Table S1 Primers used in this study

Target	FWD	REV	Amplicon size	Primer efficiency	Present in strains
16S	TAA GCC GAG GGC TTT CAC AT	AGA GAC TGA TCG GCC ACA TT	323	98%	VMC1-8
Phage 1	GTG CAA GCT TCA TCA ACA TC	AGT ATG AGC CAG CTG GTA AT	203	110%	VMC1-8
Phage 2	ATT GTT AGA TTG GCA TGC GC	GCA TAA TTT GCT TCT GCA GC	321	105%	VMC1-8
Phage 3	GAA GTT TGG GTG CTG ATG TC	TTG ATC ACT TGC ATA GCT GC	332	106%	VMC1, 4, 5, 6, 8
Phage 4	CAT GGG CTA TTT CTA TGC CA	TCG GTA AAT TGC CAG ATT GC	257	86%	VMC4, 6
Phage 5	CAT GGG CTA TTT CTA TGC CA	ATT GAG AAC TAC CCG AAC TA	294	88%	VMC2

Table S2 Participant metadata

	VMC1	VMC2	VMC3	VMC4	VMC5	VMC6	VMC7	VMC8
Age	24	34	41	39	38	29	47	29
Ethnicity[1]	AA	AA	AA	AA	NA	C	C	C
Pregnant	No	Yes	No	No	NA	No	No	No
Smoker	NA	NA	No	No	NA	Yes	No	No
Sexual partners(#)	11-20	11-20	>=21	6-10	NA	3-5	1	3-5
Sexual partners (gender)	Male	Male	Male	Male	NA	Male	Male	Male
Birth control	OCP	NA	Tubal ligation	OCP, IUD	NA	OCP, Sprintec	NA	IUD
Clinical dx	Abnormal discharge	NA	BV	NA	Normal	Normal	NA	NA
BV history	Yes	Yes	Yes	Yes	NA	No	No	No
Vaginal discharge[2]	Yes	Yes	Yes	Yes	NA	No	No	No
Vaginal odor[2]	Yes	Yes	Yes	No	NA	No	No	No

Vaginal itching[2]	Yes	Yes	Yes	Yes	NA	No	No	No

Participant metadata was collected from health history questionnaires. NA is the designation for “Not Answered”.

[1] Self-reported ethnicity: African American (AA) or Caucasian (C).

[2] Symptoms reported by the participant at the time of sampling.

Table S3 Insertion sequences identified in strains

Strain designation	Insertion Sequence Family[1]	Number of IS
VMC1	IS4_ssgr_IS4	4
	IS200_IS605_ssgr_IS1341	2
	ISL3	5
	ISLre2	1
	IS607	1
	ISNCY_ssgr_IS1202	1
	IS256	1
	IS4_ssgr_ISPepr1	2
	ISNCY	3
	IS66	1
	IS30	5
	IS3_ssgr_IS150	4
	IS200_IS605	5
	IS1182	5

	IS982	3
	IS110	6
VMC2	IS4_ssgr_IS4	4
	IS200_IS605_ssgr_IS1341	2
	ISL3	2
	IS5_ssgr_ISL2	1
	ISLre2	1
	IS607	1
	ISNCY_ssgr_IS1202	1
	IS256	18
	IS4_ssgr_ISPepr1	3
	ISNCY	3
	IS66	2
	IS30	14
	IS3_ssgr_IS150	4
	IS200_IS605	12

	IS1182	5
	IS982	17
	IS110	12
VMC3	IS4_ssgr_IS4	1
	IS200_IS605_ssgr_IS1341	2
	ISL3	4
	ISLre2	1
	IS256	14
	IS4_ssgr_ISPepr1	4
	ISNCY	2
	IS30	6
	IS3_ssgr_IS150	2
	IS200_IS605	1
	IS1182	4
	IS982	11
	IS110	1

VMC4	IS4_ssgr_IS4	4
	IS200_IS605_ssgr_IS1341	6
	ISL3	3
	ISLre2	8
	IS607	2
	ISNCY_ssgr_IS1202	1
	IS256	64
	IS4_ssgr_ISPepr1	8
	IS66	1
	IS30	21
	IS3_ssgr_IS150	6
	IS200_IS605	3
	IS1182	9
	IS982	17
	IS110	8
VMC5	IS4_ssgr_IS4	2

IS200_IS605_ssgr_IS1341	2
ISL3	5
IS5_ssgr_ISL2	1
ISLre2	1
IS607	1
ISNCY_ssgr_IS1202	1
IS256	45
IS4_ssgr_ISPepr1	5
ISNCY	3
IS66	2
IS30	13
IS3_ssgr_IS150	4
IS200_IS605	5
IS1182	8
IS982	33
IS110	8

VMC6	IS4_ssgr_IS4	4
	IS200_IS605_ssgr_IS1341	9
	ISL3	4
	IS5_ssgr_ISL2	1
	ISLre2	3
	IS607	1
	ISNCY_ssgr_IS1202	1
	IS256	64
	IS4_ssgr_ISPepr1	4
	ISNCY	3
	IS66	2
	IS30	25
	IS3_ssgr_IS150	7
	IS200_IS605	6
	IS1182	2
	IS982	13

