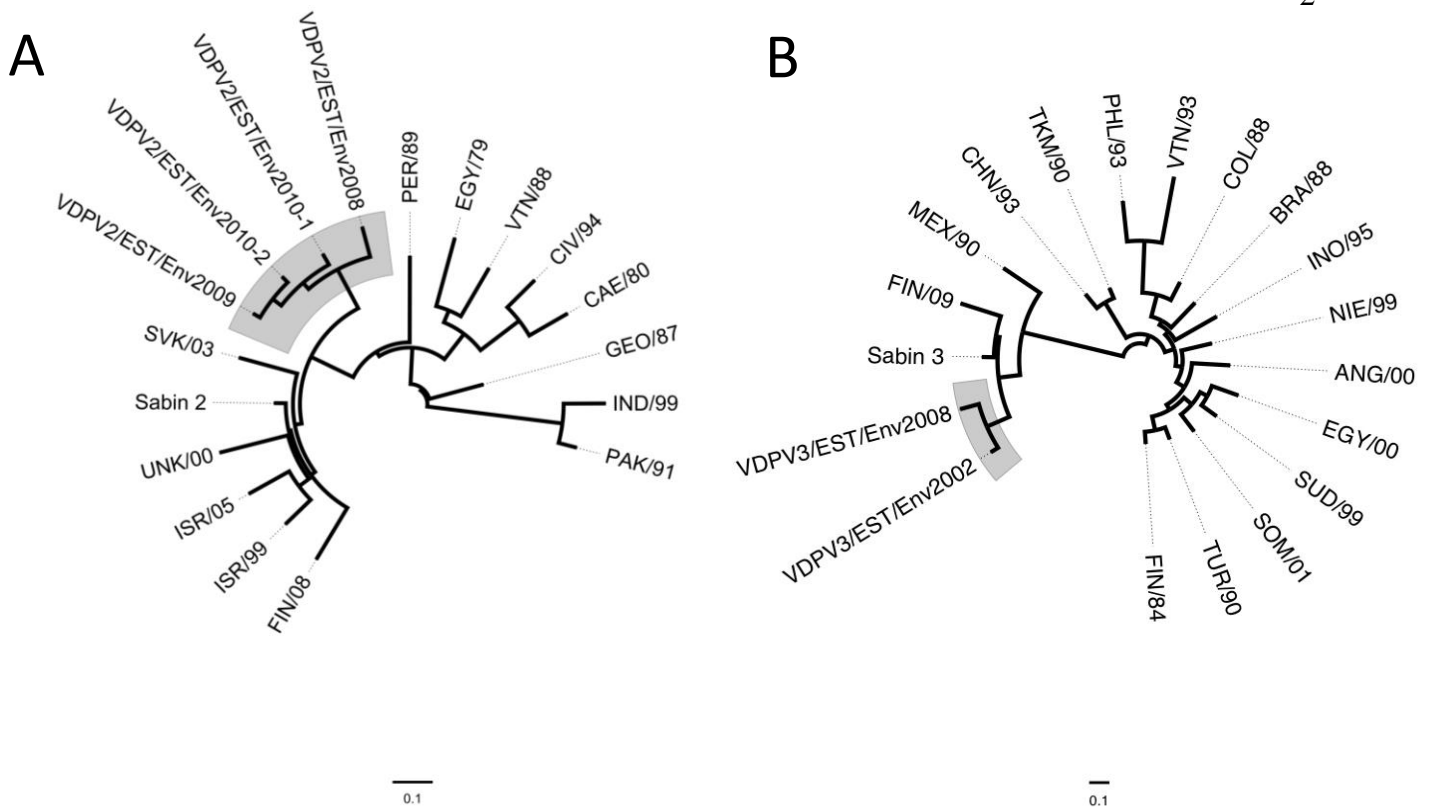# Supplemental Material

FIG S1 Maximum-likelihood trees depicting the sequence relationships in the VP1 region of the Estonian vaccine-derived poliovirus (VDPV) isolates of serotypes 2 (VDPV2) (A) and 3 (VDPV3) (B) to other environmental VDPV isolates and wild poliovirus (WPV) isolates representing different genotypes found in different parts of the world, 1979–2001. Trees with highest likelihood were inferred after a heuristic search under a general time-reversible substitution rate matrix, including the proportion of invariable sites and gamma rate heterogeneity, as determined by MODELTEST (1). Phylogenetic analyses were performed using the PAUP (2) plugin in the bioinformatics package Geneious v5.6 (http://www.geneious.com). Country abbreviations: ANG, Angola; BRA, Brazil; CAE, Cameroon; CHN, China; CIV, Côte