Figure legend - Supplementary Figure 1

Supplementary Figure 1. Alignments of 10 *pfs230* and 12 *pfg377* protein sequences obtained by PCR amplification of the respective genes from single-clone infected PNG field isolates. Sequencing by direct Sanger sequencing in one direction with nested PCR reverse primer of *pfs230* and in both directions with nested PCR forward and reverse primers of *pfg377. pfs230* nested PCR forward primer was found to be not suitable for sequencing. **A.** Protein sequence alignments of *pfs230* (left panel) and *pfg377* (right panel). Combination of two polymorphic regions resulted in increased marker diversity. 3D7 strain protein sequences derive from PlasmoDB: *pfs230,* PF3D7_0209000 and *pfg377,* PF3D7_1250100. **B.** Comparison of fragment sizes measured by capillary electrophoresis with sizes obtained by Sanger sequencing of the same isolates. Fragment sizes were underestimated by CE when compared to Sanger sequencing, which would be explained by the fact that the sized fragments were larger than the largest fragment (500 bp) of the GS500LIZ size standard (Applied Biosystems). The underestimation was less pronounced for *Pfs230* due to smaller PCR fragments.

A

pfs230 – protein alignment

```
3D7    1 RDNFVIDDEEEEEEEEEEEEE-EEEEEEEEEEEEEYDDYVYEESGDETEEQLQEEHQEEVG
S36    1 RDNFVIDDEEEEEEEEEEEEEEEEEEEEEEEEEYDDYVYEESGDETEEQLQEEHQEEVG
S47    1 RDNFVIDDEEEEEEEEEEEEE-EEEEEEEEEEEEEYDDYVYEESGDETEEQLQEEHQEEVG
S32    1 RDNFVIDDEEEEEEEEEEEEEEEEEEEEEEEEEYDDYVYEESGDETEEQLQEEHQEEVG
S39    1 RDNFVIDDEEEEEEEEEEEE---EEEEEEEEEEEYDDYVYEESGDETEEQLQEEHQEEVG
S33    1 RDNFVIDDEEEEEEEEEEEEE-EEEEEEEEEEEEEYDDYVYEESGDETEEQLQEEHQEEVG
S41    1 RDNFVIDDEEEEEEEEEEEEEEEEEEEEEEEEEYDDYVYEESGDETEEQLQEEHQEEVG
S42    1 RDNFVIDDEEEEEEEEEEEEE-EEEEEEEEEEEEEYDDYVYEESGDETEEQLQEEHQEEVG
S31    1 RDNFVIDDEEEEEEEEEEEEEEEEEEEEEEEEEYDDYVYEESGDETEEQLQEEHQEEVG
S44    1 RDNFVIDDEEEEEEEEEEEEE-EEEEEEEEEEEEEYDDYVYEESGDETEEQLQEEHQEEVG
S45    1 RDNFVIDDEEEEEEEEEEEEEEEEEEEEEEEEEYDDYVYEESGDETEEQLQEEHQEEVG

3D7   60 AESSEESFNDEDEDSVEARDGDMIRVDEYYEDQDGDTYDSTIKNEDVDE----EVGEEVG
S36   61 AESSEESFNDEDEDSVEARDGDMIRVDEYYEDQDGDTYDSTIKNEDVDE----EVGEEVG
S47   60 AESSEESFNDEDEDSVEARDGDMIRVDEYYEDQDGDTYDSTIKNEDVDEDADEEVGEEVG
S32   61 AESSEESFNDEDEDSVEARDGDMIRVDEYYEDQDGDTYDSTIKNEDADEDADEEVGEEVG
S39   58 AESSEESFNDEDEDSVEARDGDMIRVDEYYEDQDGDTYDSTIKNEDVDE----EVGEEVG
S33   60 AESSEESFNDEDEDSVEARDGDMIRVDEYYEDQDGDTYDSTIKNEDVDE----EVGEEVG
S41   61 AESSEESFNDEDEDSVEARDGDMIRVDEYYEDQDGDTYDS**SL**KNEDVDEHADAEVGEEVG
S42   60 AESSEESFNDEDEDSVEARDGDMIRVDEYYEDQDGDTYDSTIKNEDVDEDADEEVGED**A**D
S31   61 AESSEESFNDEDEDSVEARDGDMIRVDEYYEDQDGDTYDSTIKNEDVDEDADEEVGEEVG
S44   60 AESSEESFNDEDEDSVEARDGDMIRVDEYYEDQDGDTYDSTIKNEDVDEDADEEVGEE**L**G
S45   61 AESSEESFNDEDEDSVEARDGDMIRVDEYYEDQDGDTYDSTIKNEDVDEDADEEVGEEVG

3D7  116 EEVGEEVGE----EVGEEVGEEVGEEVGEEEGEEVGEGVGEEVGEEEGEEVGEEEG
S36  117 EEVGEEVGE----------------EVGEEVGEEEGEEVGEGVGEEVREEEGEEVGEEEG
S47  120 EEVGEEVGE----------EVGEEVGEEVGEEEGEEVGEGVGEEVGEEEGEEVGEEEG
S32  121 EEVGEEVGE----------EVGEEVDEEVGEE**V**GDE**E**GEGVGEELGEEEGEEVGEEEG
S39  114 EEVGEEVGE----EVGEEVGEEVDEEVGEE**V**GEE**E**GEE**V**GEEEGEEEGEEVGEEEG
S33  116 EEVGEEVGE----EVGEEVGEEVDEEVGEE**V**GEE**E**GEE**V**GEEEGEEEGEEVGEEEG
S41  121 EEVGAEVGE--------EVGEEVGEEVGEEVGEEEGEEVGEGVGEE**LRE**DEGEEVGEEEG
S42  120 EEVGEEVGE----EVGEEVGEEVGEEVGEEEGEEVGEGVGEEVGEEEGEEVGEEEG
S31  121 EEVGEEVGE----EVGEEVGEEVGEEVGEEEGEEVGEGVGEEVGEEEGEEVGEEEG
S44  120 EEVGEEVGEEVGEEVGEEVGEEVGEEVGEEEGE**D**VGEGVGEEVG**Q**EEGEEVGEEEG
S45  121 EEVGEEVGEEVGEEVGEEVGEEVGEEVGEEEGEEVGEGVGEEVGEEEGEEVGEEEG

3D7  172 EYVDEKERQG  282
S36  161 EYVDEKERQG  171
S47  168 EYVDEKERQG  178
S32  169 EYVDEKERQG  179
S39  170 EYVDEKERQG  180
S33  172 EYVDEKERQG  182
S41  173 EYVDEKERQG  183
S42  176 EYVDEKERQG  186
S31  177 EYVDEKERQG  187
S44  180 EYVDEKERQG  190
S45  181 EYVDEKERQG  191
```

