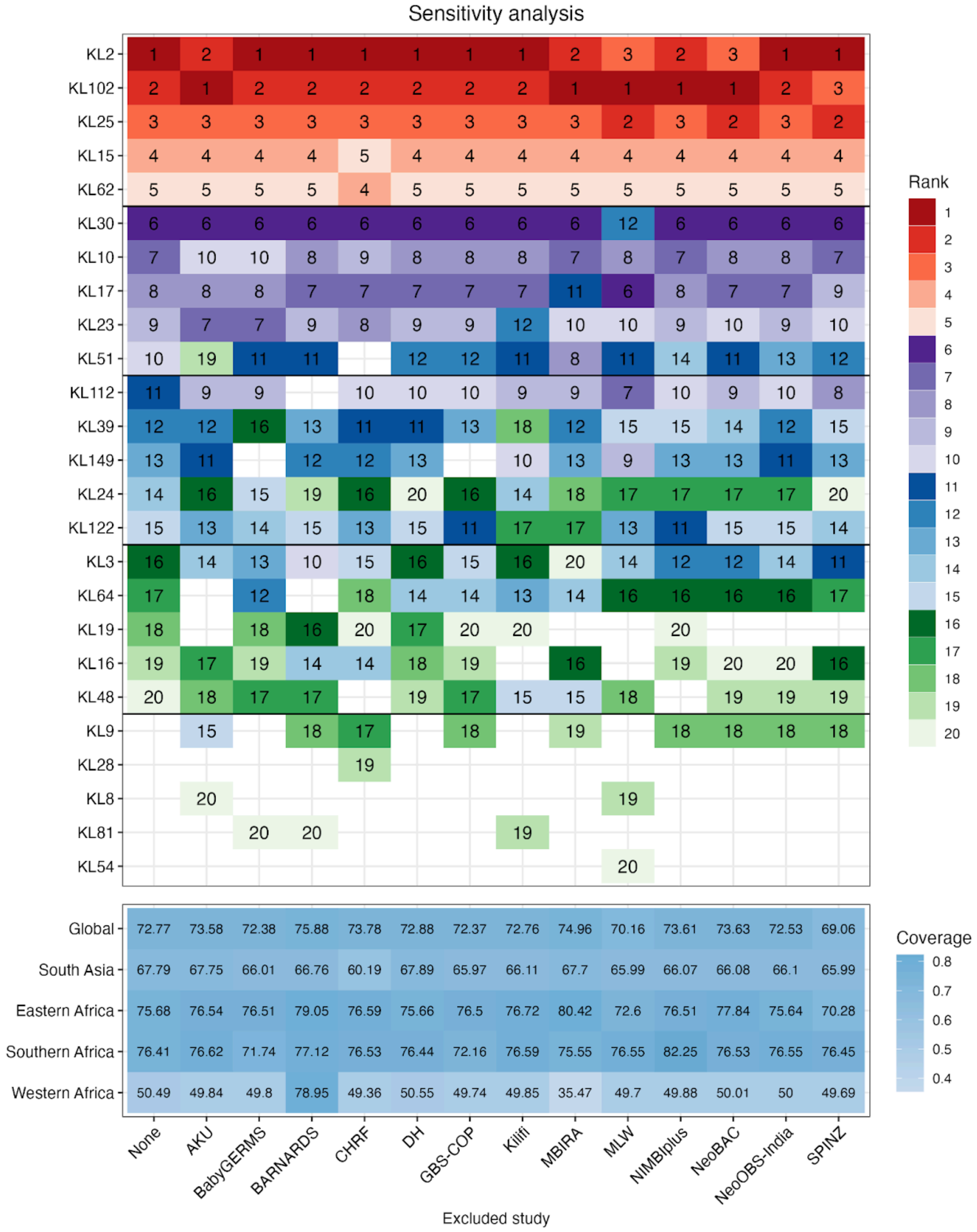
S2 Appendix: Sensitivity analyses.

S2.1 Appendix: Sensitivity of K locus global ranks to study inclusion.

Plots show the results of a leave-one-out analysis. In the first matrix, each column shows the global (overall) ranking of K loci, based on mean prevalence estimates modelled from cluster-adjusted counts, including either all studies (column 1) or all-but-one study (other columns). Each column is labelled by the study that was excluded from the meta-analysis. Cells are coloured (as per inset legend) to illustrate the rank of each K locus (row) in each meta-analysis (column). All K loci ranked in the top-20 for any meta-analysis are included. In the second matrix, cells show the coverage of the global top-20 from the full analysis.