d'Ivoire; COL, Colombia; EGY, Egypt; EST, Estonia; FIN, Finland; GEO, Georgia; IND, India; INO, Indonesia; ISR, Israel; MEX, Mexico; NIE, Nigeria; PAK, Pakistan; PER, Peru; PHL, Philippines; SOM, Somalia; SUD, Sudan; SVK, Slovakia; TKM, Turkmenistan; TUR, Turkey; UNK, United Kingdom; VTN, Vietnam. VP1 sequences of diverse VDPV2, VDPV3, WPV2, and WPV3 isolates were obtained from GenBank (**PV2:** CAE/80, AF551799; CIV/94, HQ286320; EGY/79, AF551798; FIN/08, KC784371; GEO/87, AF551802; IND/99, HQ286321; ISR/99, AM040036; ISR/05, AM158276; PAK/91, AF551805; PER/89, AF551803; SVK/03, JX913645; UNK/00, AJ544513; VTN/88, AF551804. **PV3:** ANG/00, AY221237; BRA/88, HQ286294; CHN/93, HQ286295; COL/88; HQ286296; EGY/00, GU562888; FIN/84, X04468; FIN/09, KC784373; INO/95, HQ286297; MEX/90, HQ286298; NIE/99, AY221233; PHL/93, HQ286302; SOM/01, AY221241; SUD/99, AY221243; TKM/90, KC769998; TUR/90, FJ839689; VTN/93, HQ286307).
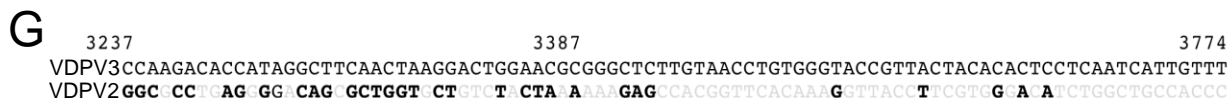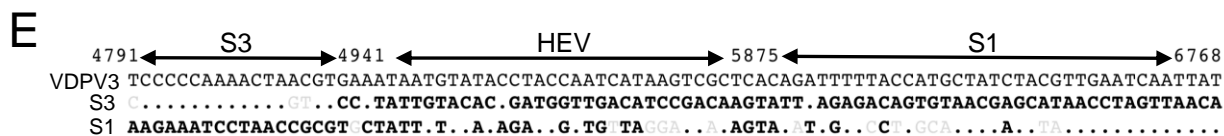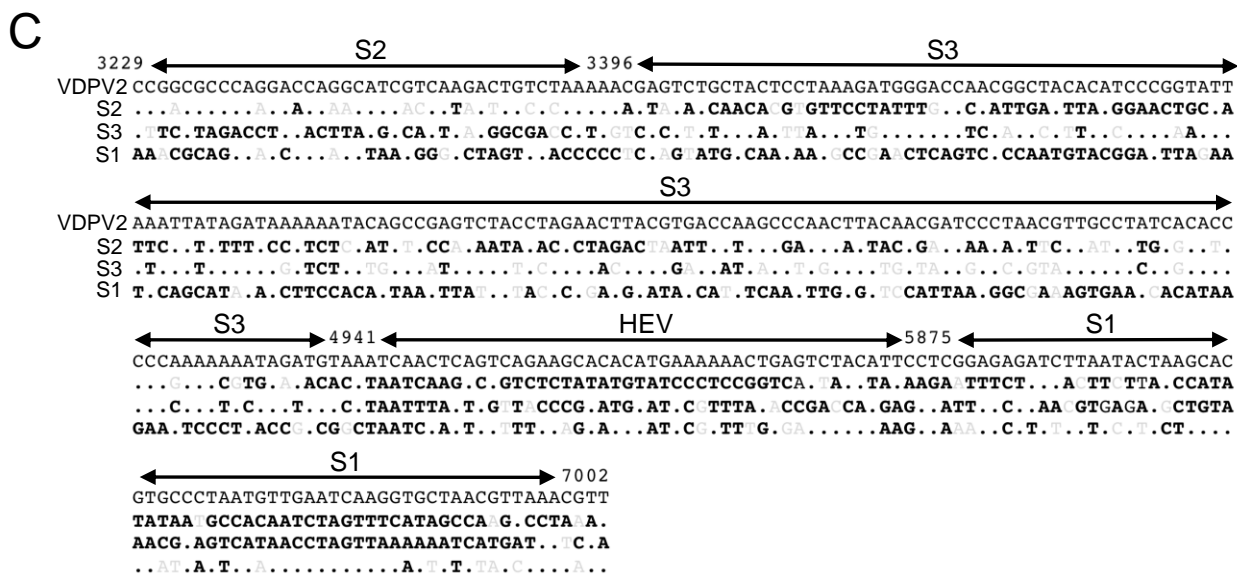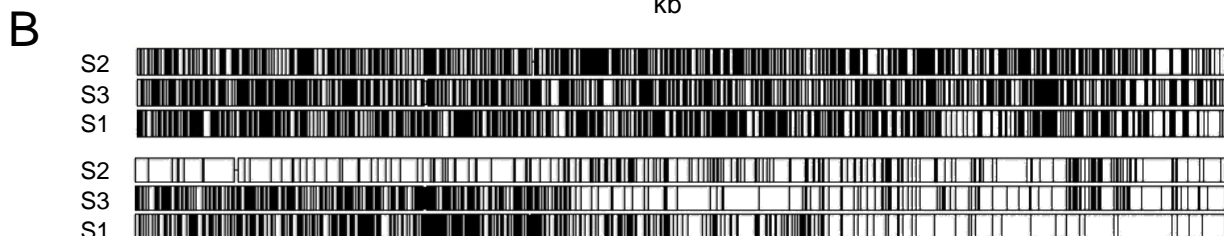
3

