Bait generation

- Donor genomic DNA (high quality, e.g., blood-derived)
- Bioruptor shearing
  - ~150bp dsDNA
  - End repair and A-tailing.
  - Ligation of a custom adaptor containing a T7 RNA polymerase recognition site and EcoO109I restriction enzyme cut site.
  - Denaturation and treatment with duplex-specific nuclease to deplete repetitive elements
  - PCR amplification followed by in vitro RNA transcription with biotin incorporation
  - DNase and restriction enzyme digestion to remove donor DNA and T7 adaptors
  - Biotinylated RNA baits

Genome-wide capture

- DNA from a non-invasive sample (highly sheared, only ~1% endogenous)
- Bioruptor shearing
  - ~200-400bp dsDNA
  - End repair, A-tailing, adaptor ligation, and 6 cycles of PCR
  - Hybridization to custom RNA baits (see A) bound to streptavidin beads
  - Non-endogenous DNA does not hybridize
  - Select hybridized, bound fragments with a magnet
  - Wash away unbound fragments
  - RNase digest to remove bait sequences. Only captured DNA is retained
  - PCR
  - Fragments are ready for sequencing
Figure S1. Schematic of RNA bait generation and hybridization reaction. (A) Schematic of RNA bait generation. High-quality genomic DNA from a baboon is fragmented to 150 bp (blue fragments). Custom adaptors (purple fragments) with a T7 RNA polymerase site and a restriction enzyme cut site are then ligated to the fragmented DNA. DSN treatment is used to reduce the representation of repetitive elements (orange fragments). Finally, the library is PCR amplified, biotinylated, and transcribed into RNA baits. (B) Schematic of hybridization and capture. A genomic library is generated from fecal DNA (fDNA) sheared to 200-400 bp fragments. This fragment pool originally contains ~1% endogenous DNA (blue fragments) and ~99% environmental/microbial DNA (black fragments). Next, the fDNA library is incubated with 750 ng biotinylated RNA baits. RNA bait-bound DNA (enriched for endogenous DNA fragments) is then separated from the supernatant with a magnet. The RNA baits are digested leaving only the enriched fDNA sample, which can be PCR amplified for high-throughput sequencing.