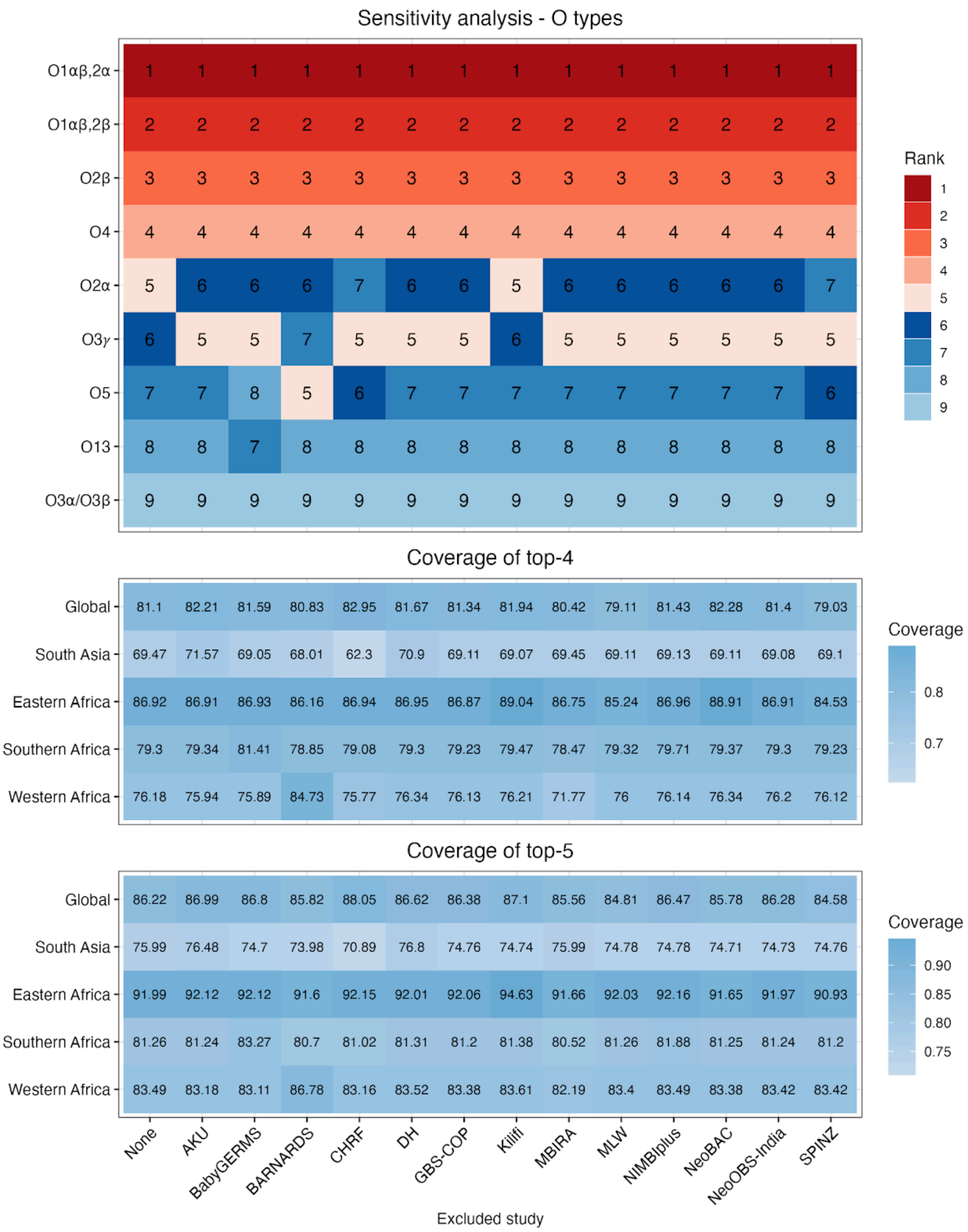


	None	AKU	BabyGERMS	BARNARDS	CHRF	DH	GBS-COP	Kilifi	MBIRA	MLW	NIMBIplus	NeoBAC	NeoOBS-India	SPINZ
KL2	1	2	1	1	1	1	1	1	2	3	2	3	1	1
KL102	2	1	2	2	2	2	2	2	1	1	1	1	2	3
KL25	3	3	3	3	3	3	3	3	3	2	3	2	3	2
KL15	4	4	4	4	5	4	4	4	4	4	4	4	4	4
KL62	5	5	5	5	4	5	5	5	5	5	5	5	5	5
KL30	6	6	6	6	6	6	6	6	6	12	6	6	6	6
KL10	7	10	10	8	9	8	8	8	7	8	7	8	8	7
KL17	8	8	8	7	7	7	7	7	11	6	8	7	7	9
KL23	9	7	7	9	8	9	9	12	10	10	9	10	9	10
KL51	10	19	11	11		12	12	11	8	11	14	11	13	12
KL112	11	9	9		10	10	10	9	9	7	10	9	10	8
KL39	12	12	16	13	11	11	13	18	12	15	15	14	12	15
KL149	13	11		12	12	13		10	13	9	13	13	11	13
KL24	14	16	15	19	16	20	16	14	18	17	17	17	17	20
KL122	15	13	14	15	13	15	11	17	17	13	11	15	15	14
KL3	16	14	13	10	15	16	15	16	20	14	12	12	14	11
KL64	17		12		18	14	14	13	14	16	16	16	16	17
KL19	18		18	16	20	17	20	20			20			
KL16	19	17	19	14	14	18	19		16		19	20	20	16
KL48	20	18	17	17		19	17	15	15	18		19	19	19
KL9		15		18	17		18		19		18	18	18	18
KL28					19									
KL8		20								19				
KL81			20	20				19						
KL54										20				

	None	AKU	BabyGERMS	BARNARDS	CHRF	DH	GBS-COP	Kilifi	MBIRA	MLW	NIMBIplus	NeoBAC	NeoOBS-India	SPINZ
Global	72.77	73.58	72.38	75.88	73.78	72.88	72.37	72.76	74.96	70.16	73.61	73.63	72.53	69.06
South Asia	67.79	67.75	66.01	66.76	60.19	67.89	65.97	66.11	67.7	65.99	66.07	66.08	66.1	65.99
Eastern Africa	75.68	76.54	76.51	79.05	76.59	75.66	76.5	76.72	80.42	72.6	76.51	77.84	75.64	70.28
Southern Africa	76.41	76.62	71.74	77.12	76.53	76.44	72.16	76.59	75.55	76.55	82.25	76.53	76.55	76.45
Western Africa	50.49	49.84	49.8	78.95	49.36	50.55	49.74	49.85	35.47	49.7	49.88	50.01	50	49.69

S2.2 Appendix: Sensitivity of O type global ranks to study inclusion.