A

5´-UTR | P1/CAPSID | P2/NONCAPSID | P3/NONCAPSID | 3´-UTR

VP4 | VP2 | VP3 | VP1 | 2A | 2B | 2C | 3A | 3C | 3D — AAAn

0 1 2 3 4 5 6 7

kb

B

S2
S3
S1

S2
S3
S1

C

```
          3229                              S2                    3396                          S3
VDPV2 CCGGCGCCCAGGACCAGGCATCGTCAAGACTGTCTAAAAACGAGTCTGCTACTCCTAAAGATGGGACCAACGGCTACACATCCCGGTATT
S2    ...A.......A..A..AA....AC..TA.T..C.....A.T.A.CAACA.G.GTTCCTATTT.G..C.ATTGA.TTA.GGAACTGC.A
S3    .TTC.TAGACCT..ACTTA.G.CA.T.A.GGCGA.C.T.GT.C.C.T..T...A.TTA...T.G.......TC.A..C.TT..C.....A...
S1    AA.CGCAG...A.C...A..TAA.GG.G.CTAGT..ACCCCC.TC..G.GT.ATG.CAA.AA..GCC.A.CTCAGTC.CCAATGTACGGA.TTA.GAA

                                            S3
VDPV2 AAATTATAGATAAAAAATACAGCCGAGTCTACCTAGAACTTACGTGACCAAGCCCAACTTACAACGATCCCTAACGTTGCCTATCACACC
S2    TTC..T.TTTT.CC.TCT.AT.T.CCA.AATA.AC.CTAGACT.ATT..T...GA...A.TAC.G..AA.A.T.C...AT..TG.G..T.
S3    .T...T......G.TCT....AT.....T.C....AC..G.AT.A..T.G....TG.TA..G..C.GTA.....C..C...
S1    T.CAGCAT..A.CTTCCACA.TAA.TTA.T..TAC.C.GA...AT.CG.TT.G..TC.CATTAA.GGC.GA.AGTGAA..C.ACATAA

          S3      4941                    HEV                    5875          S1
VDPV  CCCAAAAAAATAGATGTAAATCAACTCAGTCAGAAGCACACATGAAAAAACTGAGTCTACATTCCTCGGAGAGATCTTAATACTAAGCAC
      ...G....C.GTG..A.ACAC.TAATCAAG..C.GTCTCTATATGTATCCCTCCGGTCA..TA..TA.AAGA.TTTCT...A.TT.CTTA.CCATA
      ...C...T.C...T...C.TAATTTA.T.G.T.ACCCG.ATG.AT.C.GTTTA..CCGA.CA.GAG..ATT..C..AA.GTGAGA..CTGTA
      GAA.TCCCT.ACC..CG.CTAATC.A.T..TTT..A.G.A...AT.CG.TT.TG.GA......AAG..AAA..C.T.T..T.C.T.CT...

          S1                       7002
VDPV  GTGCCCTAATGTTGAATCAAGGTGCTAACGTTAAACGTT
      TATAA.GCCACAATCTAGTTTCATAGCCA.AG.CCTA.A.
      AACG.AGTCATAACCTAGTTAAAAAAATCATGAT..TC.A
      ...AT.A.T..A...........A.T.TA.C....A..
```