pfg377 – protein alignment

```
3D7    1 EDDEGDEVNEENVPLNHEE----------------------------ENVPLNHEEDNFPL
G19    1 EDDEGDEVNEENVPLNHEE-------------------------------------**ENV**PL
G5     1 EDDEGDEVNEENVPLNHEE-----------------ENVPLNHEE**DNF**PLNHEEDNFPL
G7     1 EDDEGDEVNEENVPLNHEE-------------------------------------DNFPL
G1     1 EDDEGDEVNEENVPLNHEE----------------------------ENVPLNHEEDNFPL
G9     1 EDDEGDEVNEENVPLNHEE-----------------ENVPLNHEEENVPLNHEEDNFPL
G2     1 EDDEGDEVNEENVPLNHEE----------------------------ENVPLNHEEDNFPL
G11    1 EDDEGDEVNEENVPLNHEE-----------------ENVPLNHEE**DNF**PLNHEEDNFPL
G4     1 EDDEGDEVNEENVPLNHEE-----------------ENVPLNHEE**DNF**PLNHEEDNFPL
G14    1 EDDEGDEVNEENVPLNHEE---------ENVPLNHEEENVPLNHEEENVPLNHEE**ENV**PL
G15    1 EDDEGDEVNEENVPLNHEE---------ENVPLNHEEENVPLNHEEENVPLNHEE**ENV**PL
G16    1 EDDEGDEVNEENVPLNHEEENVPLNHEEENVPLNHEE**DNF**PLNHEE**DNF**PLNHEE**ENV**PL
G17    1 EDDEGDEVNEENVPLNHEEENVPLNHEEENVPLNHEEENVPLNHEEENVPLNHEE**ENV**PL

3D7   34 NHEEDNFPLNHEEENVPLNHEEENVPLNHEEENVPLNHEEENYNSFLSYQKYDERDISPH
G19   25 NHEE**ENV**PLNHEEENVPLNHEEENVPLNHEEENVPLNHEEENYNSFLSYQKYDERDISPH
G5    43 NHEEDNFPLNHEEENVPLNHEEENVPLNHEEENVPLNHEEENYNSFLSYQKYDERDISPH
G7    25 NHEEDNFPLNHEEENVPLNHEEENVPLNHEEENVPLNHEEENYNSFLSYQKYDERDISPH
G1    34 NHEEDNFPLNHEEENVPLNHEEENVPLNHEEENVPLNHEEENYNSFLSYQKYDERDISPH
G9    43 NHEEDNFPLNHEEENVPLNHEEENVPLNHEEENVPLNHEEENYNSFLSYQKYDERDISPH
G2    34 NHEEDNFPLNHEE**DNF**PLNHEEENVPLNHEEENVPLNHEEENYNSFLSYQKYDERDISPH
G11   43 NHEEDNFPLNHEEENVPLNHEEENVPLNHEEENVPLNHEEENYNSFLSYQKYDERDISPH
G4    43 NHEEDNFPLNHEEENVPLNHEEENVPLNHEEENVPLNHEEENYNSFLSYQKYDERDISPH
G14   52 NHEE**ENV**PLNHEEENVPLNHEEENVPLNHEEENVPLNHEEENYNSFLSYQKYDERDISPH
G15   52 NHEE**ENV**PLNHEEENVPLNHEEENVPLNHEEENVPLNHEEENYNSFLSYQKYDERDISPH
G16   61 NHEE**ENV**PLNHEEENVPLNHEEENVPLNHEEENVPLNHEEENYNSFLSYQKYDERDISPH
G17   61 NHEE**ENV**PLNHEEENVPLNHEEENVPLNHEEENVPLNHEEENYNSFLSYQKYDERDISPH

