



System using a 2 × 150 bp run protocol. Raw reads were aligned to the mtDNA reference using BWA,<sup>3</sup> and variants were screened against the mtDNA revised Cambridge reference sequence (rCRS)<sup>4</sup> and called using the Genome Analysis Toolkit (GATK).<sup>5,6</sup> The frequencies of each variant were calculated and compared to establish shared maternal lineage.

## Analysis and Results

### Concordant Sequencing Results of HVI/II Regions

Using this method, ultradeep sequencing coverage (>10,000 reads) was reproducibly obtained for all samples. All variants previously reported in control DNA 9947A (Promega) were identified at high frequencies (Figure 3). There was concordance in variant positions and frequencies between technical replicates. Results for blood samples were concordant with those previously analyzed using CE-based Sanger sequencing.

### Human Identification from Buried Remains

Total DNA was extracted from a bone sample (995A) recovered from a grave at the Southern Battlefield Cemetery in Vietnam. A corresponding blood sample (995B) was also drawn from the presumptive brother of the interred individual and sequenced using Illumina sequencing technology. The variant positions and frequencies in the HVI/II regions were compared (Figure 4). Results indicated that the 2 samples may share maternal lineage, supporting the hypothesis of kinship.

### MiSeq mtDNA Protocols and Analysis Tools

mtDNA library preparation, sequencing, and analysis workflows and tools are available for the MiSeq System. These include a protocol for clean, intact DNA samples,<sup>7</sup> as well as one for difficult samples,<sup>8</sup> such as bone, where sequencing of the mtDNA D-loop hypervariable region is critical. Both incorporate data analysis with MiSeq Reporter and

DLOOP	GEN	Sample 917 (blood)		Sample 926 (blood)		Sample 945 (blood)		Sample 995 (bone #1)	Sample 995 (bone #2)	9947A (Control)
		917B_1	917B_2	926B_1	926B_2	945B_1	945B_2	995A_2	995B_2	9947A_1
913	G->A	0.001	0.004	0.001	0.006	0.975	0.992	0.017	0.002	0.001
973	T->C	0.918	0.951	0.011	0.025	0.024	0.020	0.028	0.032	0.027
976	C->>T	0.002	0.003	0.000	0.003	0.988	0.948	0.003	0.003	0.000
1007	C->>T	0.948	0.938	0.987	0.863	0.970	0.862	0.937	0.978	0.014
1050	C->>T	0.989	0.940	0.042	0.002	0.001	0.002	0.000	0.000	0.004
1053	A->G	0.002	0.004	0.002	0.003	0.003	0.003	0.988	0.952	0.003
1055	T->C	0.001	0.041	0.001	0.024	0.001	0.021	0.998	0.958	0.001
1062	C->>T	0.001	0.005	0.987	0.972	0.008	0.001	0.001	0.000	0.004
1078	C->>T	0.002	0.001	0.913	0.970	0.004	0.001	0.000	0.001	0.005
1081	T->C	0.002	0.008	0.002	0.003	0.994	0.990	0.004	0.006	0.002
1095	T->C	0.003	0.017	0.005	0.005	0.002	0.004	0.002	0.005	0.008
1146	T->C	0.935	0.970	0.007	0.009	0.002	0.007	0.008	0.011	0.001
1426	A->G	0.981	0.987	0.997	0.995	0.995	0.998	0.998	0.998	0.006
1446	A->G	0.006	0.001	0.004	0.003	0.006	0.003	0.002	0.004	0.001
1503	C->>T	0.995	0.996	0.992	0.990	0.990	0.992	0.990	0.997	0.000
1548	T->C	0.017	0.013	0.009	0.003	0.002	0.003	0.004	0.001	0.002
1552	T->C	0.002	0.002	0.003	0.003	0.991	0.989	0.004	0.002	0.001
1557	T->C	0.004	0.004	0.011	0.025	0.003	0.004	0.994	0.988	0.004
1567	A->G	0.007	0.003	0.015	0.002	0.006	0.001	0.004	0.003	0.012
1616	A->G	0.977	0.985	0.980	0.980	0.933	0.954	0.992	0.954	0.988

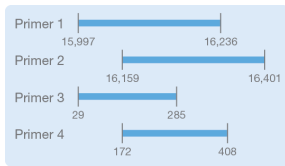
**Figure 3: SNP Results from MiSeq System Show Strong Concordance—** Raw frequencies are shown for each SNP across the D-loop. Blood and control results were concordant with samples previously sequenced with Sanger sequencing.

variant calling with the mtDNA Variant Analyzer Software tool. These same analysis capabilities, along with several usability enhancements, are also available in BaseSpace®, the Illumina genomics computing platform.

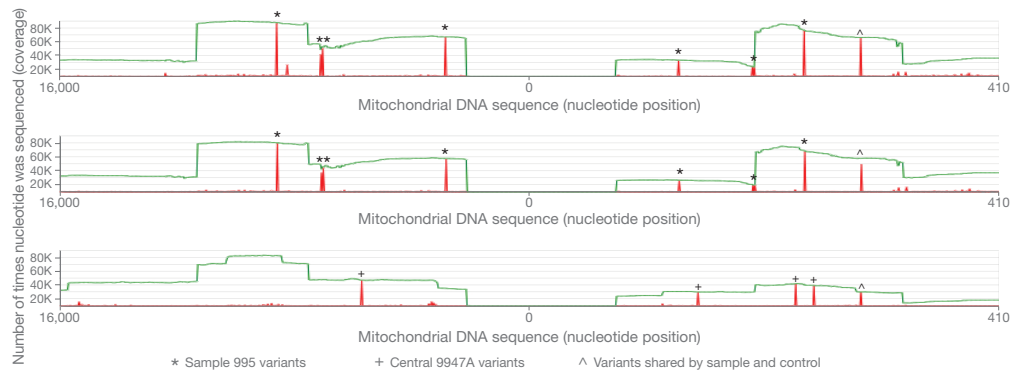
## Conclusions

The Nextera XT DNA Library Preparation Kit and MiSeq System provide robust data for human identification. Sequence derived from the HVI and HVII regions of the mtDNA D-loop region was used to establish that samples from 2 individuals shared a maternal lineage. The Illumina MiSeq System provides a reliable, reproducible method for determining human relationships from degraded samples that are not amenable to analysis by other methods.