	IS110	6
VMC7	IS4_ssgr_IS4	3
	IS200_IS605_ssgr_IS1341	3
	ISL3	2
	ISLre2	2
	IS607	1
	ISNCY_ssgr_IS1202	1
	IS256	78
	IS4_ssgr_ISPepr1	6
	ISNCY	2
	IS66	1
	IS30	7
	IS3_ssgr_IS150	3
	IS200_IS605	3
	IS1182	5
	IS982	20

	IS110	4
VMC8	IS4_ssgr_IS4	3
	IS200_IS605_ssgr_IS1341	5
	ISL3	5
	IS5_ssgr_ISL2	1
	ISLre2	1
	IS607	2
	ISNCY_ssgr_IS1202	1
	IS256	77
	IS4_ssgr_ISPepr1	3
	ISNCY	2
	IS66	1
	IS30	20
	IS3_ssgr_IS150	8
	IS200_IS605	7
	IS1182	2

IS982	23
IS110	7

[1] Descriptions of IS families can be found through the IS Finder database (https://www-is.biotoul.fr/)

Table S4 Genes exclusive to BV strains

Gene Locus	Gene annotation	Amino acid sequence
AEL93_09295	PTS cellobiose transporter subunit IIA	MAEEKKTPEQKEQETLMAAMGLIANGGNAKSLAFEAIRLAKKGDIAGAREKLKESDKSLLEAHNSQTGMLTKEAQGDHMHVTLLVVHSQDHLMNAITFRDLAGEMVDLYEKLYESGSLKKEDK
AEL93_10590	PTS cellobiose transporter subunit IIA	YFTVSRHEDSIKYMYFSRYLLIRYIITIFFFTNLMWLIIDVNYHSVLGIIVSAIMTIYSGIASIEQLTKMHNRKREVPISKVYLEVQAALNLLFIMLTFLPLGKYLFPFIENQSIMFFMTTLFLAGILLCVWSEYRIHQIMNDQDRYHKVIETFKKHQQ
AEL93_09300	PTS cellobiose transporter subunit IIB	MAEQTIMLNCSAGMSTSLLVTKMQAAAKEQGIDAEIFACPASEADDKMAQQEIDCVLLGPQVSYMKSDFENKVKGKGKDGKDIPLDVINMQDYGMMNGANVLAQAEKLIKGE
AEL93_09780	6-phospho-beta-glucosidase	MSFTKEFTWGGATAANQYEGGYDEGCKGLNAVDVLTNGSATEPRKVTWKKPNGETGATPLVWGQEFSLPKGAVPTILDGYYYPSHQGTDFYHHYQEDTKLMADMGFDVFRLSMNWSRILPNGDDEQPNEDGLAFYDKVFDECAKYGIEPLVTLSHYETPLSLITRYGGWKDRHLIDAFVHYSDIVM

		KHYKGKVRYWLTFNEINAMDMAPYMGGGLIDGSEQNRAQGAHNQFVASAKVVKLAHEIDPNNRVGQMLAYSAYYPYTCDPADQLKVLEAKQEMLFYSDVQTGGRYPNYRLKKYERDGIKLNDTPEDYELIAKYPADFLSFSCYTSNVLTTHEADAKANGNVSAGGVVNPYLESNAWGWATDPDVLRIGLNELWDRYHKPLWVVENGLGSADTLEKDGSIHDNYRIDYLRDQIKSMRDAVTIDGVDLMGYTTWSAIDLVSNGTGEMKKRYGFVYVDRDDRGNGSLKRYPKDSFYWYKKAITSNGEDLD
AEL93_09785	hypothetical protein (**ligand-binding domain superfamily**)	MITITKQFNFKAKDFFDYLEEnQLIPAIKKARGNDMPVTLAKGTKYEVDGAQVKLTDYERNKVYGAHFKTDRMELVINYVTEDNDQGVKITFSEDMLSFDREKHGKLQTMFYNFQLKMGAKKELRRMGDNVLANIAA
AEL93_09790	PTS cellobiose transporter subunit IIC	LTSTMRITTFTKTKMVAGGYEASTNGITRIAAVGIFTGLVIAWLTVQIYRYTVKHNWRIKMPASVPSGVSNSFSALIPGFCIAVVVALIELILVTLGTDIFQVLYIPFSFISAIADTWWGFLIIIFLIHFLWWFGIHGATIMSSFYTPIVLANMAANVNGANHFFAGDPMNSFVIIGGSGATLGMAIWLAFGSRSAQLKEIGKVELVPAIFNINEPLLFGLPIVYNINLLVPFICAPLASGLVGYIAVTSHLVPKIIVQQPWPTPVGLSGYLATTSWQGAVLSIVCALVAFLVWFLFIKH

		YDNVLLKKEQADATKN
AEL93_10595	PTS cellobiose transporter subunit IIC	MADSKSGSFKDALAAKAGKFAGSRFVRAIMDAGYSVISFSIIGAVFLILNVLPQAFPIPGFAAFYANTLGRFSGLFQVVYNSTMGILALVFAGTFAYSYTDIYRREEKIDLVPMNGLMMFLMAFFITFPELIWKNGAVQFITTRVVLVK