RNA using the Zymo-Seq RiboFree Total RNA Library Prep Kit (catalog no. R3000), following the manufacturer's manual (v1.3.0). Briefly, RNA was reverse-transcribed into cDNA, followed by rRNA depletion. Next, a partial P7 adapter sequence was ligated to the 3′ end of the cDNA, followed by second-strand synthesis and partial P5 adapter ligation at the 5′ end of the double-stranded DNA. Last, the libraries were amplified to incorporate full-length adapters under the following PCR conditions: initial denaturation at 95°C for 10 min; 12 cycles of denaturation at 95°C for 30 s, annealing at 60°C for 30 s, and extension at 72°C for 60 s; followed by a final extension at 72°C for 7 min. Library quality and size distribution were verified using Agilent's D1000 ScreenTape Assay on the TapeStation system. Libraries were then sequenced on an Illumina NovaSeq (150-bp paired-end sequencing) to a depth of ≥30 million read pairs per sample. Raw sequencing data were deposited in the NCBI SRA database under BioProject ID PRJNA1231518.

RNA-seq data bioinformatic analysis

The Zymo Research RNA-seq pipeline was originally adapted from the nf-core/RNAseq pipeline v1.4.2 (https://github.com/nf-core/rnaseq) (*78*). Both pipelines were built using Nextflow (https://nextflow.io/) (*79*). Briefly, quality control of raw reads was conducted using FastQC v0.11.9 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc). Adapter sequences and low-quality reads were trimmed from raw reads using Trim Galore! v0.6.6 (https://www.bioinformatics.babraham.ac.uk/projects/trim_galore). The trimmed reads were then aligned to the reference genome (GCF_000092025.1, NCBI datasets) using STAR v2.6.1d (https://github.com/alexdobin/STAR) (*80*). BAM file filtering and indexing were performed using SAMtools v1.9 (https://github.com/samtools/samtools) (*81*). RNA-seq library quality control was implemented using RSeQC v4.0.0 (http://rseqc.sourceforge.net/) (*82*) and QualiMap v2.2.2-dev (http://qualimap.conesalab.org/) (*83*). Duplicate reads were marked using Picard tools v2.23.9 (http://broadinstitute.github.io/picard/) (*84*). Library complexity was estimated using Preseq v2.0.3 (https://github.com/smithlabcode/preseq) (*85*), and duplication rate quality control was performed using dupRadar v1.18.0 (https://bioconductor.org/packages/dupRadar/) (*86*). Reads overlapping with exons were assigned to genes using featureCounts v2.0.1 (subread.sourceforge.net/featureCounts.html) (*87*). The classification of rRNA genes/exons and their reads were based on annotations and RepeatMasker rRNA tracks from the UCSC genome browser when applicable. Raw counts were normalized using the "counts()" function in DESeq2 with "normalized = TRUE" and subsequently log-transformed using the equation: $\log_2$(normalized counts + 4). Differential gene expression analysis was completed using DESeq2 v1.30.1 (https://bioconductor.org/packages/DESeq2/) (*88*). DESeq2 and PCA were performed on integrated Differential Expression and Pathway (iDEP) analysis (https://bioinformatics.sdstate.edu/idep/) (*89*). DEGs were defined as those with false discovery rate–adjusted $P < 0.05$ and fold change > 2. Last, quality control metrics and analysis results plots were visualized using MultiQC v1.9 (https://github.com/ewels/MultiQC) (*90*).

A summary of quality control, mapping, and counting statistics is provided in table S2. KEGG number assignment was performed using KofamKOALA v2024-12-02 (https://www.genome.jp/tools/kofamkoala/) (*91*) with the C58 reference proteome (GCF_000092025.1, NCBI datasets) as input. KEGG pathways were reconstructed with a KEGG mapper (https://www.genome.jp/kegg/mapper/reconstruct.html) (*47*). Plotting of normalized reads and one-way ANOVA followed by Dunnett's multiple comparisons tests were carried out in GraphPad Prism v10.4.1 (GraphPad Software, Boston, MA, USA).

Weighted gene coexpression network analysis

We performed WGCNA to identify gene modules associated with AS treatment and strain-specific expression patterns (*46*). The analysis was conducted using the WGCNA R package v1.73. Genes with expression levels above 0.5 counts per million reads were included in the analysis. Modules were constructed using a signed hybrid network with a soft-threshold power (β) of 10, which yielded a scale-free topology model fit of 0.772 and a mean connectivity of 103. The minimum module size was set to 50 genes. Module-trait relationships were assessed using the Biweight Midcorrelation method (*92*). Within each module, genes with high module membership (MM > 0.90, $P < 0.05$) were ranked by the greatest difference in gene expression between the AS and DMSO treatments or between WT and FL strains.

Tumorigenesis assay on *Kalanchoe* leaves

A *Kalanchoe* leaf–based tumorigenesis assay was performed as previously described (*29*). Briefly, *Agrobacterium* strains were cultured overnight (12 to 15 hours) in 3.5 ml of liquid YEP medium in 50-ml conical tubes, incubated at 28°C with shaking at 200 rpm, until they reached the late exponential phase. The bacterial cultures were then normalized to an OD_{600} of 0.5. A 5-μl aliquot of the bacterial suspension was applied to wounded sites on 6- to 8-week-old *Kalanchoe* leaves, prepared using hypodermic needles. Tumor formation was monitored over a period of 2 to 3 weeks, and photographs were taken to document the results.

