#### Supplemental table 1: Relative fold-change from naïve control of each cytokine. 1

	•	
	,	
4		

	IL-4*	IL-5*	IL- 13*	IFN-γ*	IL- 1β*	IL-6*	IL- 12p70*	GM- CSF*	TNF- α*	IL- 18*	Median Survival**
Naive	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	100
UgCl447	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	100
UgCl377	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	77
UgCl262	0.00	0.00	0.00	0.00	0.00	0.00	-0.51	0.00	0.00	0.00	82.5
UgCI438	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-0.90	100
UgCI466	0.00	0.00	0.00	2.32	0.00	0.00	2.07	0.00	2.57	0.00	100
UgCI250	0.00	0.00	0.00	14.54	0.00	2.22	0.00	1.38	0.00	0.00	100
UgCI549	0.00	0.00	0.00	0.00	2.25	0.00	0.00	1.39	0.94	0.00	100
UgCl332	0.00	0.00	0.00	0.00	0.00	0.00	0.00	8.78	0.00	0.00	77
UgCl382	14.42	4.14	3.94	0.62	8.97	0.00	0.00	0.00	0.00	0.00	13
UgCl243	18.47	0.00	3.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	11
UgCI538	24.73	4.16	6.68	0.00	0.00	0.00	0.00	0.00	0.00	-0.77	18
UgCI450	21.44	4.77	4.22	1.25	12.77	6.76	0.00	0.00	2.70	4.37	10
UgCI534	0.00	3.02	0.00	0.00	5.15	20.44	0.00	0.00	0.00	0.00	18
UgCI547	6.12	2.85	0.00	0.00	7.56	17.87	0.00	0.00	0.00	0.00	19
UgCl292	10.59	2.90	0.00	0.00	6.40	0.00	0.00	0.00	0.00	0.00	46
UgCI300	18.25	12.49	6.14	34.66	12.25	20.46	0.00	1.74	3.33	5.94	15
UgCl326	16.08	6.79	22.11	0.00	41.57	5.56	0.00	0.00	0.00	0.00	4.5
UgCI468	41.84	8.75	9.71	0.00	12.86	56.12	0.90	3.88	3.05	0.00	8
UgCI546	129.60	55.27	50.24	0.00	15.63	26.07	0.00	2.82	2.46	1.26	35
ΚΝ99α	40.81	12.01	5.80	0.00	13.68	26.00	0.00	0.00	0.00	0.00	0
UgCI462	30.01	6.10	4.08	99.58	10.69	87.51	0.00	4.34	0.00	2.28	11
UgCI541	34.29	14.09	4.19	0.00	33.54	166.35	0.00	7.24	0.00	0.00	-5
UgCl357	49.60	5.83	7.10	17.35	13.92	34.69	104.84	119.05	24.38	1.78	0
UgCl360	0.00	0.00	0.00	267.76	28.34	42.97	1.05	3.66	10.19	3.42	15
UgCl495	0.00	0.00	0.00	538.60	26.56	45.28	0.00	3.07	0.14	4.90	25.5
UgCl422	0.00	5.90	0.00	835.13	44.26	81.18	2.12	0.00	17.43	32.44	-3
UgCl236	0.00	0.00	3.49	802.91	0.00	0.00	0.00	6.91	0.00	0.00	-5
UgCI535	0.00	0.00	0.00	943.10	28.43	54.53	1.12	5.04	9.01	0.00	-1.5
UgCl390	0.00	0.00	0.00	609.20	31.80	69.81	0.00	6.40	11.46	0.00	14
UgCl247	0.00	0.00	3.89	1923.82	27.92	71.07	0.00	9.23	0.00	9.48	-2

3

\*Number is from three technical replicates each from five mice \*\*Median survival from 5 or 10 mice