D

S3
S1

S3
S1

E

```
       4791       S3      4941           HEV              5875        S1              6768
VDPV3 TCCCCCAAAACTAACGTGAAATAATGTATACCTACCAATCATAAGTCGCTCACAGATTTTTACCATGCTATCTACGTTGAATCAATTAT
S3    C...........GT...CC.TATTGTACAC.GATGGTTGACATCCGACAAGTATT.AGAGACAGTGTAACGAGCATAACCTAGTTAACA
S1    AAGAAATCCTAACCGCGT.CTATT.T..A.AGA..G.TG.TA.GGA...A.AGTA.AT.G..CCT.GCA....A..TA............
```

F

VDPV3/2
VDPV3/2

G

```
      3237                          3387                                          3774
VDPV3 CCAAGACACCATAGGCTTCAACTAAGGACTGGAACGCGGGCTCTTGTAACCTGTGGGTACCGTTACTACACACTCCTCAATCATTGTTT
VDPV2 GGC.CCC..AG.GG.CAG.GCTGGT.CT.GTCT.CTA.A.AAA.GAG.GCCACGGTTCACAAA.G.GTTACC.T.CGTG.GACA.TCTGGCTGCCACCC
```
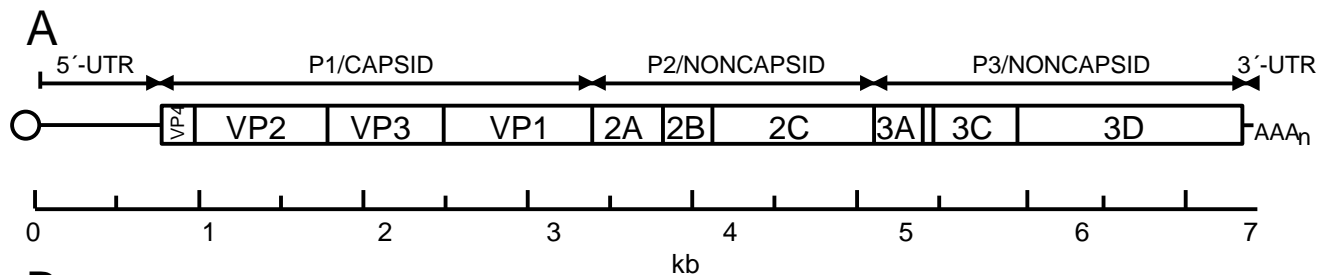
FIG S2 Approximate location of recombination sites in the Estonian VDPV2 and VDPV3 isolates. Panel A: In the schematic of the poliovirus genome, the single open reading frame is represented by an open rectangle flanked by the 5′- and 3′- untranslated regions (UTRs), represented by lines. Panel B: Alignment of the VDPV2/EST/Env2008 sequence (reference) with blocks marking sequence differences from Sabin 2, Sabin 3, and Sabin 1 (S1, S2, and S3) (upper trio: total nucleotide differences; lower trio: transversion differences). Panel C: Alignment of VDPV2/EST/Env2008 sequence (top, reference) showing nucleotide sequence differences from Sabin 2, Sabin 3, and Sabin 1 at sites of transversions (boldface type) relative to at least one of the Sabin strains. Transitions are shown in grey type. Panel D: Alignment