Evaluation of virulence via RUBY expression assay

Plant materials and growth conditions

N. benthamiana plants were grown in Metro-Mix 360 (Sungro, MA, USA) supplemented with Plant Food (Expert Gardener) in plastic trays. Plants were maintained in a growth chamber under controlled conditions: 150 μmol m^{-2} s^{-1} light intensity, a 16-hour light/8-hour dark photoperiod, and a constant temperature of 24° ± 1°C.

Agrobacterium-mediated transient expression assay (agroinfiltration)

Agrobacterium strains were transformed with the *RUBY* reporter construct pCBL101-RUBY (*93*). The *RUBY* reporter converts tyrosine into purple pigment betalain (*41*), and its expression serves as a quantitative indicator of transient T-DNA delivery efficiency (*42*). All tested strains were grown on YEP agar plates containing spectinomycin (100 mg/liter) and thymidine (50 mg/liter). After 2 days, bacterial colonies were inoculated into YEP liquid medium [supplemented with spectinomycin (100 mg/liter) and thymidine (50 mg/liter)] and cultured overnight at 28°C in the dark on a shaking incubator. Bacterial cells were harvested by centrifugation at 4000*g* for 10 min and resuspended in agroinfiltration buffer [10 mM 2-morpholinoethanesulfonic acid (MES, pH 5.6), 10 mM $MgCl_2$, and 200 μM AS]. The cell suspension was adjusted to an OD_{600} of 0.5, incubated at room temperature in the dark for 2 to 3 hours, and then infiltrated into fully expanded leaves of 5-week-old *N. benthamiana* plants. Agroinfiltration buffer alone was used as a mock control for the experiment.

Betalain content assay

Two days after agroinfiltration, purple pigmentation was observed in the leaf areas infiltrated with the tested *Agrobacterium* strains but not in mock control. Betalain pigments were extracted and quantified using the method described previously (*42*). Briefly, the infiltrated areas were excised from each *N. benthamiana* leaf using clean scissors, and pigments were extracted in 50% methanol using a hand grinder and plastic pestle. The homogenized samples were vortexed for 5 to 10 s and centrifuged at 13,000*g* for 5 min. The resulting supernatant was diluted with distilled H_2O, and absorbance at 538 nm was measured using a SpectraMax iD3 Multi-Mode Microplate Reader (Molecular Devices LLC, San Jose, CA, USA) in 96-well plates. Betalain content (mg/kg fresh tissue) was calculated using the formula: Betalain content = $[(A \times DF \times MW \times 1000)/(\varepsilon \times L)]$, where A is the absorbance value at 538 nm, DF is the dilution factor (10× sample dilution ratio), MW is the molecular weight of betalain (550 g/mol), ε is the molar extinction coefficient (60,000 liters/mol·cm in H_2O), and L is the path length (0.2643 cm for a 96-well plate)

UV light sensitivity assay of *recA* mutants

The effects of UV-B light on *Agrobacterium* strains were tested as described by Aliu *et al.* (*29*), with minor modifications. Briefly, overnight cultures of each strain were prepared by inoculating a single colony into a 50-ml conical tube containing 3 ml of YEP medium and incubating at 28°C with shaking at 200 rpm for 14 to 18 hours. The cultures were then diluted in fresh YEP medium to an OD_{600} of 0.6 to 0.7. Serial dilutions (10 to 10^6 fold) of the bacterial suspension were prepared in a six-well plate and exposed to UV-B irradiation for 2 hours in the dark using a 6-W handheld UV lamp (365 nm, model UVM; Analytik Jena US LLC, Upland, CA, USA) positioned 10 cm above the plate. After exposure, 4 μl of each bacterial suspension was pipetted onto YEP agar plates, which were then covered with aluminum foil and incubated at 28°C for 34 hours. As a control, duplicate dilutions of bacterial suspensions that were not exposed to UV irradiation were similarly plated and incubated. The *recA*-deficient *Agrobacterium* AGL1 strain (*94*) was included as a DNA repair–deficient control and sensitivity reference for the experiment.

Peroxide stress assay of *Agrobacterium* strains

The H_2O_2 stress tolerance test for all *Agrobacterium* strains was conducted using a disk diffusion method (*95*) on YEP agar plates. In this procedure, 7-mm filter paper disks were soaked in 30% H_2O_2 for 2 min and air-dried for 6 min before being placed on YEP agar plates that had been uniformly streaked with *Agrobacterium* strains. For each strain, three replicate plates were tested. The plates were sealed with parafilm and incubated at 28°C in the dark for 48 hours. After incubation, the diameter of the inhibition zone surrounding each disk was measured, and the average was calculated to evaluate strain sensitivity to oxidative stress.