5 Supplemental table 2: Ordinal categorical variables association table for ST93A GWAS

	Marker	Alt Freq	Alt Counts	Missing Rate	Р	beta	seBeta
					value		
	4_995956	0.18	6	0.15	0.0005	1.98	0.63
L-DOPA 30°C	6_989732	0.29	10	0.15	0.0026	-1.48	0.51
	12_503401	0.42	16	0.05	0.0047	1.24	0.44
	13_11108	0.15	6	0	0.0017	2.34	0.79
SDS	5_836479	0.25	10	0	0.0024	1.95	0.65
	1_975397	0.39	14	0.1	0.0033	1.90	0.65
	7_165873	0.35	14	0	0.0003	-1.75	0.50
Growth on 37°C	12_503049	0.16	6	0.05	0.0017	2.03	0.67
	12_15014	0.39	14	0.1	0.0019	1.68	0.55
Virulence	9_6619	0.25	10	0	0.0014	2.63	0.84
category	10_15302	0.33	12	0.1	0.0033	-1.99	0.70

#### 6 from variables analyzed using a proportional odds logistic mixed model.

## 8 Supplemental table 3: SNP and gene results from three GWASs

	ST93 all SNPs Human GWAS	ST93 all SNPs Mouse GWAS	ST93A Mouse GWAS
SNPs	207	161	108
Genes	115	75	45
Filtered SNPs <sup>+</sup>	145	127	93
Filtered genes+	40	38	32
SNPs from human and mouse		42	34
Genes from human and mouse		16	17

9

+ Gene with only one significant SNP associated with only one trait were filtered from analysis

Gene	Predicted Phenotype
CNAG_00012	Hypothetical; orthology to: oxioreductase or NAD binder
CNAG_01241	Enzyme regulator
CNAG_01461	Sodium/bile acid co-transporter
CNAG_01491	Hypothetical; orthology to: ankyrin repeat protein
CNAG_02176	Hypothetical
CNAG_02475	Flavin-containing monooxygenase
CNAG_02487	phs1
CNAG_03387	Hypothetical
CNAG_04101	Oxidation of branched chain-fatty acids
CNAG_04922	Hypothetical
CNAG_05185	Hypothetical
CNAG_05329	myo-inositol 2-dehydrase
CNAG_05661	FACT complex subunit
CNAG_05662	itr4
CNAG_05664	Branched chain amino acid transaminase
CNAG_05746	Swiss dependent recombination DNA repair
CNAG_05913	Alpha glucosidase
CNAG_05937	Hypothetical
CNAG_05987	Hypothetical
CNAG_06169	Hypothetical; orthology to s-glutathione dioxygenase
CNAG_06256	Hypothetical
CNAG_06574	app1
CNAG_06876	Alpha ketoglutarate-dependent taurine dioxygenase
CNAG_07497	Hypothetical
CNAG_07528	DNA binding protein; putative transposon
CNAG_07586	Hypothetical
CNAG_07728	Solute carrier family 3a (Zinc transporter)
CNAG_07748	t-RNA methyltransferase
CNAG_07837	Hypothetical
CNAG_07874	Sugar transporter
CNAG_07950	Hypothetical
CNAG_08006	Hypothetical

# 10 Supplemental table 4: Function of genes identified in mouse GWAS

	ΚΝ99α*			itr4∆*		
PHS1	11.52	11.70	11.71	9.04	9.07	8.87
CNAG_04101	11.84	11.90	11.78	11.72	11.62	11.73
CNAG_05664	11.60	11.58	11.65	0.00	0.00	0.00
ITR4	14.69	14.68	14.67	0.00	0.00	0.00
CNAG_05329	8.36	8.55	8.34	8.00	7.90	8.17
CNAG_01241	10.94	11.19	11.01	11.33	11.25	11.47
CNAG_07950	3.63	2.25	2.45	2.48	4.07	2.45
CNAG_07528	4.63	3.84	4.36	3.37	4.23	5.22
CNAG_06256	-0.18	-0.07	1.45	2.84	2.79	2.71