3D7   94 NVNQHKKEIFHHRNITPYHVNHHNQHMDQHDHHIDHHN-------HHIDHHNHHIDHHNH
G19   85 NVNQHKKEIFHHRNITPYHVNHHNQHMDQHDHHIDHHN----------------------
G5   103 NVNQHKKEIFHHRNITPYHVNHHNQHMDQHDHHIDHHN----------------------
G7    85 NVNQHKKEIFHHRNITPYHVNHHNQHMDQHDHHIDHHN-------------HHIDHHNH
G1    94 NVNQHKKEIFHHRNITPYHVNHHNQHMDQHDHHIDHHN--------------------H
G9   103 NVNQHKKEIFHHRNITPYHVNHHNQHMDQHDHHIDHHN----------------------
G2    94 NVNQHKKEIFHHRNITPYHVNHHNQHMDQHDHHIDHHN--------------HHIDHHNH
G11  103 NVNQHKKEIFHHRNITPYHVNHHNQHMDQHDHHIDHHN--------------------H
G4   103 NVNQHKKEIFHHRNITPYHVNHHNQHMDQHDHHIDHHNQHMDQHD--------------H
G14  112 NVNQHKKEIFHHRNITPYHVNHHNQHMDQHDHHIDHHN--------------HHIDHHNH
G15  112 NVNQHKKEIFHHRNITPYHVNHHNQHMDQHDHHIDHHN-------HHIDHHNHHIDHHNH
G16  121 NVNQHKKEIFHHRNITPYHVNHHNQHMDQHDHHIDHHN--------------HHIDHHNH
G17  121 NVNQHKKEIFHHRNITPYHVNHHNQHMDQHDHHIDHHN-------HHIDHHNHHIDHHNH

3D7  147 HIDHHNHHIDHQDHHIDHHNHHIDHHNHHIDHHDHHIDHHDHHIDHHNHHIDHKSNNQFLQ  207
G19  123 ------HHIDH**HN**HHIDHH**D**HHIDHHNHHIDHHDHHIDHHDHHIDHHNHHIDHKSNNQFLQ  176
G5   141 --------------------HHIDHHNHHIDHHDHHIDHHDHHIDHHNHHIDHKSNNQFLQ  181
G7   131 HIDHHNHHIDHQDHHIDHHNHHIDHHNHHIDHHDHHIDHHDHHIDHHNHHIDHKSNNQFLQ  190
G1   133 HIDHHNHHIDH**HN**HHIDH**QD**HHIDHHNHHIDHHDHHIDHHDHHIDHHNHHIDHKSNNQFLQ  193
G9   141 ------HHIDH**HN**HHIDH**QD**HHIDHHNHHIDHHDHHIDHHDHHIDHHNHHIDHKSNNQFLQ  194
G2   140 HIDHHNHHIDHQDHHIDHHNHHIDHHNHHIDHHDHHIDHHDHHIDHHNHHIDHKSNNQFLQ  199
G11  142 HIDHHNHHIDH**HN**HHIDH**QD**HHIDHHNHHIDHHDHHIDHHDHHIDHHNHHIDHKSNNQFLQ  201
G4   149 HIDHHNHHIDH**HN**HHIDH**QD**HHIDHHNHHIDH**H**HHIDHHDHHIDHHNHHIDHKSNNQFLQ  209
G14  158 HIDHHNHHIDH**HN**HHIDH**QD**HHIDHHNHHIDHHDHHIDHHDHHIDHHNHHIDHKSNNQFLQ  217
G15  165 HIDHHNHHIDHQDHHIDHHNHHIDHHNHHIDHHDHHIDHHDHHIDHHNHHIDHKSNNQFLQ  225
G16  167 HIDHHNHHIDHQDHHIDHHNHHIDHHNHHIDHHDHHIDHHDHHIDHHNHHIDHKSNNQFLQ  227
G17  174 HIDHHNHHIDHQDHHIDHHNHHIDHHNHHIDHHDHHIDHHDHHIDHHNHHIDHKSNNQFLQ  233
```