Stable transformation in maize

Maize (*Zea mays* L.) inbred line B104 seeds were provided by E. Vollbrecht (Iowa State University). B104 plants were grown in the Agronomy Greenhouse at Iowa State University under conditions described by Kang *et al.* (*96*). Immature embryos were harvested 12 to 14 days after pollination, when they reached a size of 1.8 to 2.0 mm, for *Agrobacterium* infection. The T-DNA vector construct pCBL101-RUBY and the helper plasmid pKL2299A (*44*, *93*) were introduced to both WT EHA105 and the EHA105Cir strains. *Agrobacterium* infection and transformation procedures were performed as described by Lee *et al.* (*93*) and Kang *et al.* (*97*).

Briefly, *Agrobacterium* strains were initially cultured on mother plates containing spectinomycin (100 mg/liter), gentamycin (50 mg/liter), and thymidine (50 mg/liter). One day before infection, cells were streaked onto fresh YEP plates with the same antibiotics and thymidine and incubated at 28°C for 16 to 20 hours. The infection solution was prepared by harvesting *Agrobacterium* cells from YEP plates and suspending them in infection medium to an OD_{550} of 0.4 to 0.6, immediately before the infection. Immature embryos were dissected, collected in 2-ml microcentrifuge tubes, and incubated in 1 ml of the infection medium for 5 min at room temperature (22°C). Following infection, embryos were placed scutellum side up on cocultivation medium and incubated at 20°C in the dark for 16 to 20 hours. Embryos were then transferred to resting medium and incubated at 28°C in the dark for 1 week, followed by two successive 2-week incubations on maturation medium amended with G418 (75 mg/liter), also at 28°C in the dark. Regenerated shoots were transferred to rooting medium with G418 (75 mg/liter) and incubated in a light chamber with a 16-hour light/8-hour dark photoperiod at 28°C for 2 weeks. At this stage, rooted plantlets, with or without *RUBY* expression, were counted and sampled for further analysis. Transformation

frequency was calculated as the percentage of rooted plantlets relative to the total number of infected embryos.

Statistical analysis

Differences between sample means were analyzed using a two-tailed Student's *t* test, with $P \leq 0.05$ considered statistically significant. For comparisons involving more than two groups, a one-way ANOVA followed by Dunnett's or Tukey's multiple comparisons test was performed. See the "Growth and cell viability assay" and "RNA-seq data bioinformatic analysis" sections for details on ANOVA applications.

Use of artificial intelligence

Large language model ChatGPT (GPT-4o), developed by OpenAI, was used for minor refinements in sentence structure and grammar throughout the text. When using it, we entered the sentences written by ourselves and asked GPT for improvement. For example, our original sentence: "As reported by Liao *et al.* (*21*), the replication origins of both the circular and linear chromosomes, and their partitioning loci, *parAB* and *repABC*, were essential in all four strains, indicating that the four strains remained the same replication and segregation program." GPT improvement: "As reported by Liao *et al.* (*21*), the replication origins of both the circular and linear chromosomes, along with their partitioning loci, *parAB* and *repABC*, were essential in all four strains, indicating that they maintained the same replication and segregation program."

Supplementary Materials

This PDF file includes:

Figs. S1 to S18
Tables S1 to S6

REFERENCES AND NOTES

1. J. Cairns, The bacterial chromosome and its manner of replication as seen by autoradiography. *J. Mol. Biol.* **6**, 208–213 (1963).
2. C. Baril, C. Richaud, G. Baranton, I. S. Saint Girons, Linear chromosome of *Borrelia burgdorferi*. *Res. Microbiol.* **140**, 507–516 (1989).
3. J. Hinnebusch, K. Tilly, Linear plasmids and chromosomes in bacteria. *Mol. Microbiol.* **10**, 917–922 (1993).
4. J. N. Volff, J. Altenbuchner, A new beginning with new ends: Linearisation of circular chromosomes during bacterial evolution. *FEMS Microbiol. Lett.* **186**, 143–150 (2000).
5. G. C. diCenzo, T. M. Finan, The divided bacterial genome: Structure, function, and evolution. *Microbiol. Mol. Biol. Rev.* **81**, e00019-17 (2017).
6. A. Shi, F. Fan, J. R. Broach, Microbial adaptive evolution. *J. Ind. Microbiol. Biotechnol.* **49**, kuab076 (2022).
7. F. Choulet, B. Aigle, A. Gallois, S. Mangenot, C. Gerbaud, C. Truong, F. X. Francou, C. Fourrier, M. Guérineau, B. Decaris, V. Barbe, Evolution of the terminal regions of the *Streptomyces* linear chromosome. *Mol. Biol. Evol.* **23**, 2361–2369 (2006).
8. T. J. Krogh, J. Møller-Jensen, C. Kaleta, Impact of chromosomal architecture on the function and evolution of bacterial genomes. *Front. Microbiol.* **9**, 2019 (2018).
9. E. L. Torrance, A. Diop, L. M. Bobay, Homologous recombination shapes the architecture and evolution of bacterial genomes. *Nucleic Acids Res.* **52**, gkae1265 (2024).
10. M. Y. Galperin, Linear chromosomes in bacteria: No straight edge advantage? *Environ. Microbiol.* **9**, 1357–1362 (2007).
11. P. R. Marri, L. K. Harris, K. Houmiel, S. C. Slater, H. Ochman, The effect of chromosome geometry on genetic diversity. *Genetics* **179**, 511–516 (2008).
12. S. C. Slater, B. S. Goldman, B. Goodner, J. C. Setubal, S. K. Farrand, E. W. Nester, T. J. Burr, L. Banta, A. W. Dickerman, I. Paulsen, L. Otten, G. Suen, R. Welch, N. F. Almeida, F. Arnold, O. T. Burton, Z. Du, A. Ewing, E. Godsy, S. Heisel, K. L. Houmiel, J. Jhaveri, J. Lu, N. M. Miller, S. Norton, Q. Chen, W. Phoolcharoen, V. Ohlin, D. Ondrusek, N. Pride, S. L. Stricklin, J. Sun, C. Wheeler, L. Wilson, H. Zhu, D. W. Wood, Genome sequences of three *Agrobacterium* biovars help elucidate the evolution of multichromosome genomes in bacteria. *J. Bacteriol.* **191**, 2501–2511 (2009).
13. S. B. Gelvin, *Agrobacterium*-mediated plant transformation: The biology behind the "gene-jockeying" tool. *Microbiol. Mol. Biol. Rev.* **67**, 16–37 (2003).
14. P. W. Harrison, R. P. Lower, N. K. Kim, J. P. W. Young, Introducing the bacterial 'chromid': Not a chromosome, not a plasmid. *Trends Microbiol.* **18**, 141–148 (2010).
15. A. J. Weisberg, Y. Wu, J. H. Chang, E. M. Lai, C.-H. Kuo, Virulence and ecology of agrobacteria in the context of evolutionary genomics. *Annu. Rev. Phytopathol.* **61**, 1–23 (2023).
16. W. M. Huang, J. DaGloria, H. Fox, Q. Ruan, J. Tillou, K. Shi, H. Aihara, J. Aron, S. Casjens, Linear chromosome-generating system of *Agrobacterium tumefaciens* C58: Protelomerase generates and protects hairpin ends. *J. Biol. Chem.* **287**, 25551–25563 (2012).
17. M. H. Ramírez-Bahena, L. Vial, F. Lassalle, B. Diel, D. Chapulliot, V. Daubin, X. Nesme, D. Muller, Single acquisition of protelomerase gave rise to speciation of a large and diverse clade within the *Agrobacterium/Rhizobium* supercluster characterized by the presence of a linear chromid. *Mol. Phylogenet. Evol.* **73**, 202–207 (2014).
18. B. Goodner, G. Hinkle, S. Gattung, N. Miller, M. Blanchard, B. Qurollo, B. S. Goldman, Y. Cao, M. Askenazi, C. Halling, L. Mullin, K. Houmiel, J. Gordon, M. Vaudin, O. Iartchouk, A. Epp, F. Liu, C. Wollam, M. Allinger, D. Doughty, C. Scott, C. Lappas, B. Markelz, C. Flanagan, C. Crowell, J. Gurson, C. Lomo, C. Sear, G. Strub, C. Cielo, S. Slater, Genome sequence of the plant pathogen and biotechnology agent *Agrobacterium tumefaciens* C58. *Science* **294**, 2323–2328 (2001).