Plots show the results of a leave-one-out analysis. Top matrix shows the results of a leave-one-out analysis. Each column shows the global (overall) ranking of O types, based on mean prevalence estimates modelled from cluster-adjusted counts, including either all studies (column 1) or all-but-one study (other columns). Each column is labelled by the study that was excluded from the meta-analysis. Cells are coloured (as per inset legend) to illustrate the rank of each O type (row) in each meta-analysis (column). Only the top nine O types are shown (the other six loci are excluded as they were rarely observed: O1αβ,2αγ, n=4; O2αγ, n=1; O12, n=2; O10, n=7; O14, n=1; O15, n=2). In the matrices below, cells show the coverage of the global top-4 or top-5 from the full analysis.



Sensitivity analysis - O types

	None	AKU	BabyGERMS	BARNARDS	CHRF	DH	GBS-COP	Kilifi	MBIRA	MLW	NIMBIplus	NeoBAC	NeoOBS-India	SPINZ
O1αβ,2α	1	1	1	1	1	1	1	1	1	1	1	1	1	1
O1αβ,2β	2	2	2	2	2	2	2	2	2	2	2	2	2	2
O2β	3	3	3	3	3	3	3	3	3	3	3	3	3	3
O4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
O2α	5	6	6	6	7	6	6	5	6	6	6	6	6	7
O3γ	6	5	5	7	5	5	5	6	5	5	5	5	5	5
O5	7	7	8	5	6	7	7	7	7	7	7	7	7	6
O13	8	8	7	8	8	8	8	8	8	8	8	8	8	8
O3α/O3β	9	9	9	9	9	9	9	9	9	9	9	9	9	9

Coverage of top-4

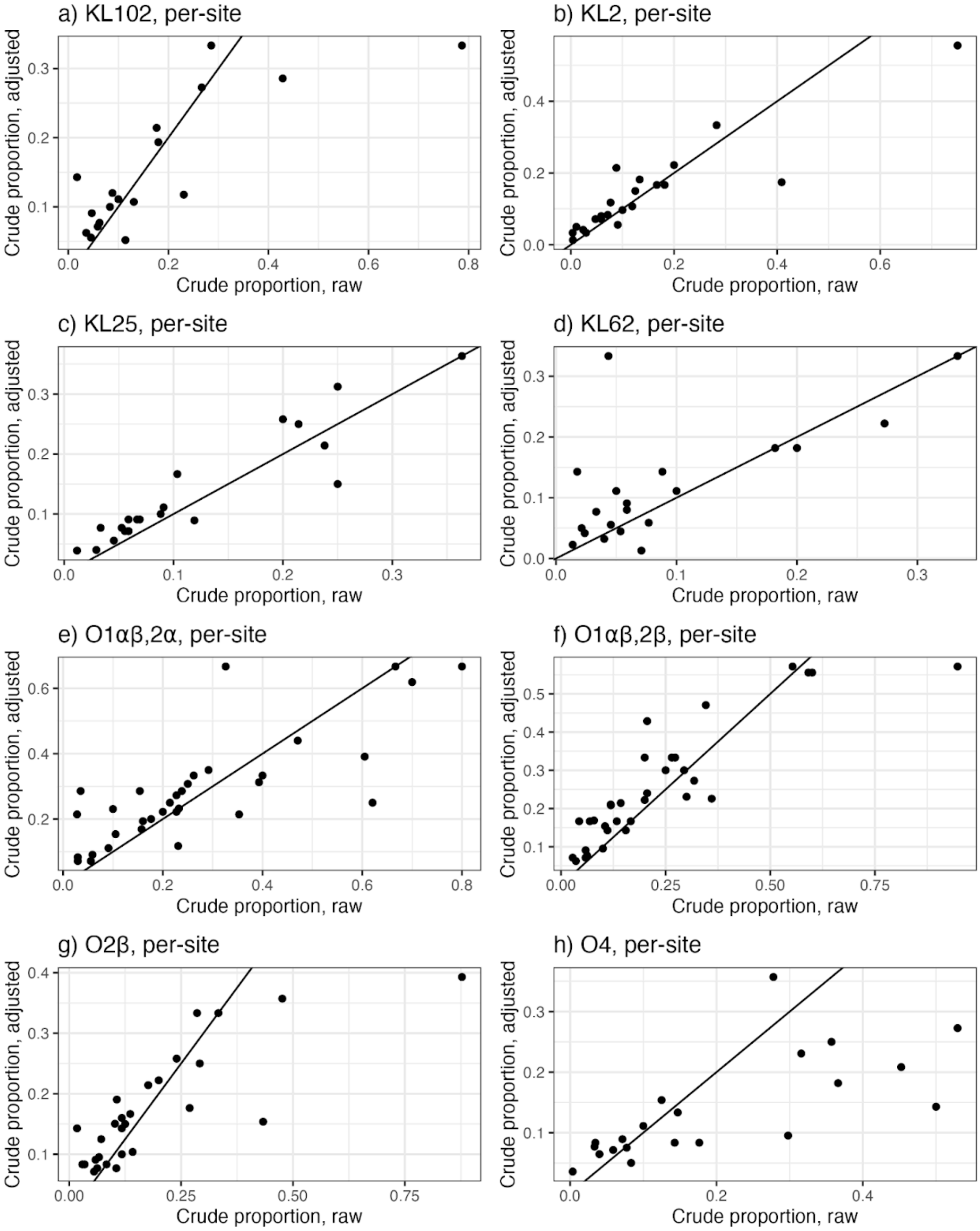
	None	AKU	BabyGERMS	BARNARDS	CHRF	DH	GBS-COP	Kilifi	MBIRA	MLW	NIMBIplus	NeoBAC	NeoOBS-India	SPINZ
Global	81.1	82.21	81.59	80.83	82.95	81.67	81.34	81.94	80.42	79.11	81.43	82.28	81.4	79.03
South Asia	69.47	71.57	69.05	68.01	62.3	70.9	69.11	69.07	69.45	69.11	69.13	69.11	69.08	69.1
Eastern Africa	86.92	86.91	86.93	86.16	86.94	86.95	86.87	89.04	86.75	85.24	86.96	88.91	86.91	84.53
Southern Africa	79.3	79.34	81.41	78.85	79.08	79.3	79.23	79.47	78.47	79.32	79.71	79.37	79.3	79.23
Western Africa	76.18	75.94	75.89	84.73	75.77	76.34	76.13	76.21	71.77	76	76.14	76.34	76.2	76.12

