SUPPLEMENTAL DATA

QUANTIFICATION OF AMPLIFIABLE GENOMES IN SPERM SAMPLES BY KT-PCR

We quantified the approximate number of amplifiable genomes per mass unit DNA for each sperm sample with a ktPCR on an AB7700 using the Expand Long Template PCR System (Roche). Unmodified, non allele-specific primers were used for this experiment (listed below). To do this, a standard curve using human genomic DNA (BD Biosciences) was used as a basis for comparison, assuming that for high-quality DNA, 10ng = 3000 amplifiable genomes. We measured the amplification yields for the "unknown" sperm samples relative to the curve generated by yields of four standard DNA amounts in a 50uL reaction (120ng, 92ng, 40ng, 12ng). Replicates for each sample were averaged, and a "correction factor" was estimated for each sperm DNA. The amount of sperm DNA added to each AS-PCR was then adjusted (10ng x sample-specific "correction factor") to bring the number of estimated amplifiable genomes to 3000 for each reaction.

Reaction conditions:

Each 120uL mastermix for each sample contained the following: 220ng sperm DNA, 400nM each primer, 0.1x buffer 1 (Expand Long Template System), 0.1x DMSO (Sigma), 0.12x Betain (Sigma) 350uM each dNTP (Bioline), 1.8uL Expand enzyme mix, and 0.01x SybrGreen (Molecular Probes). Each of two 50uL replicates was then aliquoted to the PCR plate.

Primer sequences:

C10_RT_QUANT_F: CTAGGCTGGCCTCAGTGATCCTCCT

C10_RT_QUANT_R: CCTTCACCCTCAAGCAGAGCTGTG

Cycling conditions on AB7700:

1. 92 ºC 2:00
2. 92 ºC 0:10
3. 63.2 ºC 0:30
4. 68 ºC 4:00
5. 68 ºC 1:00 data collection
6. Go to 2 for 40 cycles
7. 68 ºC 7:00
8. 4 ºC forever

OPTIMIZATION OF THE AMPLIFICATION OF RECOMBINANT MOLECULES IN SPERM POOLS BY AS-PCR

We performed serial dilution experiments to determine the appropriate cycling conditions for the amplification of few recombinant molecules out of a pool containing predominantly non-recombinant molecules, and with no evidence of non-specific PCR in the negative controls. In each dilution series, we used an estimated 3000 amplifiable genomes, under the assumption that, uncorrected for variable number of amplifiable genomes, 10 ng of genomic DNA equal 3000 genomes. First, 5 serial dilutions were made in which a positive control, i.e. blood DNA from an individual carrying the "recombinant" C-C-A-C haplotype, was diluted in 1:1 mix of DNA from

two homozygotes for the "non-recombinant" haplotypes (C-C-C-G and the T-G-A-C). The most concentrated dilution contained equal numbers of "recombinant" and "non-recombinant" molecules; the least concentrated dilution contained 3 "recombinant" molecules for every 3000 "non-recombinant" molecules. A sample with no "recombinant" molecules was also included as a negative control for the specificity of AS-PCR. Finally, to determine if the optimized PCR conditions were sensitive enough to detect a single recombinant molecule in a pool of non-recombinants, we amplified 8 DNA pools containing 1 "recombinant" molecule in 3000 "non-recombinant" molecules and 8 negative controls with no "recombinant" molecules. Under the expectations of a Poisson distribution, 50-60% (4 or 5 out of 8) of the 1:3000 dilution reactions should be positive. We determined that appropriate specificity and sensitivity of PCR conditions was achieved when the 1:3000 dilution reactions yielded the Poisson expected fraction of positives and the no "recombinant" controls were all negative.

ALLELE-SPECIFIC CONDITIONS FOR NESTED PCR OF RECOMBINANTS

We used the Roche Expand Long Template PCR System for both the first external and the internal nested AS-PCR, with a total reaction volume of 15uL. The last four phosphodiester bonds on the 3' end of each AS primer were replaced with phosphorothioate (*) to increase specificity. Additionally, mismatches were placed (shown in bold) in the 3’ end of the internal primers to increase specificity of the nested PCR (see primer sequences below). The amount of sperm DNA (for both heterozygous test individuals and homozygous controls) added was adjusted (see above) to approximately 3000 genomes per each DNA pool. The DNA was added

to the master mix of the 1st PCR before dividing into individual aliquots to ensure a homogeneous DNA concentration in each reaction. An equal number of pools with DNA from two individuals, each homozygous for the non-recombinant haplotypes, were set up in the same manner as the test individuals as negative controls. The first PCR was performed on an MJ Tetrad Thermocycler (see cycling conditions). As template for the nested PCR, 1 uL of the undiluted 1st PCR product (for 2nd PCR on MJ Tetrad) or 1uL of a 1:40 dilution of the 1st PCR product (ktPCR on AB7700) was added to the wells before aliquoting the PCR mastermix. The second PCR was performed on either the MJ Tetrad or on the AB7700.