19. D. W. Wood, J. C. Setubal, R. Kaul, D. E. Monks, J. P. Kitajima, V. K. Okura, Y. Zhou, L. Chen, G. E. Wood, N. F. Almeida Jr., L. Woo, Y. Chen, I. T. Paulsen, J. A. Eisen, P. D. Karp, D. Bovee Sr., P. Chapman, J. Clendenning, G. Deatherage, W. Gillet, C. Grant, T. Kutyavin, M. J. Li, E. McClelland, A. Palmieri, C. Raymond, G. Rouse, C. Saenphimmachak, Z. Wu, P. Romero, D. Gordon, S. Zhang, H. Yoo, Y. Tao, P. Biddle, M. Jung, W. Krespan, M. Perry, B. Gordon-Kamm, L. Liao, S. Kim, C. Hendrick, Z. Y. Zhao, M. Dolan, F. Chumley, S. V. Tingey, J. F. Tomb, M. P. Gordon, M. V. Olson, E. W. Nester, The genome of the natural genetic engineer *Agrobacterium tumefaciens* C58. *Science* **294**, 2317–2323 (2001).
20. J. De Saeger, J. Park, H. S. Chung, J.-P. Hernalsteens, M. Van Lijsebettens, D. Inzé, M. Van Montagu, S. Depuydt, Agrobacterium strains and strain improvement: Present and outlook. *Biotechnol. Adv.* **53**, 107677 (2021).
21. Q. Liao, Z. Ren, E. E. Wiesler, C. Fuqua, X. Wang, A dicentric bacterial chromosome requires XerC/D site-specific recombinases for resolution. *Curr. Biol.* **32**, 3609–3618.e7 (2022).
22. R. Sanath-Kumar, A. Rahman, Z. Ren, I. P. Reynolds, L. Augusta, C. Fuqua, A. J. Weisberg, X. Wang, Linear dicentric chromosomes in bacterial natural isolates reveal common constraints for replicon fusion. *mBio* **16**, e01046-25 (2025).
23. X. Guo, M. Flores, P. Mavingui, S. I. Fuentes, G. Hernández, G. Dávila, R. Palacios, Natural genomic design in *Sinorhizobium meliloti*: Novel genomic architectures. *Genome Res.* **13**, 1810–1817 (2003).
24. C. Chapman, M. Henry, K. A. Bishop-Lilly, J. Awosika, A. Briska, R. N. Ptashkin, T. Wagner, C. Rajanna, H. Tsang, S. L. Johnson, V. P. Mokashi, P. S. G. Chain, S. Sozhamannan, Scanning the landscape of genome architecture of non-O1 and non-O139 *Vibrio cholerae* by whole genome mapping reveals extensive population genetic diversity. *PLOS ONE* **10**, e0120311 (2015).
25. J. F. Mori, R. A. Kanaly, Natural chromosome-chromid fusion across rRNA operons in a *Burkholderiaceae* bacterium. *Microbiol. Spectr.* **10**, e02225-21 (2022).
26. S. J. Klein, R. J. O'Neill, Transposable elements: Genome innovation, chromosome diversity, and centromere conflict. *Chromosome Res.* **26**, 5–23 (2018).
27. S. E. Klompe, P. L. H. Vo, T. S. Halpin-Healy, S. H. Sternberg, Transposon-encoded CRISPR–Cas systems direct RNA-guided DNA integration. *Nature* **571**, 219–225 (2019).
28. P. L. H. Vo, C. Ronda, S. E. Klompe, E. E. Chen, C. Acree, H. H. Wang, S. H. Sternberg, CRISPR RNA-guided integrases for high-efficiency, multiplexed bacterial genome engineering. *Nat. Biotechnol.* **39**, 480–489 (2021).
29. E. Aliu, K. Lee, K. Wang, CRISPR RNA-guided integrase enables high-efficiency targeted genome engineering in *Agrobacterium tumefaciens*. *Plant Biotechnol. J.* **20**, 1916–1927 (2022).
30. E. R. Morton, P. M. Merritt, J. D. Bever, C. Fuqua, Large deletions in the pAtC58 megaplasmid of *Agrobacterium tumefaciens* can confer reduced carriage cost and increased expression of virulence genes. *Genome Biol. Evol.* **5**, 1353–1364 (2013).
31. S. B. Gelvin, *Agrobacterium* virulence gene induction, in *Agrobacterium Protocols*, K. Wang, Ed. (Humana Press, ed. 2, 2006), pp. 77–84.
32. B. Gompertz, XXIV. On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. In a letter to Francis Baily, Esq. F. R. S. *Philos. Trans. R. Soc. Lond.* **115**, 513–583 (1825).
33. C. P. Winsor, The Gompertz curve as a growth curve. *Proc. Natl. Acad. Sci. U.S.A.* **18**, 1–8 (1932).
34. M. H. Zwietering, I. Jongenburger, F. M. Rombouts, K. van 't Riet, Modeling of the bacterial growth curve. *Appl. Environ. Microbiol.* **56**, 1875–1881 (1990).
35. S. E. Stachel, E. Messens, M. Van Montagu, P. Zambryski, Identification of the signal molecules produced by wounded plant cells that activate T-DNA transfer in *Agrobacterium tumefaciens*. *Nature* **318**, 624–629 (1985).
36. S. E. Stachel, E. W. Nester, P. Zambryski, A plant cell factor induces *Agrobacterium tumefaciens* vir gene expression. *Proc. Natl. Acad. Sci. U.S.A.* **83**, 379–383 (1986a).