of the VDPV3/EST/Env2002 sequence (top, reference) with blocks as described in panel B marking sequence differences from Sabin 3 and Sabin 1. Panel E: Alignment of the VDPV3/EST/Env2002 sequences (reference) showing nucleotide sequence differences from Sabin 3 and Sabin 1 at sites of transversions (boldface type). Symbols are as described in panel C. Panel F: Alignment of pairwise differences between VDPV3/EST/Env2002 (reference) and VDPV2/EST/Env2008 in total nucleotides (top) and transversions (bottom). Panel G: Pairwise transversion differences between VDPV3/EST/Env2002 (reference) and VDPV2/EST/Env2008 near the S2/S3 recombination site. Nucleotide positions are numbered according to Toyoda et al. (3) with the polymorphic sites distributed non-uniformly along the genome.
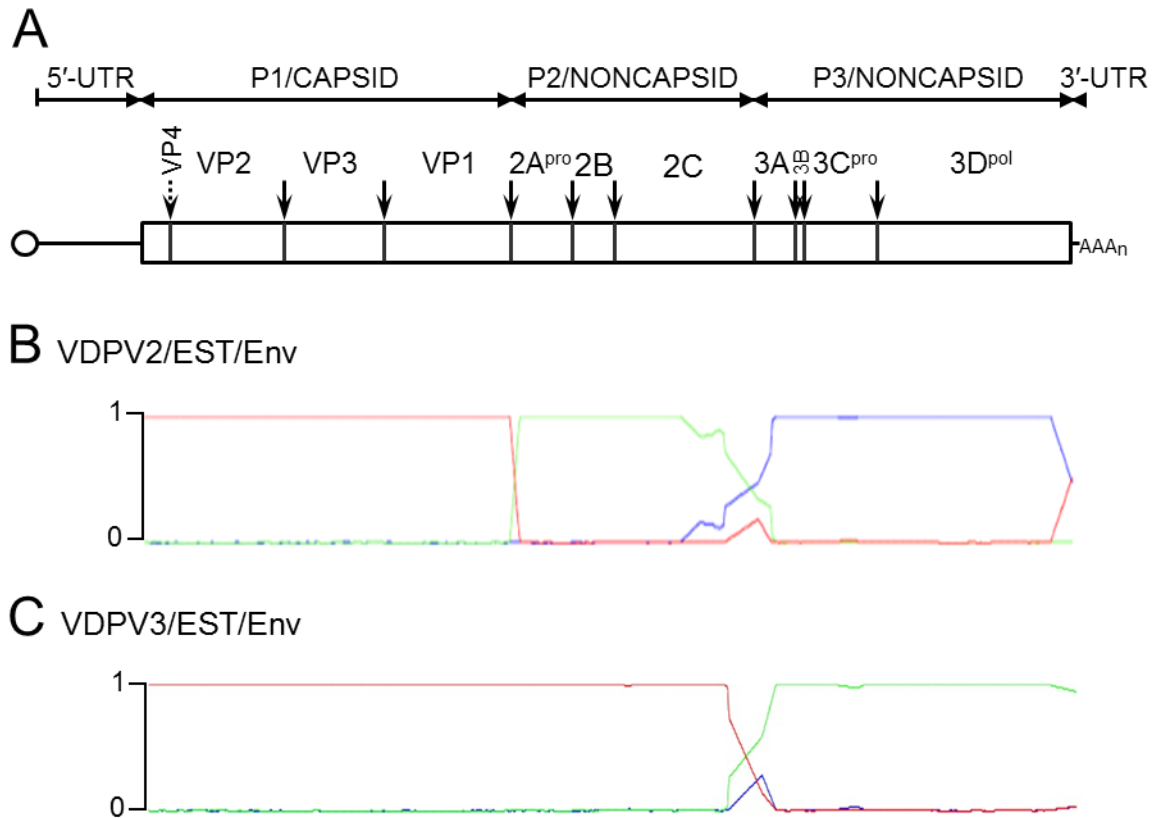
FIG S3 Schematic of the poliovirus genome (A) aligned with DualBrothers (4, 5) recombination detection profiles for VDPV2/EST/2008 (B) and VDPV3/EST/2002 (C). In each profile, as the scan moves from 5′ to 3′, color coding of the donor sequence in the recombinant blocks shifts from red to green to blue. Therefore, a particular donor sequence may be assigned a different color code in different plots. Profile B (5′ to 3′): Sabin 2 (red)— Sabin 3 (green)—non-Sabin—Sabin 1 (blue). Profile C (5′ to 3′): Sabin 3 (red)—non-Sabin— Sabin 1 (green). Plots are scaled to the posterior probability of the recombinant topology. The profiles scored only for transversional differences across donor sequences, which greatly reduced noise in the genetic signal.

