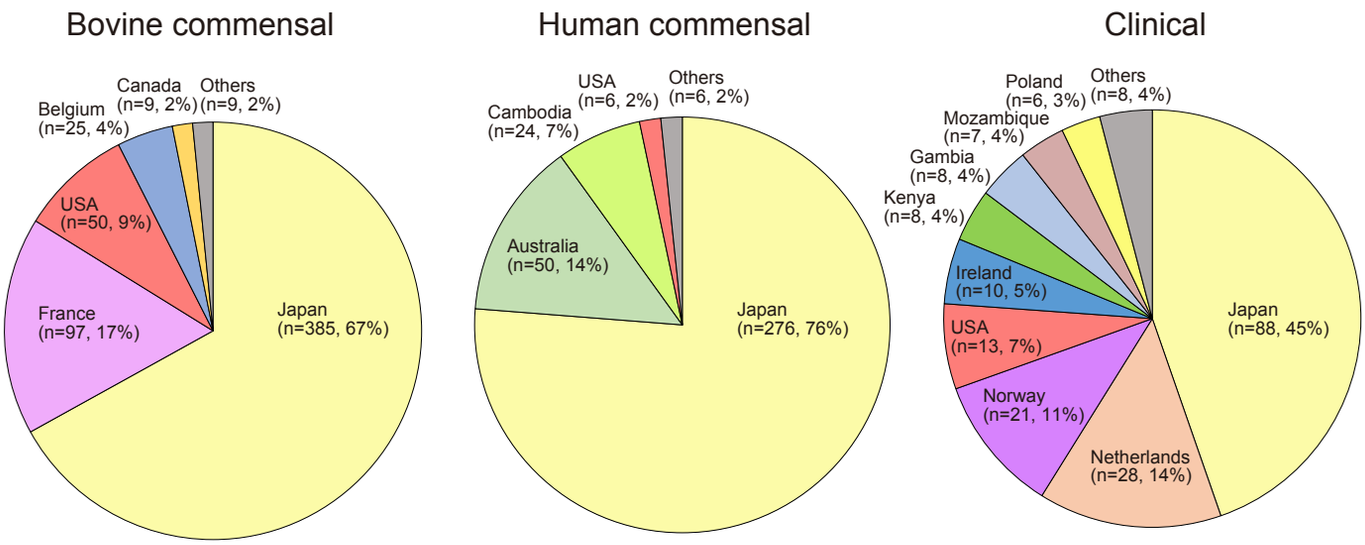
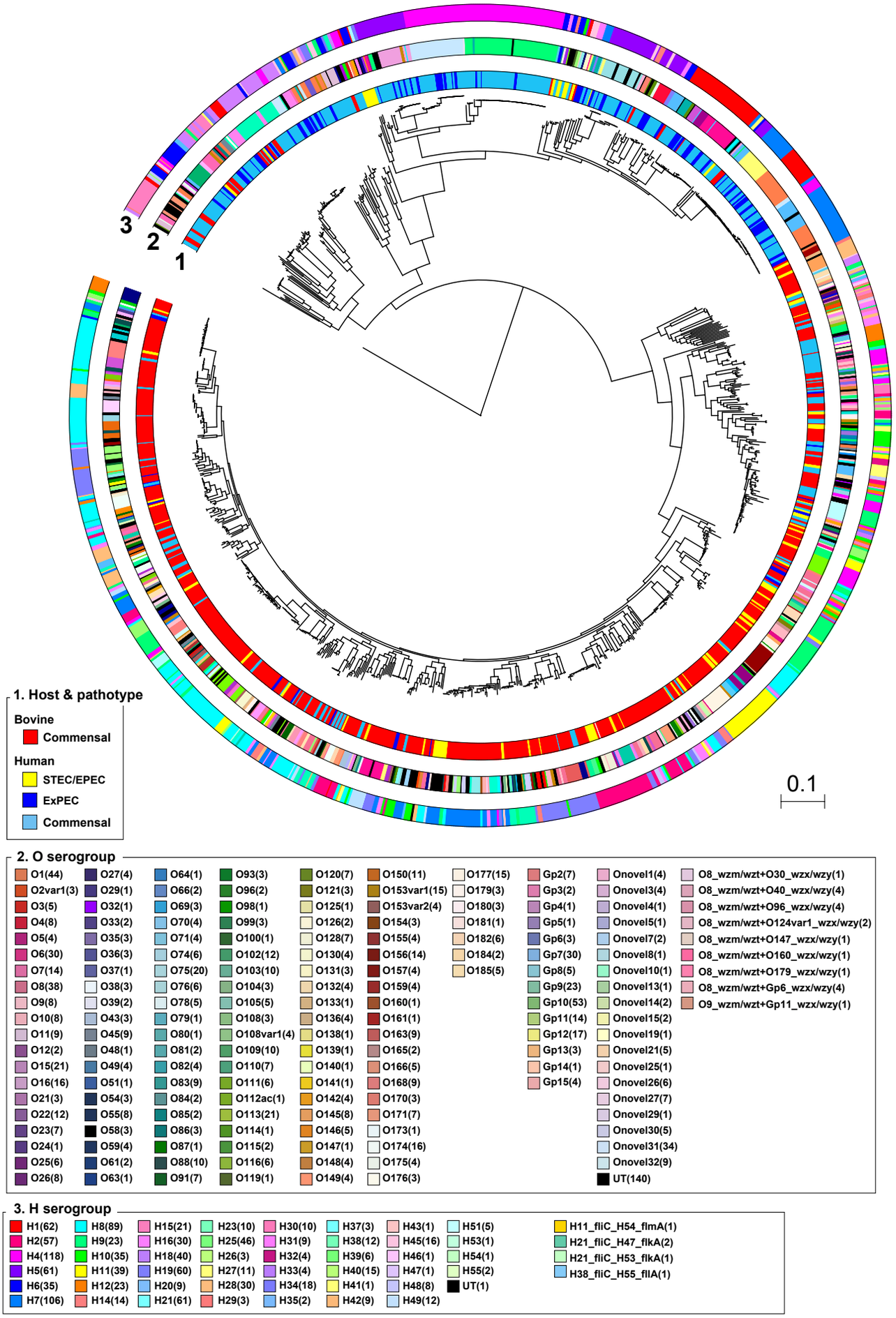
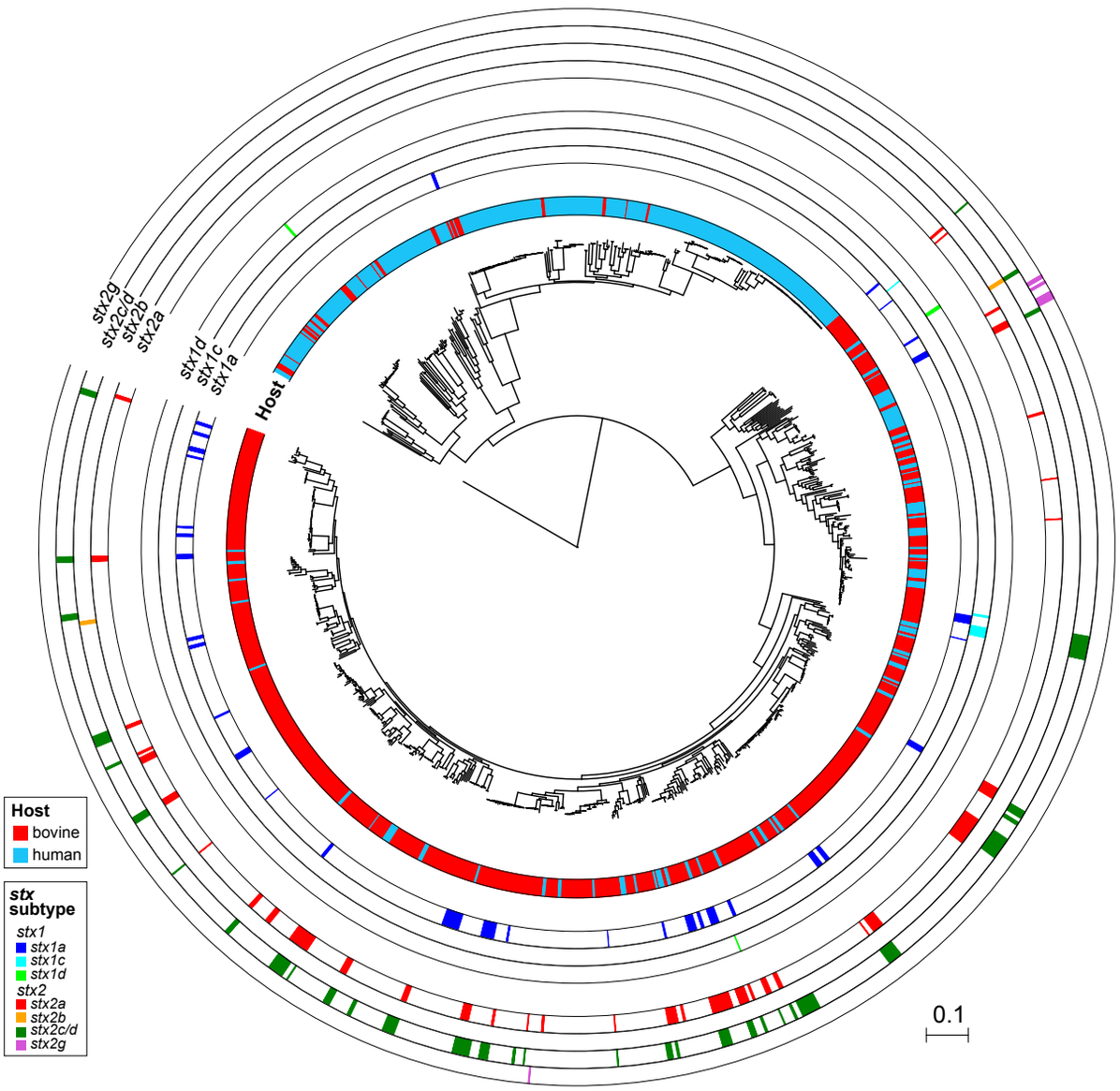


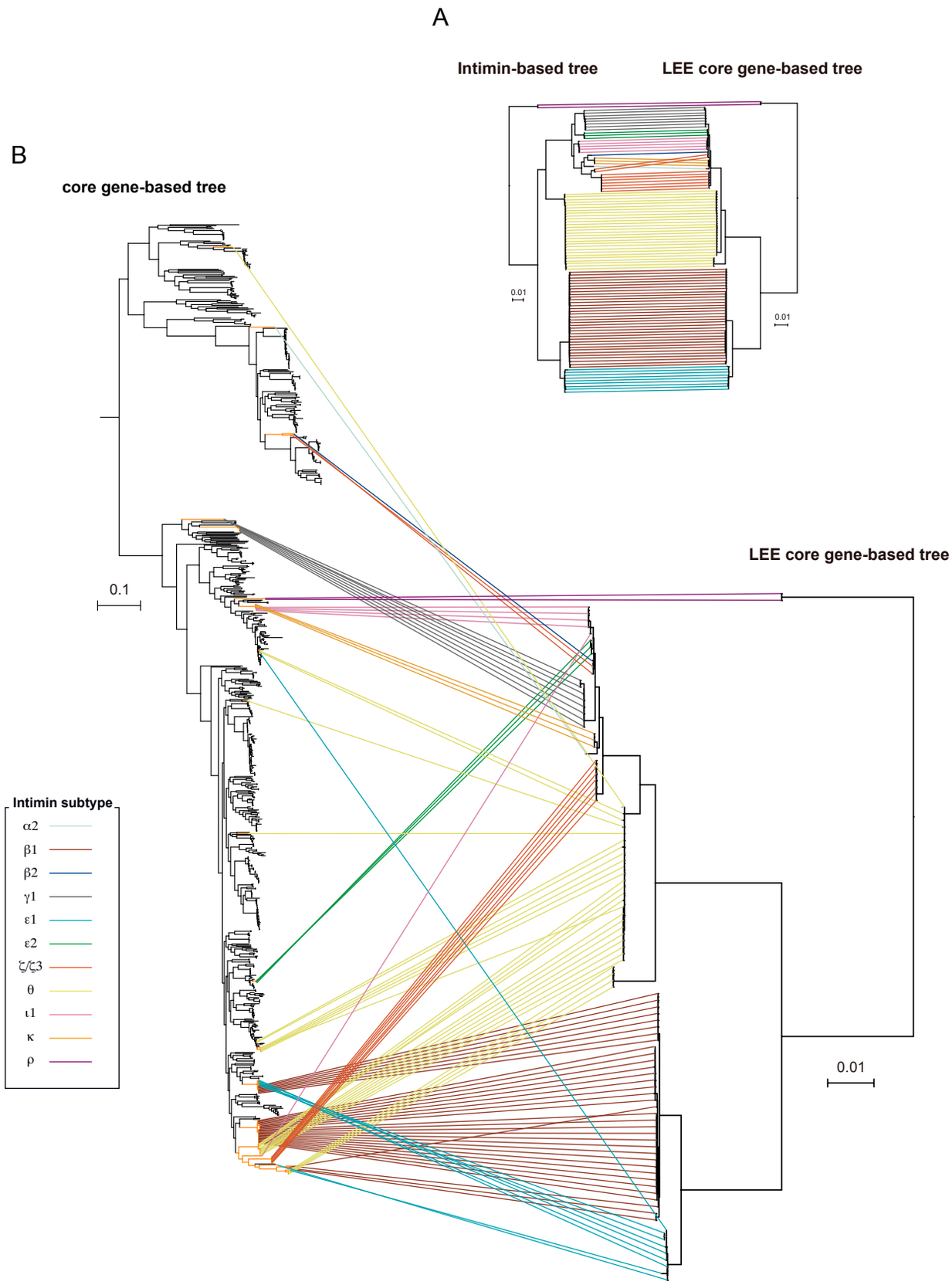
Supplemental figure S1. Geographic distribution of the strains analyzed in this study. Proportions of countries where the strains were isolated are shown for three groups of strains, respectively. Clinical isolates included STEC, EPEC and ExPEC strains which had clear indications that caused diseases in humans.



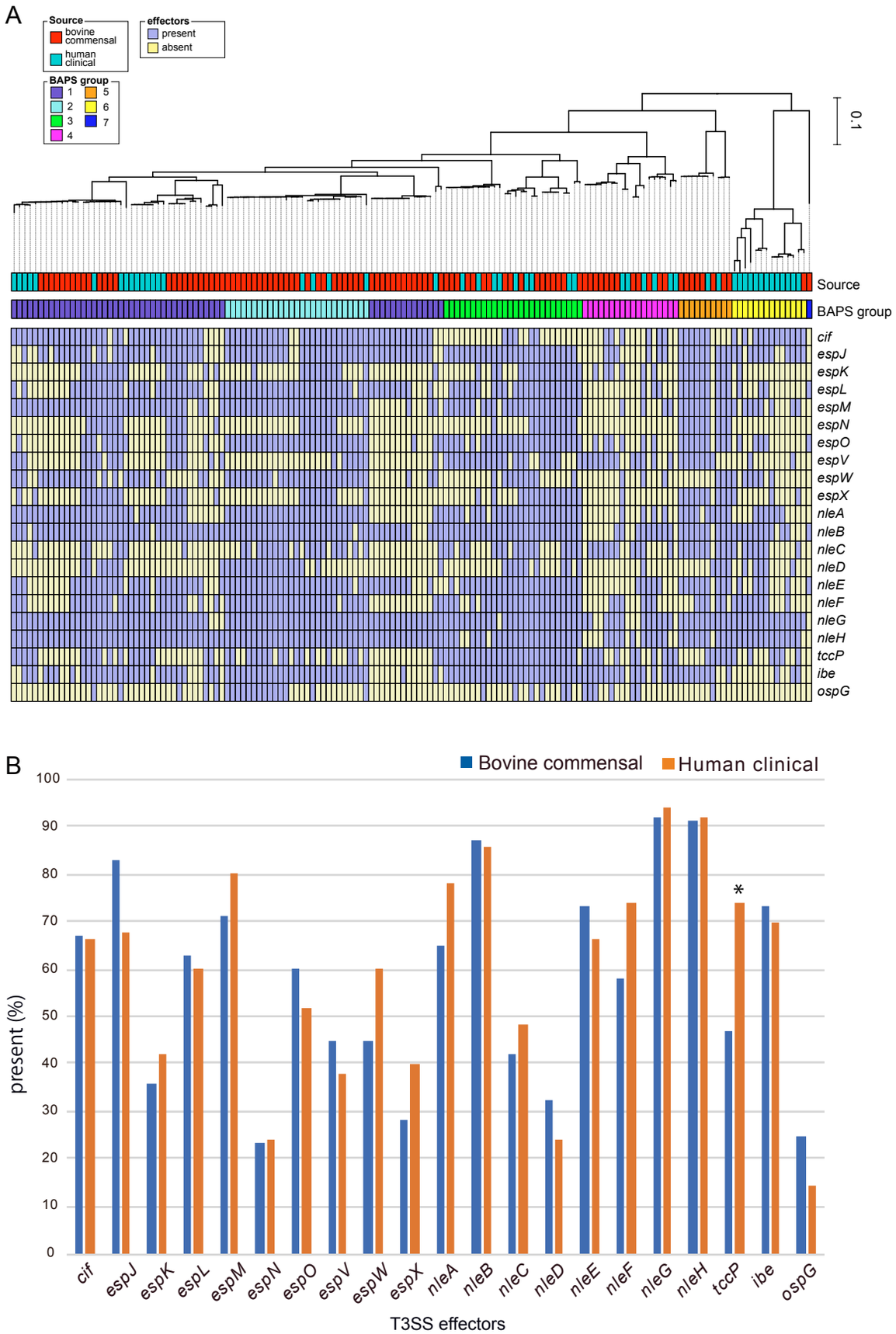