37. S. E. Stachel, P. Zambryski, *virA* and *virG* control the plant-induced activation of the T-DNA transfer process of *A. tumefaciens*. *Cell* **46**, 325–333 (1986).
38. L. S. Melchers, T. J. Regensburg-Tuïnk, R. B. Bourret, N. J. Sedee, R. A. Schilperoort, P. J. Hooykaas, Membrane topology and functional analysis of the sensory protein VirA of *Agrobacterium tumefaciens*. *EMBO J.* **8**, 1919–1925 (1989).
39. A. M. Ashby, M. D. Watson, C. H. Shaw, A Ti-plasmid determined function is responsible for chemotaxis of *Agrobacterium tumefaciens* towards the plant wound product acteosyringone. *FEMS Microbiol. Lett.* **41**, 189–192 (1987).
40. A. Das, S. Stachel, P. Ebert, P. Allenza, A. Montoya, E. Nester, Promoters of *Agrobacterium tumefaciens* Ti-plasmid virulence genes. *Nucleic Acids Res.* **14**, 1355–1364 (1986).
41. Y. He, T. Zhang, H. Sun, H. Zhan, Y. Zhao, A reporter for noninvasively monitoring gene expression and plant transformation. *Hortic. Res.* **7**, 152 (2020).
42. D. Pramanik, K. Lee, K. Wang, A simple and efficient method for betalain quantification in *RUBY*-expressing plant samples. *Front. Plant Sci.* **15**, 1449409 (2024).
43. E. E. Hood, S. B. Gelvin, L. S. Melchers, A. Hoekema, New *Agrobacterium* helper plasmids for gene transfer to plants. *Transgenic Res.* **2**, 208–218 (1993).
44. E. Aliu, Q. Ji, A. Wlazlo, S. Grosic, M. K. Azanu, K. Wang, K. Lee, Enhancing *Agrobacterium*-mediated plant transformation efficiency through improved ternary vector systems and auxotrophic strains. *Front. Plant Sci.* **15**, 1429353 (2024).
45. T. C. Charles, E. W. Nester, A chromosomally encoded two-component sensory transduction system is required for virulence of *Agrobacterium tumefaciens*. *J. Bacteriol.* **175**, 6614–6625 (1993).
46. P. Langfelder, S. Horvath, WGCNA: An R package for weighted correlation network analysis. *BMC Bioinformatics* **9**, 559 (2008).
47. M. Kanehisa, S. Goto, KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* **28**, 27–30 (2000).
48. R. Stasi, H. I. Neves, B. Spira, Phosphate uptake by the phosphonate transport system PhnCDE. *BMC Microbiol.* **19**, 79 (2019).
49. A. Schrell, J. Alt-Moerbe, T. Lanz, J. Schroeder, Arginase of *Agrobacterium* Ti plasmid C58: DNA sequence, properties, and comparison with eucaryotic enzymes. *Eur. J. Biochem.* **184**, 635–641 (1989).
50. H. Zanker, J. von Lintig, J. Schröder, Opine transport genes in the octopine (*occ*) and nopaline (*noc*) catabolic regions in Ti plasmids of *Agrobacterium tumefaciens*. *J. Bacteriol.* **174**, 841–849 (1992).
51. K. R. Piper, S. B. von Bodman, S. K. Farrand, Conjugation factor of *Agrobacterium tumefaciens* regulates Ti plasmid transfer by autoinduction. *Nature* **362**, 448–450 (1993).
52. B. C. Heckel, A. D. Tomlinson, E. R. Morton, J.-H. Choi, C. Fuqua, *Agrobacterium tumefaciens* exoR controls acid response genes and impacts exopolysaccharide synthesis, horizontal gene transfer, and virulence gene expression. *J. Bacteriol.* **196**, 3221–3233 (2014).
53. J. Hołówka, D. Trojanowski, K. Ginda, B. Wojtaś, B. Gielniewski, D. Jakimowicz, J. Zakrzewska-Czerwińska, HupB is a bacterial nucleoid-associated protein with an indispensable eukaryotic-like tail. *mBio* **8**, e01272–17 (2017).
54. M. Aktas, K. A. Jost, C. Fritz, F. Narberhaus, Choline uptake in *Agrobacterium tumefaciens* by the high-affinity ChoXWV transporter. *J. Bacteriol.* **193**, 5119–5129 (2011).
55. J. A. Budnick, L. M. Sheehan, M. J. Ginder, K. C. Failor, J. M. Perkowski, J. F. Pinto, K. A. Kohl, L. Kang, P. Michalak, L. Luo, J. E. Heindl, C. C. Caswell, A central role for the transcriptional regulator VtlR in small RNA-mediated gene regulation in *Agrobacterium tumefaciens*. *Sci. Rep.* **10**, 14968 (2020).
56. A. Jordan, F. Åslund, E. Pontis, P. Reichard, A. Holmgren, Characterization of *Escherichia coli* NrdH: A glutaredoxin-like protein with a thioredoxin-like activity profile. *J. Biol. Chem.* **272**, 18044–18050 (1997).
57. M. Torres, A. Jiquel, E. Jeanne, D. Naquin, Y. Dessaux, D. Faure, *Agrobacterium tumefaciens* fitness genes involved in the colonization of plant tumors and roots. *New Phytol.* **233**, 905–918 (2022).
58. A. González-Mula, J. Lang, C. Grandclément, D. Naquin, M. Ahmar, L. Soulère, Y. Queneau, Y. Dessaux, D. Faure, Lifestyle of the biotroph *Agrobacterium tumefaciens* in the ecological niche constructed on its host plant. *New Phytol.* **219**, 350–362 (2018).
59. A. Staroń, T. Mascher, General stress response in α-proteobacteria: PhyR and beyond. *Mol. Microbiol.* **78**, 271–277 (2010).
60. A. Fiebig, J. Herrou, J. Willett, S. Crosson, General stress signaling in the Alphaproteobacteria. *Annu. Rev. Genet.* **49**, 603–625 (2015).
61. H. Zhu, S. Shuman, Characterization of *Agrobacterium tumefaciens* DNA ligases C and D. *Nucleic Acids Res.* **35**, 3631–3645 (2007).
62. B. Prapagdee, W. Eiamphungporn, P. Saenkham, S. Mongkolsuk, P. Vattanaviboon, Analysis of growth phase regulated KatA and CatE and their physiological roles in determining hydrogen peroxide resistance in *Agrobacterium tumefaciens*. *FEMS Microbiol. Lett.* **237**, 219–226 (2004).
63. H. Wang, M. Zhang, E. Wang, R. Xiao, S. Zhang, M. Guo, *Agrobacterium fabrum* gene atu1420 regulates the pathogenicity by affecting the degradation of growth-and virulence-associated phenols. *Res. Microbiol.* **174**, 104011 (2023).