	NAgS1												NAgS2												NAgS3						NAgS4	
	VP1												VP2							VP1					VP3					VP2	VP3
	91	92	93	94	95	96	97	98	99	100	101	102	164	165	166	168	169	170	270	221	222	223	224	226	58	59	60	71	73	72	76
Sabin 2	E	V	D	N	D	A	P	T	K	R	A	S	T	N	A	N	P	A	R	A	S	T	E	D	T	S	Q	E	S	R	A
VDPV2/EST/Env2008	-	-	-	-	-	-	-	-	Q	-	-	-	N	-	-	-	-	T	-	S	T	S	-	-	-	-	-	-	-	S	-
VDPV2/EST/Env2009	-	-	-	-	-	-	-	-	Q	-	T	-	*	D	-	S	-	-	-	S	T	S	-	-	-	-	-	-	-	N	-
VDPV2/EST/Env2010-1	-	-	-	-	-	-	-	-	Q	-	-	-	N	-	-	S	-	-	-	S	T	S	-	-	-	-	-	-	-	N	-
VDPV2/EST/Env2010-2	-	-	-	-	-	-	-	-	Q	S	-	-	N	-	-	S	-	-	-	S	T	S	-	-	-	-	-	-	-	N	-

	NAgS1												NAgS2										NAgS3											NAgS4			
	VP1												VP2				VP1					VP3					VP1					VP2	VP3		
	89	90	91	92	93	94	95	96	97	98	99	100	164	166	167	172	221	222	223	224	226	58	59	60	71	73	286	287	288	289	290	72	76	77	79
Sabin 3	E	V	D	N	E	Q	P	T	T	R	A	Q	N	V	T	E	D	Q	I	G	S	E	S	T	T	S	R	N	N	L	D	G	A	D	S
VDPV3/EST/Env2002	-	-	-	-	-	-	-	-	A	-	T	-	T	T	-	K	-	-	V	-	-	-	R	-	-	-	K	-	-	-	E	-	-	-	-
VDPV2/EST/Env2008	-	-	-	-	-	-	-	-	A	-	-	-	T	T	-	-	-	-	V	-	-	-	G	-	-	-	K	-	-	-	E	-	-	-	-

FIG S4 Substitutions in neutralizing antigenic (NAg) sites in selected Estonian VDPV strains as compared to the parental Sabin 2 and 3 strains. *, amino acid deletion.

References

1.   **Posada D, Crandall KA.** 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics **14:**817-818.

2.   **Swofford DL.** 2002. PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4. Sinaeur Associates, Sunderland, Massachusetts.

3.   **Toyoda H, M. Kohara M, Kataoka Y, Suganuma T, Omata T, Imura N, Nomoto A.** 1984. Complete nucleotide sequences of all three poliovirus serotype genomes: Implication for genetic relationship, gene function and antigenic determinants. J. Mol. Biol. **174:**561-585.

4.   **Suchard MA, Weiss RE, Dorman KS, Sinsheimer JS.** 2003. Inferring spatial phylogenetic variation along nucleotide sequences: a multiple changepoint model. J. Amer. Stat. Assoc. **98:**427-437.

5.   **Minin VN, Dorman KS, Fang F, Suchard MA.** 2005. Dual multiple change-point model leads to more accurate recombination detection. Bioinformatics **21:**3034-3042.