Coverage of top-5

	None	AKU	BabyGERMS	BARNARDS	CHRF	DH	GBS-COP	Kilifi	MBIRA	MLW	NIMBIplus	NeoBAC	NeoOBS-India	SPINZ
Global	86.22	86.99	86.8	85.82	88.05	86.62	86.38	87.1	85.56	84.81	86.47	85.78	86.28	84.58
South Asia	75.99	76.48	74.7	73.98	70.89	76.8	74.76	74.74	75.99	74.78	74.78	74.71	74.73	74.76
Eastern Africa	91.99	92.12	92.12	91.6	92.15	92.01	92.06	94.63	91.66	92.03	92.16	91.65	91.97	90.93
Southern Africa	81.26	81.24	83.27	80.7	81.02	81.31	81.2	81.38	80.52	81.26	81.88	81.25	81.24	81.2
Western Africa	83.49	83.18	83.11	86.78	83.16	83.52	83.38	83.61	82.19	83.4	83.49	83.38	83.42	83.42

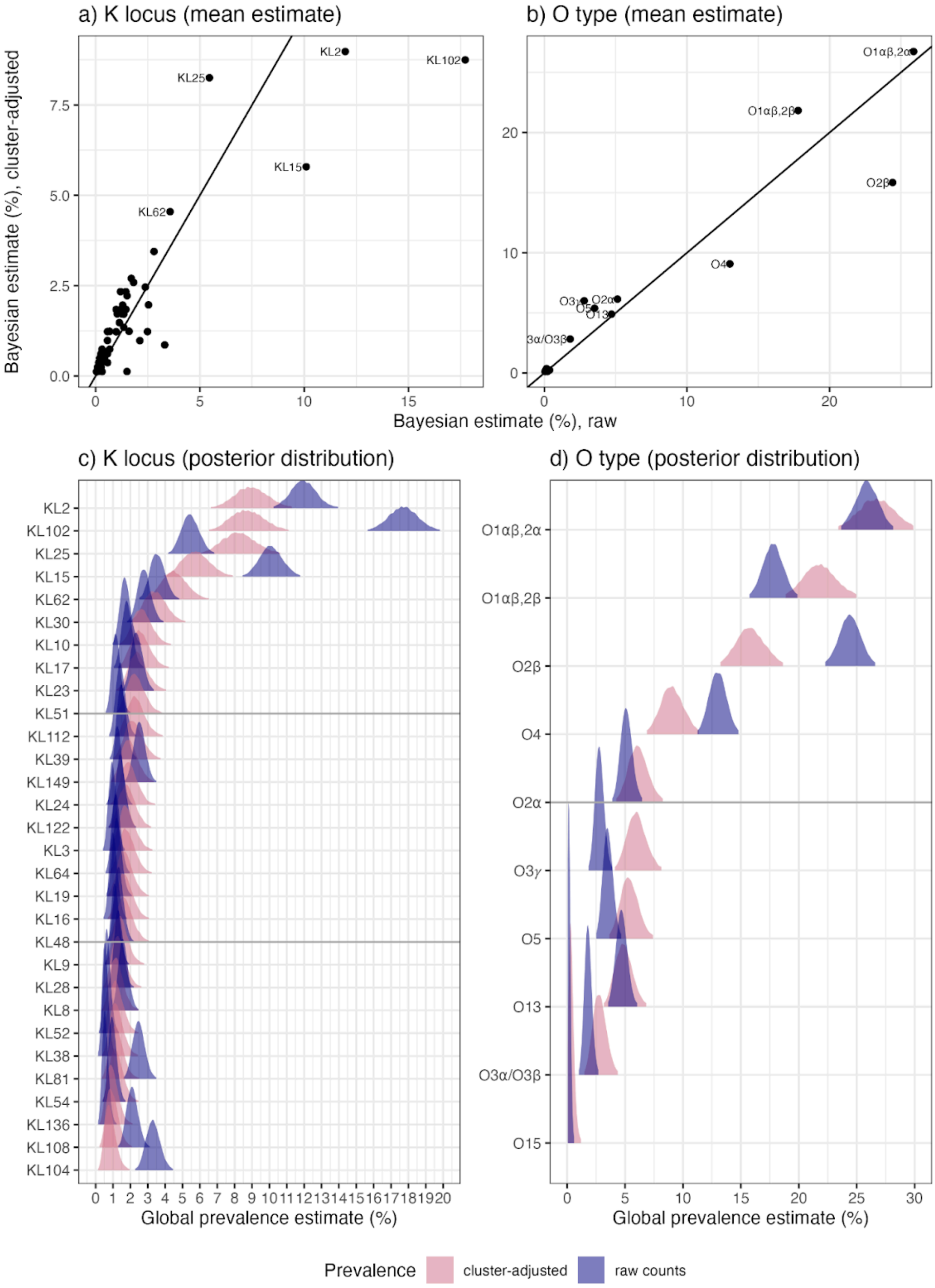
S2.3 Appendix: Effect of cluster adjustment on crude proportions per site.