B

pfs230 – nucleotide sequence size

Sample	Size by CE (bp)*	Size by Sequencing (bp)
3D7	538.5	544
S36	510.1	511
S47	528.3	532
S32	531.0	535
S39	533.2	538
S33	538.0	544
S41	541.1	547
S42	547.1	556
S31	550.0	559
S44	558.8	568
S45	559.6	571

*Using Size Standard GS500LIZ (Applied Biosystem)

pfg377 – nucleotide sequence size

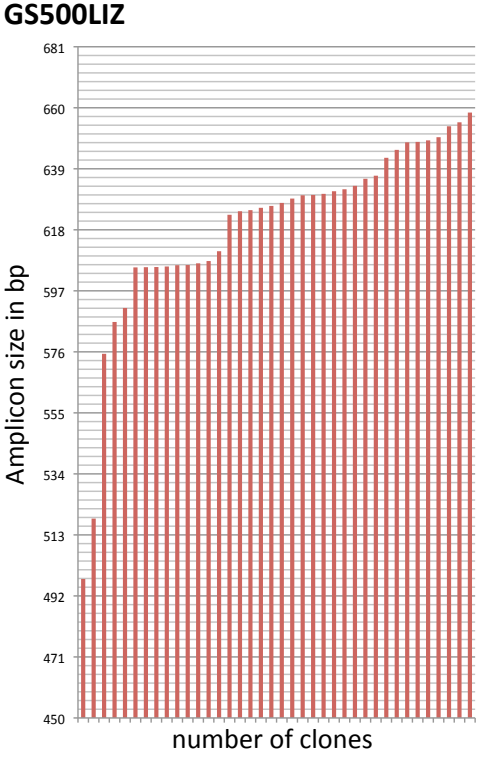
Sample	Size by CE (bp)*	Size by Sequencing (bp)
3D7	591.2	621
G19	523.3	531
G5	534.2	543
G7	557.5	573
G1	563.4	579
G9	566.6	585
G2	576.8	600
G11	582.9	606
G4	598.8	627
G14	614.6	654
G15	629.2	675
G16	631.0	681
G17	641.3	702