Supplemental figure S2. Serogroup distribution in bovine and human commensal isolates and human clinical isolates. Distribution of O and H serogroups across the core gene-based ML tree (the same tree shown in Fig. 1 in the main text) of 937 bovine and human commensal isolates and 197 human clinical isolates. The number of strains for each serogroup is shown in parentheses. UT: untypeable



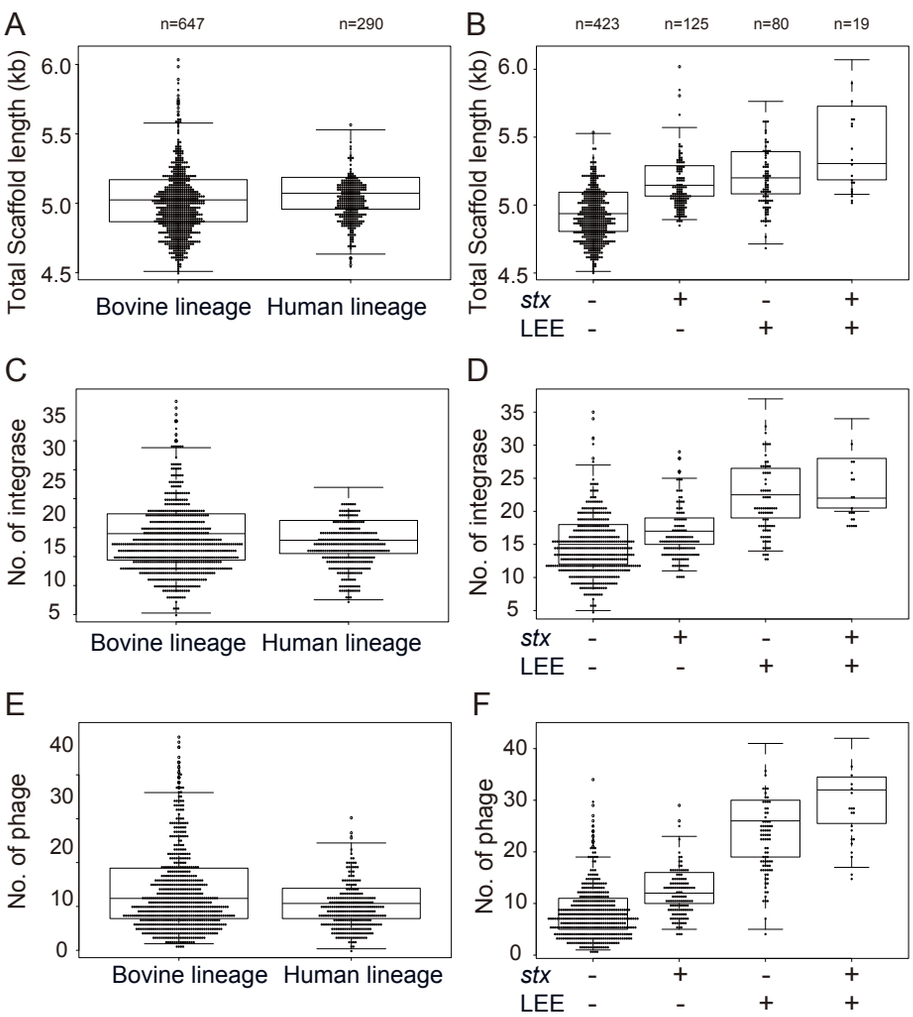
Supplemental figure S3. Distribution of *stx1* and *stx2* subtypes among bovine and human commensal *E. coli* isolates. Distribution of *stx1* and *stx2* subtypes across the core gene-based ML tree (the same tree shown in Fig. 2 in the main text) of 937 bovine and human commensal isolates. Each subtype is indicated by different colors.



Supplemental figure S4. Phylogenetic analysis of the LEE elements found in commensal strains and their distribution in the whole genome core gene-based tree. (*A*) Comparison of the intimin-based NJ tree and the six LEE core gene-based NJ tree. Only N-terminal conserved regions of the intimin genes (position 1 to 1,971) was used for the intimin-based tree construction. The six LEE core genes included *escS*, *escC*, *escJ*, *escV*, *escN*, and *cesD2*. Intimin subtypes are indicated by lines with different colors. (*B*) Comparison of the whole genome core gene-based ML tree of 937 bovine and human commensal isolates (the same tree shown in Fig. 2 in the main text) and the LEE core gene-based NJ tree of the LEE elements found in 104 commensal isolates.



Supplemental figure S5. Comparison of LEE-positive human clinical isolates and LEE-positive bovine commensal isolates. (*A*) Distribution of non-LEE effectors among LEE-positive human clinical isolates (n=50) and LEE-positive bovine commensal isolates (n=100) in the core gene-based ML tree constructed based on 160,444 SNPs sites on the 2,206 core genes of the 150 LEE-positive isolates. The presence and absence of each non-LEE effector family are indicated by light purple and beige, respectively. (*B*) Conservation of each effector in human clinical isolates and bovine commensal isolates. Statistical significance was analyzed by the Fisher's exact test with the Bonferroni correction for multiple comparisons (*P<0.05).



Supplemental figure S6. Comparison of the total scaffold lengths and the numbers of integrase and prophage between commensal isolates with respects to their phylogeny and possession of *stx* and LEE. Boxplots of total scaffold length (*A, B*) and numbers of integrase (*C, D*) and phage (*E, F*) between bovine and human lineages (*A, C, E*) and among bovine lineages with respect to the presence of *stx* and/or LEE (*B, D, F*). Solid horizontal lines and box indicate the median and the 25%-75% quartile range, respectively. Stems represent the minimum and maximum values, and dots represent the value of each strain. Statistical analysis is shown in supplemental Table S4.