64. L. S. Ma, A. Hachani, J. S. Lin, A. Filloux, E. M. Lai, *Agrobacterium tumefaciens* deploys a superfamily of type VI secretion DNase effectors as weapons for interbacterial competition *in planta*. *Cell Host Microbe* **16**, 94–104 (2014).
65. D. Dar, R. Sorek, Extensive reshaping of bacterial operons by programmed mRNA decay. *PLOS Genet.* **14**, e1007354 (2018).
66. M. de Virgilio, A. Lombardi, R. Caliandro, M. S. Fabbrini, Ribosome-inactivating proteins: From plant defense to tumor attack. *Toxins* **2**, 2699–2737 (2010).
67. A. A. Slavokhotova, E. A. Rogozhin, Defense peptides from the α-hairpinin family are components of plant innate immunity. *Front. Plant Sci.* **11**, 465 (2020).
68. S. Tabata, P. J. Hooykaas, A. Oka, Sequence determination and characterization of the replicator region in the tumor-inducing plasmid pTiB6S3. *J. Bacteriol.* **171**, 1665–1672 (1989).
69. M. A. Ramirez-Romero, N. Soberon, A. Perez-Oseguera, J. Tellez-Sosa, M. A. Cevallos, Structural elements required for replication and incompatibility of the *Rhizobium etli* symbiotic plasmid. *J. Bacteriol.* **182**, 3117–3124 (2000).
70. S. Yamamoto, K. Kiyokawa, K. Tanaka, K. Moriguchi, K. Suzuki, Novel toxin-antitoxin system composed of serine protease and AAA-ATPase homologues determines the high level of stability and incompatibility of the tumor-inducing plasmid pTiC58. *J. Bacteriol.* **191**, 4656–4666 (2009).
71. D. G. Gibson, L. Young, R. Y. Chuang, J. C. Venter, C. A. Hutchison III, H. O. Smith, Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6**, 343–345 (2009).
72. Z. Ren, Q. Liao, X. Karaboja, I. S. Barton, E. G. Schantz, A. Mejia-Santana, C. Fuqua, X. Wang, Confirmation and dynamic interactions of the multipartite genome in *Agrobacterium tumefaciens*. *Proc. Natl. Acad. Sci. U.S.A.* **119**, e2115854119 (2022).
73. L. Ferrières, G. Hémery, T. Nham, A. M. Guérout, D. Mazel, C. Beloin, J. M. Ghigo, Silent mischief: Bacteriophage Mu insertions contaminate products of *Escherichia coli* random mutagenesis performed using suicidal transposon delivery plasmids mobilized by broad-host-range RP4 conjugative machinery. *J. Bacteriol.* **192**, 6418–6427 (2010).
74. T. van Opijnen, D. W. Lazinski, A. Camilli, Genome-wide fitness and genetic interactions determined by Tn-seq, a high-throughput massively parallel sequencing method for microorganisms. *Curr. Protoc. Microbiol.* **36**, 1E.3.1–1E.3.24 (2015).
75. A. Untergasser, I. Cutcutache, T. Koressaar, J. Ye, B. C. Faircloth, M. Remm, S. G. Rozen, Primer3-new capabilities and interfaces. *Nucleic Acids Res.* **40**, e115 (2012).
76. K. Lee, X. Huang, C. Yang, D. Lee, V. Ho, K. Nobuta, J. B. Fan, K. Wang, A genome-wide survey of highly expressed non-coding RNAs and biological validation of selected candidates in *Agrobacterium tumefaciens*. *PLOS ONE* **8**, e70720 (2013).
77. K. J. Livak, T. D. Schmittgen, Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. *Methods* **25**, 402–408 (2001).
78. H. Patel, P. Ewels, A. Peltzer, R. Hammarén, O. Botvinnik, G. Sturm, C. Davenport, nf-core/rnaseq: nf-core/rnaseq v3.0 - Silver Shark (Version 3.0). Zenodo (2020).
79. P. Di Tommaso, M. Chatzou, E. W. Floden, P. P. Barja, E. Palumbo, C. Notredame, Nextflow enables reproducible computational workflows. *Nat. Biotechnol.* **35**, 316–319 (2017).
80. A. Dobin, C. A. Davis, F. Schlesinger, J. Drenkow, C. Zaleski, S. Jha, P. Batut, M. Chaisson, T. R. Gingeras, STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
81. P. Danecek, J. K. Bonfield, J. Liddle, J. Marshall, V. Ohan, M. O. Pollard, A. Whitwham, T. Keane, S. A. McCarthy, R. M. Davies, H. Li, Twelve years of SAMtools and BCFtools. *Gigascience* **10**, giab008 (2021).
82. L. Wang, S. Wang, W. Li, RSeQC: Quality control of RNA-seq experiments. *Bioinformatics* **28**, 2184–2185 (2012).
83. F. García-Alcalde, K. Okonechnikov, J. Carbonell, L. M. Cruz, S. Götz, S. Tarazona, J. Dopazo, T. F. Meyer, A. Conesa, Qualimap: Evaluating next-generation sequencing alignment data. *Bioinformatics* **28**, 2678–2679 (2012).
84. "Picard Toolkit." Broad Institute, GitHub Repository (Broad Institute, 2019). http://broadinstitute.github.io/picard/.
85. T. Daley, A. D. Smith, Predicting the molecular complexity of sequencing libraries. *Nat. Methods* **10**, 325–327 (2013).
86. S. Sayols, D. Scherzinger, H. Klein, dupRadar: A Bioconductor package for the assessment of PCR artifacts in RNA-seq data. *BMC Bioinformatics* **17**, 428 (2016).
87. Y. Liao, G. K. Smyth, W. Shi, FeatureCounts: An efficient general-purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).
88. M. I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
89. S. X. Ge, E. W. Son, R. Yao, iDEP: An integrated web application for differential expression and pathway analysis of RNA-seq data. *BMC Bioinformatics* **19**, 534 (2018).
90. P. Ewels, M. Magnusson, S. Lundin, M. Käller, MultiQC: Summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* **32**, 3047–3048 (2016).
91. T. Aramaki, R. Blanc-Mathieu, H. Endo, K. Ohkubo, M. Kanehisa, S. Goto, H. Ogata, KofamKOALA: KEGG ortholog assignment based on profile HMM and adaptive score threshold. *Bioinformatics* **36**, 2251–2252 (2020).
92. C. H. Zheng, L. Yuan, W. Sha, Z. L. Sun, Gene differential coexpression analysis based on biweight correlation and maximum clique. *BMC Bioinformatics* **15**, S3 (2014).