Plots show crude proportion per site, estimated using raw (x-axis) vs cluster-adjusted (y-axis) counts, for the 4 most common K loci **(a-d)** and O types **(e-h)**. The impact of cluster-adjustment on the global prevalence estimates are shown in **S2.4 Appendix**.



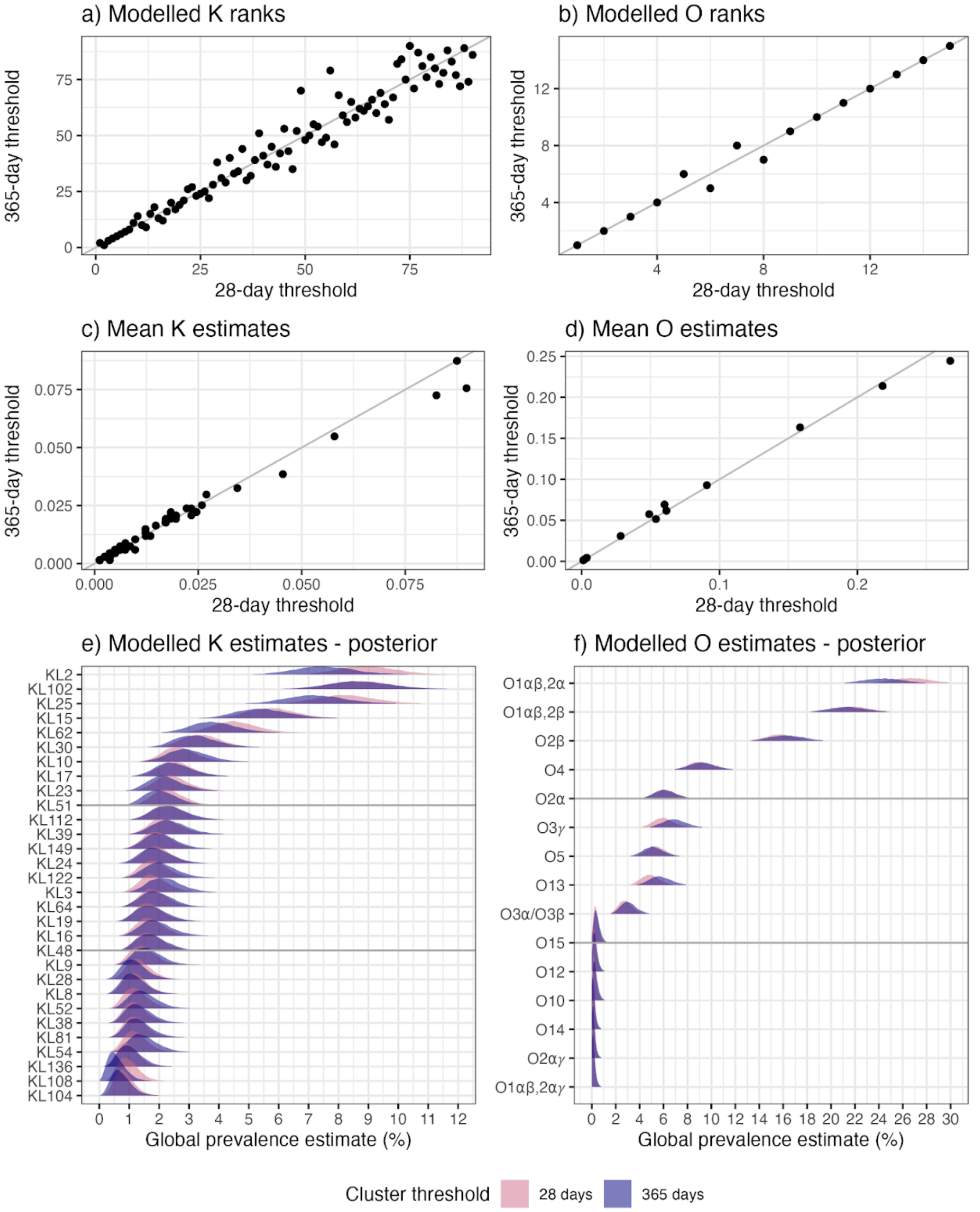
S2.4 Appendix: Effect of cluster adjustment on Bayesian modelled estimates of global K locus prevalence.