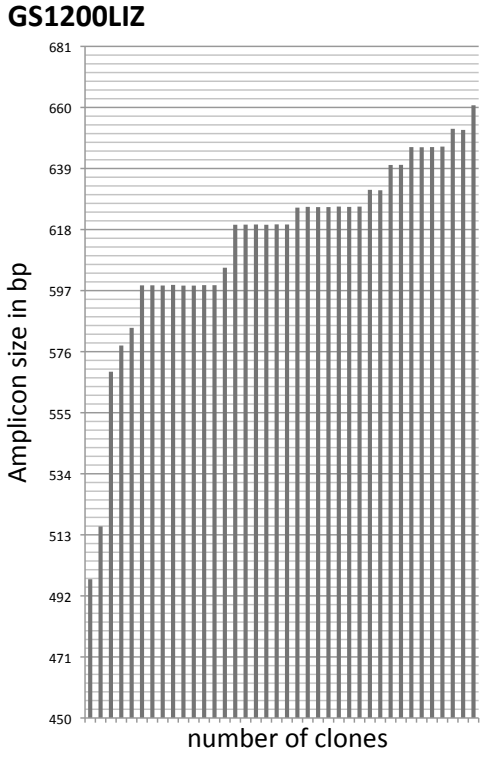
*Using Size Standard GS500LIZ (Applied Biosystem)

Figure Legend - Supplementary Figure 2

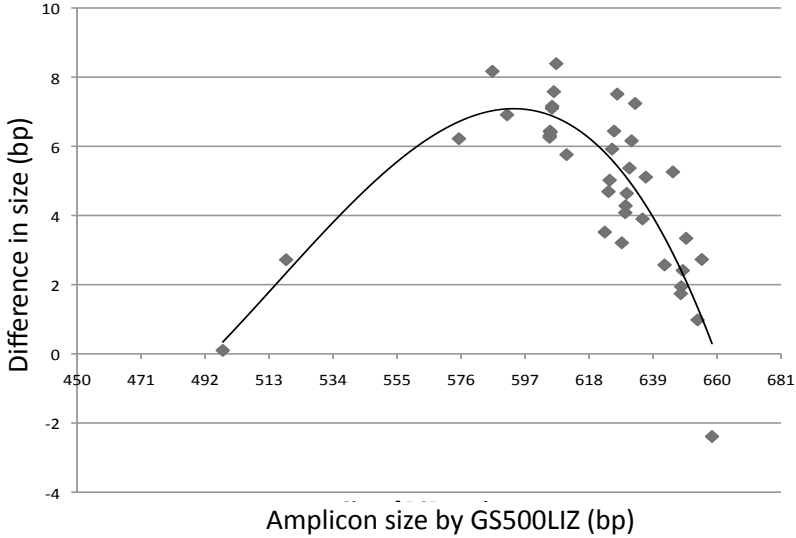
Supplementary Figure 2. Comparison of fragment sizes obtained by using two different capillary electrophoresis (CE) size standards applied to 38 *pfg377* amplicons detected in 13 blood samples. To investigate whether a size standard containing fragments up to 1200 bp would provide more accurate *pfg377* fragment sizing, a subset of 13 samples was simultaneously sized by CE for *pfg377* using either the GS500LIZ or the GS1200LIZ (Applied Biosystems). **A.** Distribution of amplicon sizes by using GS500LIZ (left panel) or GS1200LIZ (right panel). Better resolution, especially in the size range over 600 bp, was obtained for GS1200LIZ. **B.** Difference of sizes by GS500LIZ minus sizes by GS1200LIZ plotted over the amplicon size of GS500LIZ. Curve shows a polynomic equation: $y = -5E\text{-}06x^3 + 0.0083x^2 - 4.1946x + 689.48$. A non-linear overestimation of size by GS500LIZ was found. This leads to the conclusion that GS500LIZ size standard is sufficient for *pfs230,* as amplicons are <600 bp. For *pfg377* with amplicons >700 bp the GS1200 size standard provides improved sizing.

A


GS500LIZ
Amplicon size in bp
681
660
639
618
597
576
555
534
513
492
471
450
number of clones


GS1200LIZ
Amplicon size in bp
681
660
639
618
597
576
555
534
513
492
471
450
number of clones

B


Difference in size (bp)
10
8
6
4
2
0
-2
-4
450
471
492
513
534
555
576
597
618
639
660
681
Amplicon size by GS500LIZ (bp)