93. K. Lee, M. Kang, Q. Ji, S. Grosic, K. Wang, New T-DNA binary vectors with *NptII* selection and *RUBY* reporter for efficient maize transformation and targeted mutagenesis. *Plant Physiol.* **192**, 2598–2603 (2023).
94. G. R. Lazo, P. A. Stein, R. A. Ludwig, A DNA transformation-competent Arabidopsis genomic library in *Agrobacterium*. *Biotechnology* **9**, 963–967 (1991).
95. M. L. Hossain, L. Y. Lim, K. Hammer, D. Hettiarachchi, C. Locher, A review of commonly used methodologies for assessing the antibacterial activity of honey and honey products. *Antibiotics* **11**, 975 (2022).
96. M. Kang, K. Lee, T. Finley, H. Chappell, V. Veena, K. Wang, An improved *Agrobacterium*-mediated transformation and genome-editing method for maize inbred B104 using a ternary vector system and immature embryos. *Front. Plant Sci.* **13**, 860971 (2022).
97. M. Kang, M. K. Azanu, K. Lee, K. Wang, A rapid *Agrobacterium*-mediated transformation method using maize B104 immature embryos. *Cold Spring Harb. Protoc.*, doi: 10.1101/pdb.prot108595 (2024).

Acknowledgments: We thank A. Weisberg for providing the WT *A. tumefaciens* C58 strain, S. Grosic for providing maize materials for transformation, A. Dalia for providing the transposon plasmid (pTND2823), and C. Fuqua for discussion. We acknowledge the use of ChatGPT, a language model developed by OpenAI, for minor refinements in sentence structure and grammar throughout the text. We also acknowledge publication subvention support from the Office of the Vice President for Research and the Crop Bioengineering Center at Iowa State University. **Funding:** This work was partially supported by the National Science Foundation (NSF) Plant Genome Research Program award IOS-1917138 and NSF Established Program to Stimulate Competitive Research's Research Infrastructure Improvement Program award 2121410 to K.W. and K.L., by the seed grant fund from Crop Bioengineering Center of Iowa State University to K.L., by the National Institute of Food and Agriculture of United State Department of Agriculture Hatch project #IOW04714 and State of Iowa funds to K.W., and by National Institutes of Health awards R01GM141242, R01GM143182, and R01AI172822 and NSF GEMS Biology Integration Institute 2022049 to X.W. K.W.'s contribution to this work is partially supported by (while serving at) the NSF. **Author contributions:** Conceptualization: K.W., K.L., and E.A. Funding acquisition: K.W., K.L., and X.W. Supervision: K.W., K.L., and X.W. Methodology: E.A., K.L., K.W., and X.W. Investigation: E.A., K.L., R.S.-K., C.Y., D.P., and Q.J. Formal analysis: E.A., K.L., R.S.-K., C.Y., L.G., X.W., and K.W. Data curation: C.Y. Visualization: E.A., C.Y., K.W., K.L., X.W., R.S.-K., and D.P. Writing—original draft: E.A., K.L., K.W., C.Y., X.W., and R.S.-K. Writing—review and editing: E.A., K.L., R.S.-K., C.Y., L.G., D.P., Q.J., X.W., and K.W. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials or Zenodo repository 15122195. The RNA-seq dataset is deposited in the NCBI Sequence Read Archive (SRA) under accession number PRJNA1231518. Hi-C and WGS data used for marker frequency analysis are available in the NCBI Gene Expression Omnibus (GEO) under accession number GSE293397. The bacterial strains and plasmid DNAs can be provided by K.W. (kanwang@iastate.edu) pending scientific review and a completed material transfer agreement.

Submitted 2 April 2025
Accepted 29 August 2025
Published 3 October 2025
10.1126/sciadv.adx7408