12 Supplemental table 5: RNAseq Transcript Data

13 \*Average of three technical replicates

## 14 Supplemental table 6: Primers used in this study

Primer name	Sequence (5'-3')	Description
CX5	GTAAAACGACGGCCAG	M13F
CX6	CAGGAAACAGCTATGAC	M13R
JH8994	TGTGGATGCTGGCGGAGGATA	JH8994
CX249	GAGGCTATTCGGCTATGACTGG	NEO split F
CX2281	AGGCTGGCAAATCAAGCGTG	ITR4 F1
CX2282	CTGGCCGTCGTTTTACCGTGGTGATGGTGGTGCTCGAG	ITR4 R1
CX2283	GTCATAGCTGTTTCCTGACGGGGGAGAAAGGCGTAGAGG	ITR4 F2
CX2284	CAACTTCCCAATACATCATG	ITR4 R2
CX1010	CTGAGATTGCTCCCGCTAGG	ITR4 F3
CX1011	AATGAACACCACCCAAGCCA	ITR4 R3
CX1009	TGCAGTTTACATTTCCAATCGTC	ITR4 F4
CX2285	CAGGTATTTGACTAGTCTGC	ITR4 R4
CX1419	TACCGAGCTCGGATCCGACTCTCACGTCGTTGTATAC	pJAF1-ITR4 F
CX1420	CGTTACTAGTGGATCCTCCAAGCACCCATCACTACATG	pJAF1-ITR4 R
CX1800	CAACATGTCTGGATCCATGTCCACGCTTGACTACAAG	pCXU200-ITR4 F
CX1801	TAGAACTAGTGGATCCAGCCTTCGACCGCTTTTCATT	pCXU200-ITR4 R
CX115	CATCGCTTCCGCATTCACTCACTC	ITR4 QPCR F
CX116	TTGCCGGTACCCTTGACGATAACA	ITR4 QPCR R
CX2365	CACCGATGGTTCATCCCTATAC	CNAG_05664 QPCR F
CX2366	GGACATAGCGTATCCCTTCTTC	CNAG_05664 QPCR R









23 Supplemental figure 2: Mice infected with clinical isolates showed four distinct virulence manifestations. Five or 10 mice were infected with 5x10<sup>4</sup> cells of each clinical isolate and 24 25 monitored for 100 days. Mouse virulence was categorized into four categories, depending on 26 relative median survival and CFUs in the mouse brain. A) All survival curves are shown. The 27 survival curves were normalized to a matched KN99a control. B) Mouse CFUs were collected 28 from the lungs and brain of mice at terminal endpoint. For mice infected with typical CNS, typical 29 non-CNS, and hypervirulent isolates, CFUs are all from mice that succumbed to infection rather than mice sacrificed at 100 days post infection. For mice infected with latent isolates, the CFUs 30 are from mice that survived to 100 days post infection. Significance was determined using 31 32 Kruskal-Wallis nonparametric test with Dunn's multiple comparison correction. # marks mice 33 that are not significantly different from the KN99α control. (n=3-10; exact P-values provided in Source Data file). Source data are provided as a Source Data file. 34



Supplemental figure 3: *in vitro* phenotypes do not correlate with disease manifestation. *in vitro* phenotype data was collected for each clinical isolate. For cell growth stressors and cell wall and membrane stressors, cells were diluted from 10<sup>1</sup>-10<sup>4</sup> and spot plated in triplicate on media containing each stressor. Growth of each isolate was determined relative to growth on unadulterated YPD media. Melanin production was determined by spotting 10<sup>6</sup> cells in triplicate