DLOOP	GEN	995A_2	995B_2	9947A_1
1007	C->>T	0.927	0.918	0.014
1053	A->G	0.998	0.992	0.003
1055	T->C	0.999	0.999	0.001
1095	T->C	0.002	0.005	0.000
1426	A->G	0.958	0.982	0.006
1446	A->G	0.002	0.004	0.001
1503	C->>T	0.990	0.991	0.000
1548	T->C	0.004	0.001	0.000
1557	T->C	0.006	0.009	0.004
1567	A->G	0.004	0.003	0.004
1616	A->G	0.952	0.954	0.938



Note: Two pairs of amplicons cover each of the HV regions, with overlap of approximately 240 bp within HVI and 110 bp overlap within HVII.



**Figure 4: Illumina Sequencing of Bone Samples Confirms Putative Relationships—** DNA from a bone sample recovered from a grave (995A) and a blood sample (995B) were compared. The high degree of concordance in identified SNPs and profile characteristics suggests the 2 individuals are brothers.

### Learn More

To learn more about the MiSeq System, visit:  
[www.illumina.com/miseq](http://www.illumina.com/miseq)

For more on Illumina sequencing technology for forensic applications, visit: [www.illumina.com/areas-of-interest/forensic-genomics.html](http://www.illumina.com/areas-of-interest/forensic-genomics.html).

To access mtDNA Processor and mtDNA Variant Analyzer in BaseSpace, log in to BaseSpace at [www.illumina.com](http://www.illumina.com), go to the Apps tab, and search for mtDNA Processor and mtDNA Variant Analyzer.

### References

1. Vigilant L, Pennington R, Harpending H, Kocher T, and Wilson A. Mitochondrial DNA sequences in single hairs from a southern African population. *Proc Natl Acad Sci.* 1989; 86: 9350–9354.
2. Wilson M, DiZinno J, Polanskey D, Replogle J, and Budowle B. Validation of mitochondrial DNA sequencing for forensic casework analysis. *Int J Legal Med.* 1995;108: 68–74.
3. Li H and Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics.* 2010; 26: 589–595.
4. Andrews R, Kubacka I, Chinnery P, Lightowlers R, Turnbull D, and Howell N. Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nat Genet.* 1999; 23: 147.
5. McKenna A, Hanna M, Banks E, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 2010; 20: 1297–303.
6. DePristo M, Banks E, Poplin R, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet.* 2011; 43: 491–498.
7. Human mtDNA Genome Protocol ([support.illumina.com/downloads/human\\_mtdna\\_genome\\_guide\\_15037958.ilmn](http://support.illumina.com/downloads/human_mtdna_genome_guide_15037958.ilmn)). Accessed 26 January 2016.
8. Human mtDNA D-loop Hypervariable Region Protocol, ([support.illumina.com/downloads/human\\_mtdna\\_d\\_loop\\_hypervariable\\_region\\_guide\\_15034858.ilmn](http://support.illumina.com/downloads/human_mtdna_d_loop_hypervariable_region_guide_15034858.ilmn)). Accessed 26 January 2016.

AAAGAATGATAACAGTAAACACACTTCTGTAAACCTTAAGATTACTTGATCCACTGATTC AACGTACCGTAAACGACGTATCAATTGAGACTAAATATTAACGTACCATAAGAGCTACCGTGCAACGACGAAAAGAATGATAACAGTAAACACACTTCTGTAAAC  
AATCAACGTACCGTAAACGACGTATCAATTGAGATTACTTGATCCACTGATTC AACGTACCGTAAACGACGTATCAATTGAGACTAAATATTAACGTACCATAAGAGCTACCGTGCAACGACGAAAAGAATGATAACAGTAAACACACTTCTGTAAAC  
AACGACGAAAAGAATGATAACAGTAAACACACTTCTGTAAACCTTAAGATTACTTGATCCACTGATTC AACGTACCGTAAACGACGTATCAATTGAGACTAAATATTAACGTACCATAAGAGCTACCGTGCAACGACGAAAAGAATGATAAC  
TTAAAGGTACCATTAAAGAGCTACCGTGCAACGACGAAAAGAATGATAACAGTAAACACACTTCTGTAAACCTTAAGATTACTTGATCCACTGATTC AACGTACCGTAAACGACGTATCAATTGAGACTAAATATTAACGTACCATAAGAGCTACCGTGCAACGACGAAAAGAATGATAAC  
AAAGAAATGATAACAGTAAACACACTTCTGTAAACCTTAAGATTACTTGATCCACTGATTC AACGTACCGTAAACGACGTATCAATTGAGACTAAATATTAACGTACCATAAGAGCTACCGTGCAACGACGAAAAGAATGATAAC  
AAAGATTACTTGATCCACTGATTC AACGTACCGTAAACGACGTATCAATTGAGACTAAATATTAACGTACCATAAGAGCTACCGTGCAACGACGAAAAGAATGATAAC  
AACGTATCAATTGAGACTAAATATTAACGTACCATAAGAGCTACCGTGCAACGACGAAAAGAATGATAAC