Global (overall) prevalence estimates modelled for **(a)** each K locus or **(b)** each predicted O type, using either raw counts (x-axis) vs cluster-adjusted counts (y-axis).



S2.5 Appendix: Sensitivity analysis showing effect of temporal clustering thresholds on K locus proportions and ranks (modelled global estimates).

The clustering threshold used for all primary analyses was 28 days; modelling was repeated using the 365-day threshold to assess sensitivity to clustering parameters. Data plotted here summarise the Bayesian modelled estimates for global prevalence, using counts adjusted for clustering using different temporal thresholds (≤28 days vs ≤365 days). In panels (**a-d**) each point represents a single K locus or O type. Panels (**e-f**) show the posterior distributions of prevalence estimates.



a) Modelled K ranks

b) Modelled O ranks

c) Mean K estimates

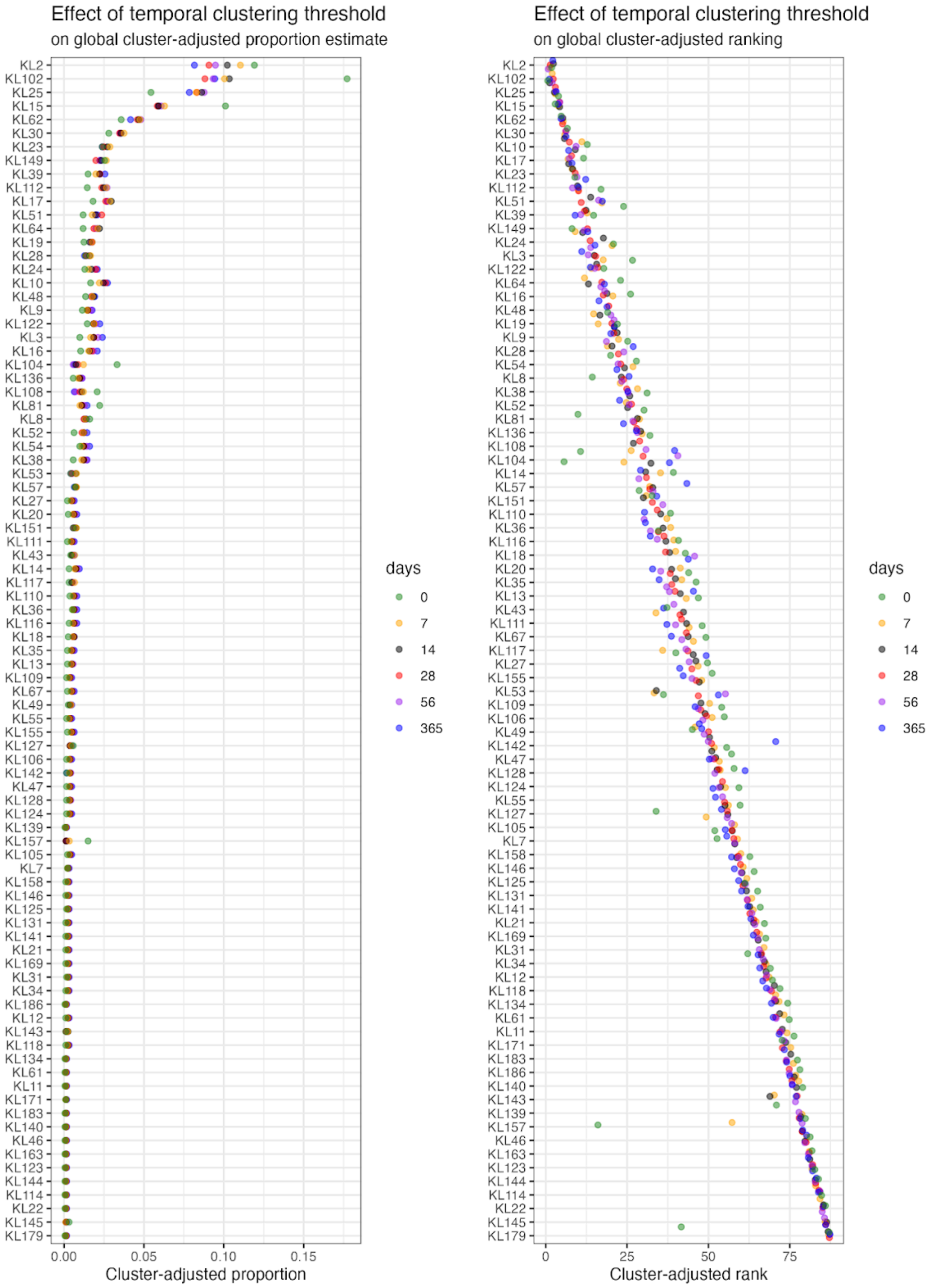
d) Mean O estimates

e) Modelled K estimates - posterior

f) Modelled O estimates - posterior

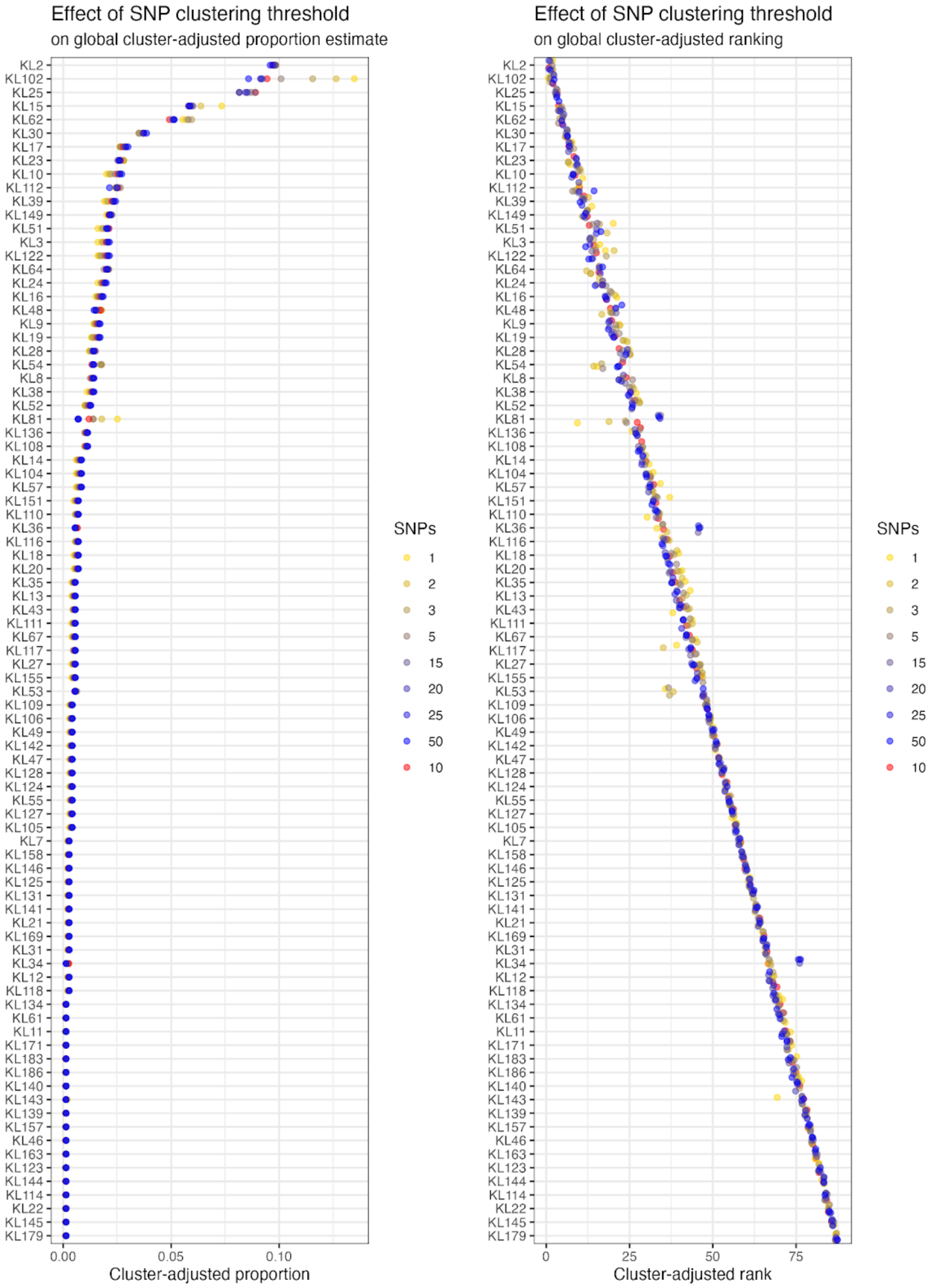
S2.6 Appendix: Sensitivity analysis showing effect of temporal clustering thresholds on K locus proportions and ranks (crude proportions).

Plotted data are based on crude proportions (rather than Bayesian modelled estimates of prevalence), using a SNP threshold of 10. The thresholds used for primary analysis were 28 days (red) and 10 SNPs.



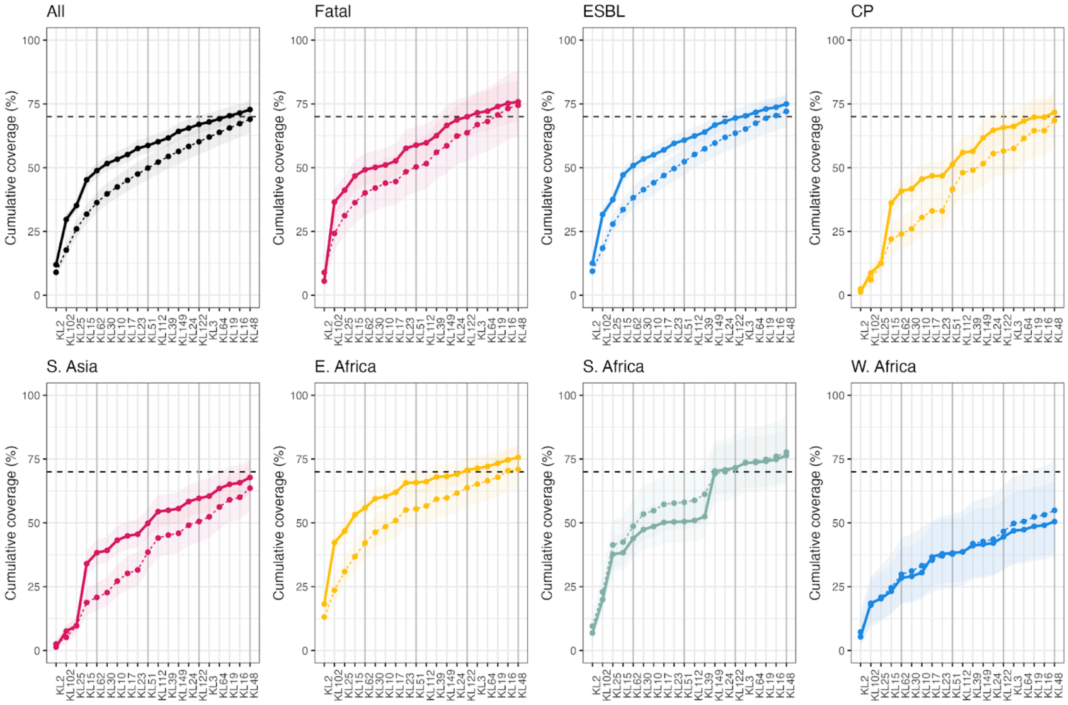
S2.7 Appendix: Sensitivity analysis showing effect of SNP clustering thresholds on K locus proportions and ranks (crude proportions).