40 on either Niger seed or L-DOPA plates and grown at either room temperature or at 30°C. Each isolate was given a score of 1-5 depending on the shade of brown or black for each colony. 41 Urease production was determined by plating 10<sup>1</sup>-10<sup>4</sup> cells on Christensen's urea agar and 42 determining the dilution at which the zone of clearance occurred. Cell morphology was 43 44 determined by growing cells in water supplemented with FBS for three days (62). Cells were imaged and 200 cells from each isolate was measured to determine cell body size, capsule 45 size, total cell size, and ability to form titan cells. Isolates were assigned a numerical score 46 depending on each size. Sample clustering was performed using an average clustering linkage 47 48 method and Euclidean distance measurement (63), and no associations were observed. Four representative isolates from each virulence category were selected and grouped manually in 49 order of virulence. The heat map was normalized to KN99α and each isolate was compared to 50 51 KN99α. Scale bar describes relative ability to grow under stress, produce effector proteins, etc. 52 in comparison to KN99α; i.e. "more" or "less" than KN99α. In general, there were no associations between in vitro phenotypes and disease manifestation. Source data are provided 53 as a Source Data file. 54



Supplemental figure 4: in vivo immune response associates with disease outcome. A) 56 57 Sample clustering was performed using an average clustering linkage method and Euclidean distance measurement (63) and normalized by row. B) Heatmap describing the normalized fold 58 change. Mice were sacrificed on day 17 or 21 post infection, and lungs removed. The lung 59 60 supernatant was collected, and cytokine levels determined using ThermoFisher Th1/Th2 61 cytokine panel and a Luminex Magpix. The fold change from a naïve control was calculated for each sample that had a significant (P < 0.05) increase in cytokines from uninfected mice. 62 63 Significance was determined using Kruskal-Wallis nonparametric test with Dunn's multiple comparison correction. Sample clustering was performed using an average clustering linkage 64 method and Euclidean distance measurement (63). The heat map was normalized. 65 Representative samples are shown. Median survival was normalized to KN99α survival. Latent 66 isolates (purple) clustered with the uninfected control and showed a mostly undetectable 67 68 immune response. Typical CNS (yellow) and typical non-CNS (blue) isolates generally showed 69 a type 2 immune response, with increases in IL-4, IL-5, and IL-13, and clustered with KN99a. 70 Hypervirulent isolates (red) generally showed an increase in type 1 cytokines, including IFNy, or

- proinflammatory cytokines, including IL-1 $\beta$ , IL-6, and GM-CSF. Source data are provided as a
- 72 Source Data file.







- 85 Supplemental figure 6: Principal component analysis on all mutations in the ST93
- 86 **population.** A PCA was performed using all SNPs/INDELs in the ST93 isolates. No clusters
- 87 were observed.



- 90 SNPs in population. B) Scree plot for 562 SNPs predicted to influence gene coding or
- 91 regulation.

92

89



Supplemental figure 8: K-means clustering revealed three clusters of isolates. Distance between points was calculated with Euclidean distance. Gap statistic (65) and silhouette score (66) were used to calculate the ideal number of clusters. A) For all variates, gap statistic showed no clusters and K-means suggested that K=2 was the optimal number of clusters. Each isolate was colored according to the K=2 assignment. B) The coding variants found that K=3 was the optimal number of clusters. Each isolate was colored corresponding to the K=3 assignment.



99Supplemental figure 9: Cytokine results from all deletion mutants. 5 mice were sacrificed100on day 17-20 post infection and lungs removed. The lung supernatant was collected, and101cytokine levels determined. Fold expression was calculated relative to an uninfected control.102Significance was determined using a Kruskal-Wallis nonparametric test without Dunn's multiple103comparison correction (\* = P < 0.05; \*\* = P < 0.01; \*\*\* = P < 0.001; n=5; exact p-values in

104 Source Data file). Source data are provided as a Source Data file.



- 105 Supplemental figure 10: SNP variants (highlighted) in CNAG\_07528, a putative DNA
- 106 transposase, associated with hypervirulence in clinical isolates. Signature features of DNA
- 107 transposon Class II mobile element sequences are indicated including an endonucleolytic DDE
- 108 catalytic domain and putative 44-bp terminal inverted repeats (TIRs) flanking the transposase
- 109 gene.