Reaction conditions for the 1st PCR:

Each 15uL reaction contained the following: sperm DNA (adjusted to 3000 amplifiable genomes), 400nM each primer, 0.1x buffer 1 (Expand Long Template System), 0.1x DMSO (Sigma), 0.12x Betain (Sigma), 350uM each dNTP (Bioline), and 0.225uL Expand enzyme mix.

Primer sequences for 1st PCR:

C10_ST_30952_C ctggcctcagtgatcctcctgcctt*g*g*c

C10_ST_35309_C CCCTCAAGCAGAGCTGTGCGAGGA*G*C*C*G

1st PCR Cycling Conditions (MJ Tetrad Thermocycler):

1. 92 ºC 2:00
2. 92 ºC 0:10
3. 69 ºC 0:30
4. 68 ºC 3:30
5. Go to 2 for 10 cycles

6. 92 ºC 0:10
7. 69 ºC 0:30
8. 68 ºC 3:30 +10sec/cycle
9. go to 6 for 20 cycles
10. 68 ºC 7:00
11. 4 ºC forever

Reaction conditions for the 2nd PCR:

Each 15uL reaction contained the following: 1uL of 1st PCR product (or 1uL of 1:40 diluted product for ktPCR) 400nM each primer, 0.1x buffer 1 (Expand Long Template System), 0.1x DMSO (Sigma), 0.12x Betain (Sigma) 350uM each dNTP (Bioline), 0.225uL Expand enzyme mix, and 0.01x SybrGreen (Molecular Probes).

Primers for 2nd PCR:

C10_ST_31449_3*T_C tgatttgggccttgct*g***t***c*c

C10_ST_35020_2*A_A CAGGTGTGAGCCGGGG*C*G***A***T

2nd PCR Cycling Conditions (MJ Tetrad for gel imaging or AB7700 for ktPCR):

1. 92 ºC 2:00
2. 92 ºC 0:10
3. 57.2 ºC 0:30
4. 68 ºC 3:30
5. Go to 2 for 10 cycles