Plotted data are based on crude proportions (rather than Bayesian modelled estimates of prevalence), using a temporal threshold of 4 weeks. The thresholds used for primary analysis were 10 SNPs (red) and 4 weeks.



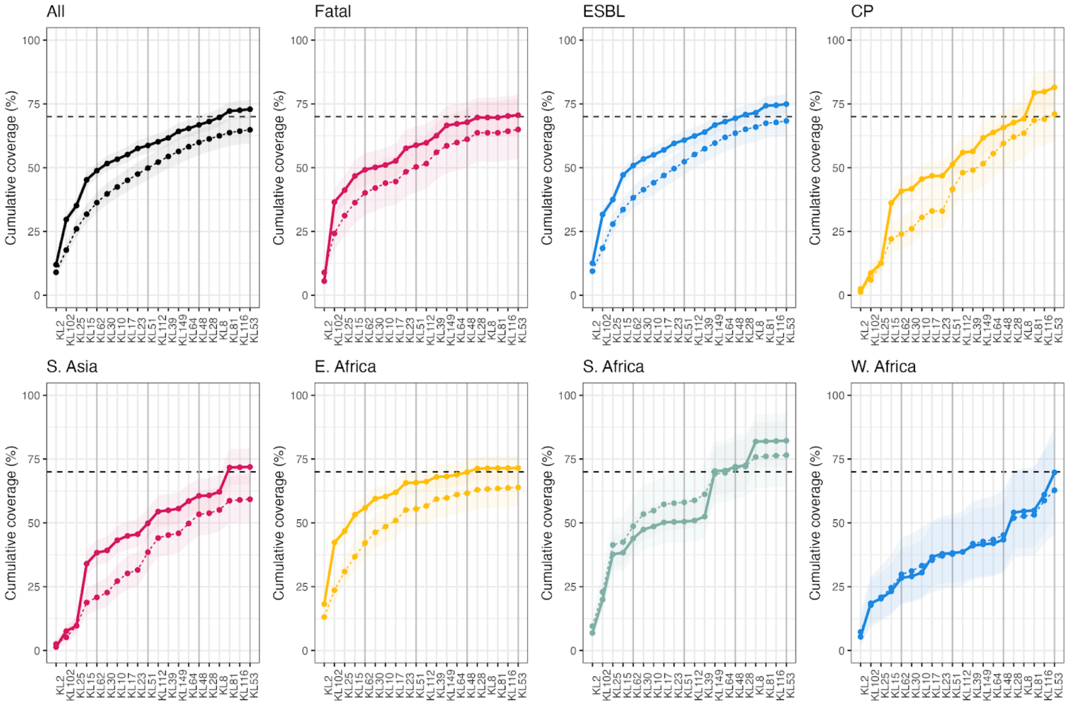
S2.8 Appendix: Coverage estimates for KL set 1 (global top 20)

Cumulative coverage estimates for the top 20 K loci (KL), ordered by cluster-adjusted global prevalence estimate. The first panel shows estimates for all cases (“All”), remaining panels show coverage estimated for the subset of cases indicated in the panel title (ESBL, extended-spectrum beta-lactamase; CP, carbapenemase-producing). Coverage estimated from raw counts are shown as solid lines, and from cluster-adjusted counts as dashed lines, each with 95% credible intervals shaded.



S2.9 Appendix: Coverage estimates for KL set 2 (top-8 per region)

Cumulative coverage estimates for KL set 2, comprising the top-8 K loci within each region, ranked according to cluster-adjusted global prevalence estimate. The first panel shows estimates for all cases ("All"), remaining panels show coverage estimated for the subset of cases indicated in the panel title (ESBL, extended-spectrum beta-lactamase; CP, carbapenemase-producing). Coverage estimated from raw counts are shown as solid lines, and from cluster-adjusted counts as dashed lines, each with 95% credible intervals shaded.



S2.10 Appendix: Coverage estimates for O types

Cumulative coverage estimates for the top 10 O types, ordered by cluster-adjusted global prevalence estimate. The first panel shows estimates for all cases (“All”), remaining panels show coverage estimated for the subset of cases indicated in the panel title (ESBL, extended-spectrum beta-lactamase; CP, carbapenemase-producing). Coverage estimated from raw counts are shown as solid lines, and from cluster-adjusted counts as dashed lines, each with 95% credible intervals shaded.