6. 92 ºC 0:10

7. 57.2 ºC 0:30

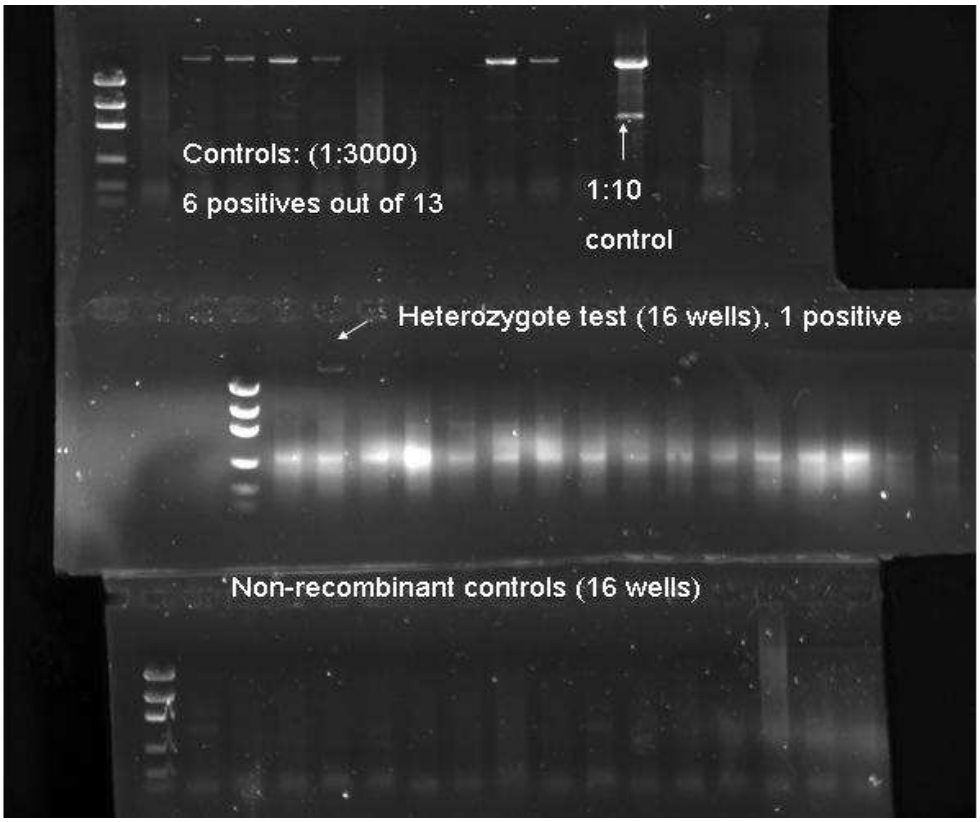
8. 68 ºC 3:30 +10sec/cycle

9. go to 6 for 22 cycles

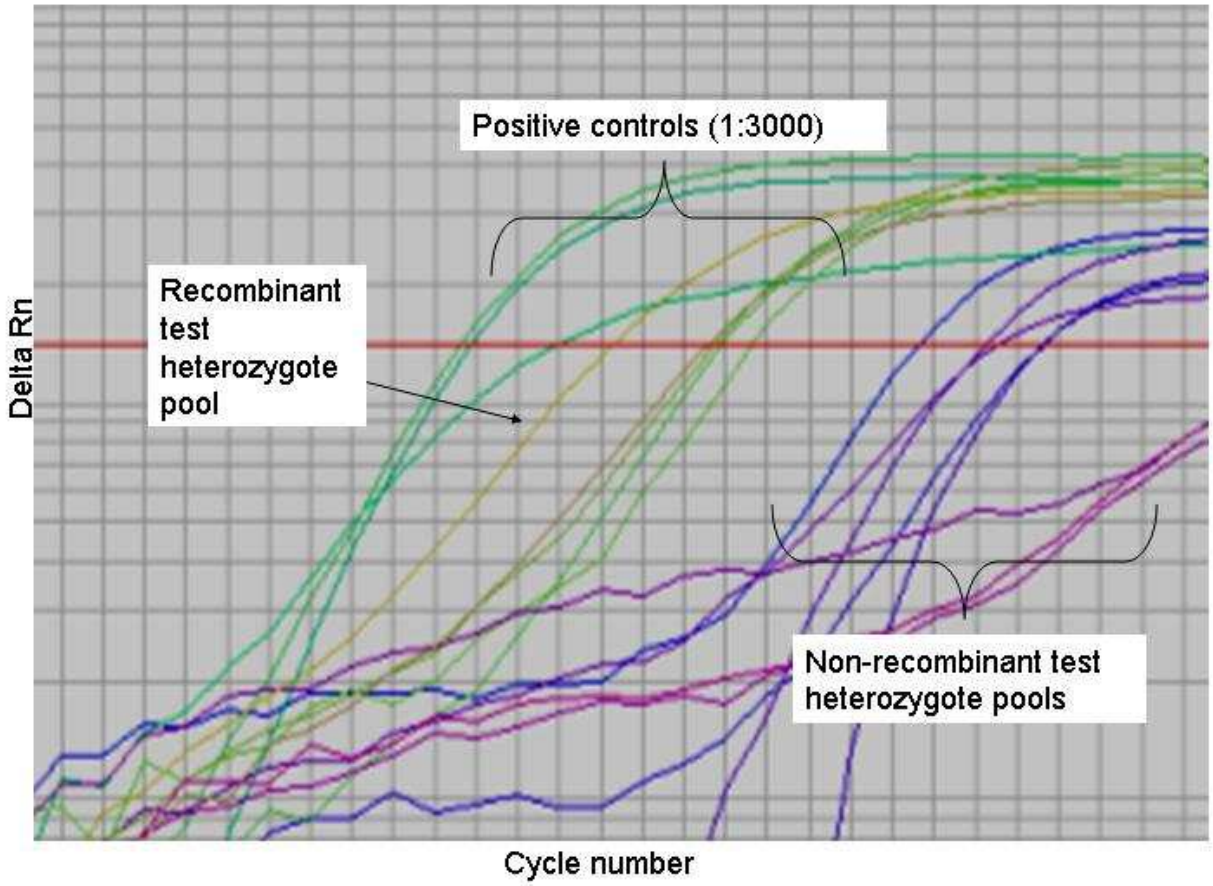
10. 68 ºC 7:00

11. 4 ºC forever

SUPPLEMENTAL FIGURE 1. Agarose gel image of 2nd PCR product. Experiment includes 13 (1:3000) positive control reactions and one 1:10 control (bands on top row), 16 heterozygote test pools (1/16 are positive) (middle row), and 16 non-recombinant negative control reactions (bottom row).



SUPPLEMENTAL FIGURE 2. Results of nested PCR (ktPCR) on AB7900. Early set of amplification plots (to the left) denote positive controls (1:3000 dilutions) and a heterozygote test reaction in which a recombinant molecule is detected. The later set of amplification plots (to the right) denote heterozygote test reactions in which no recombinants are detected.



SUPPLEMENTAL TABLE 1. The proportion of 1000 simulated data sets in which the estimate of ρ_{LD} was as high or higher than observed

	Western chimpanzee	Hausa	Italian	Chinese
Standard Neutral				
$f = 0$	0.530	0.000	0.068	0.261
$f = 2$, $L = 500$	0.592	0.000	0.138	0.364
$f = 5$, $L = 60$	0.575	0.000	0.115	0.315
Other demography				
$f = 0$	--	0.000	0.290	0.326
$f = 2$, $L = 500$	--	0.000	0.325	0.390
$f = 5$, $L = 60$	--	0.000	0.278	0.339