mtDNAmanager:

A Forensic Mitochondrial DNA Database Aimed at Supporting Data Quality Control and Generating Reliable Frequency Estimates

Hwan Young Lee[1], Injee Song[2], Eunho Ha[3], Sung-Bae Cho[2], Woo Ick Yang[1], Kyoung-Jin Shin[1]

(1) Department of Forensic Medicine, Yonsei University College of Medicine, Seoul, Korea
(2) Department of Computer Science, Yonsei University, Seoul, Korea
(3) Department of Information and Statistics, Yonsei University at Wonju, Gangwon-do, Korea

ISFG

YONSEI UNIVERSITY COLLEGE OF MEDICINE

mtDNAmanager (http://mtmanager.yonsei.ac.kr)

- The goal of mtDNAmanager is to provide a web-based forensic mitochondrial DNA bioinformatics resource for supporting **data quality control** and generating **reliable frequency estimates** using a new approach based on haplogroup estimation and data comparison with the contents of a given database.

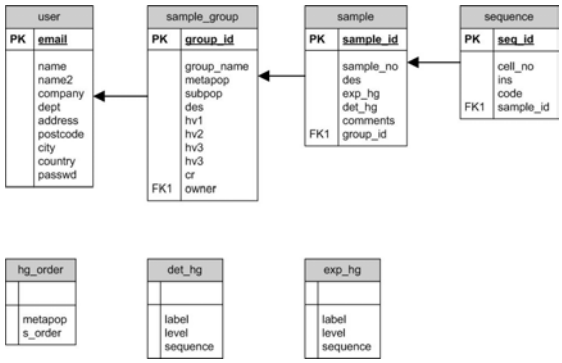
- mtDNAmanager consists of previously reported **high quality mtDNA sequences**, and a set of **bioinformatics tools**, able to automatically characterize newly submitted data by **estimating its haplogroup according to the haplogroup-specific control region mutation motif**.

ISFG

YONSEI UNIVERSITY COLLEGE OF MEDICINE

mtDNAmanager structure

- The database is of relational type, designed and implemented using the multithreaded, multi-user **SQL database management system (DBMS), MySQL**.
- Web interfaces have all been implemented using Asynchronous JavaScript and XML (AJAX) technique based on **Javascript and PHP**.
- The system is optimized with **internet explorer v. 6.0 or more**.



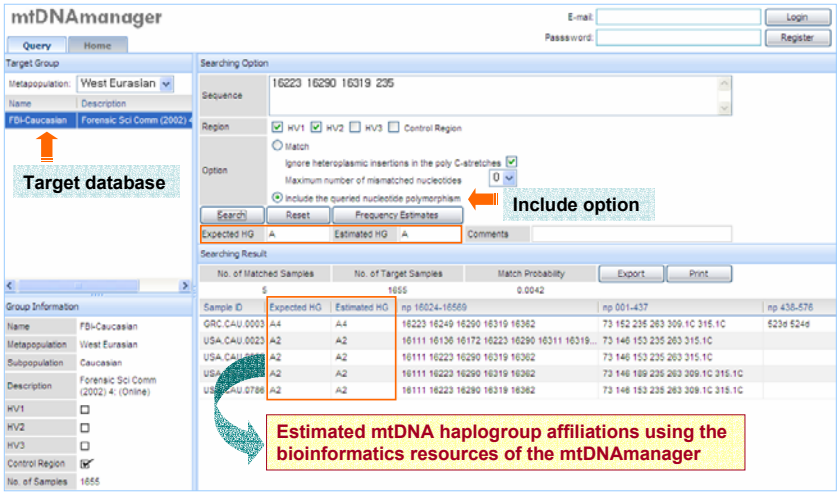
mtDNAmanager Database

- The mtDNAmanager's first release contains **4839 mtDNA control region sequences from FBI** and **593 Korean mtDNA control region sequences**. We will continue the effort to collect and integrate high quality mtDNA control region sequence data as much as possible.
- The mtDNAmanager database will be extended as follows in the near future:

Metapopulation	Subpopulation	No. of sequences	References
African	African	1148	Forensic Sci Commun (2002) 4 *Online*
	Kenya	100	Int J Legal Med (2004) 118: 294-306
West Eurasian	Caucasian	1655	Forensic Sci Commun (2002) 4 *Online*
	Austrian	273	Forensic Sci Int (2007) 166: 164-175
	Finnish	200	Forensic Sci Int (2007) *In Press*
	German	313	Forensic Sci Int (2007) *In Press*
	Hungarian	416	Int J Legal Med (2007) *In Press*
	Rumanian	360	Ann Hum Genet (2007) *In Press*
East Asian	East Asian	753	Forensic Sci Commun (2002) 4 *Online*
	Japanese	211	Int J Legal Med (2003) 117:218-225
	Korean	593	Int J Legal Med (2006) 120:5-14
Oceanian	-	-	-
Admixed	Hispanic	686	Forensic Sci Commun (2002) 4 *Online*
Total		**6708**	

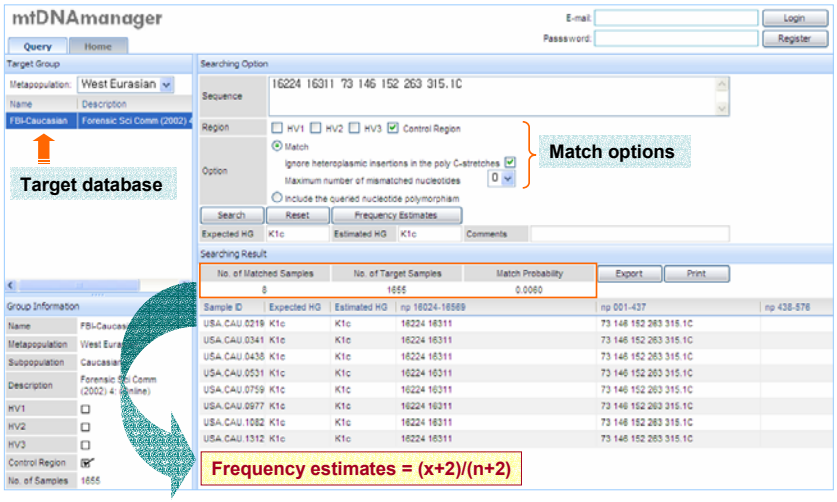
Database search

- By default but with an exchangeable setting, a **query system** retrieves sequences which **include the queried nucleotide polymorphisms** therein from a user-selected database.



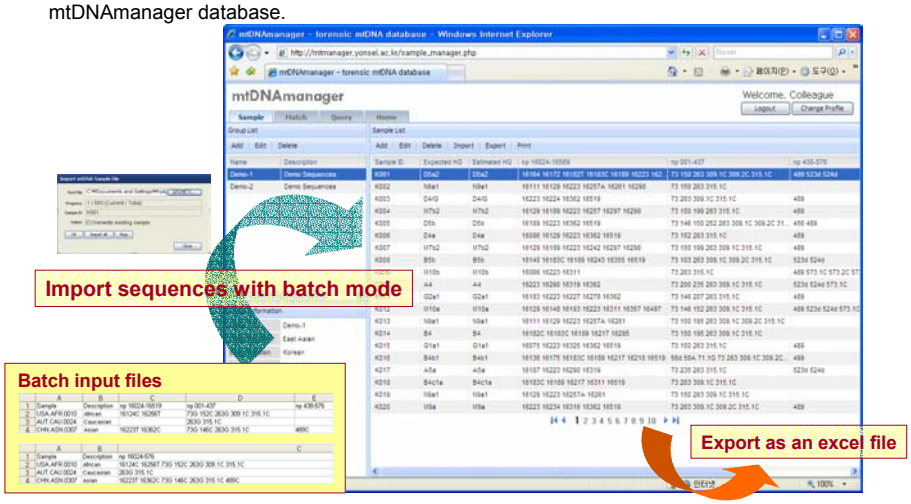
Database search and statistics

- A query system of the mtDNAmanager also searches sequences which show **a match with the queried sequence data** from the database.



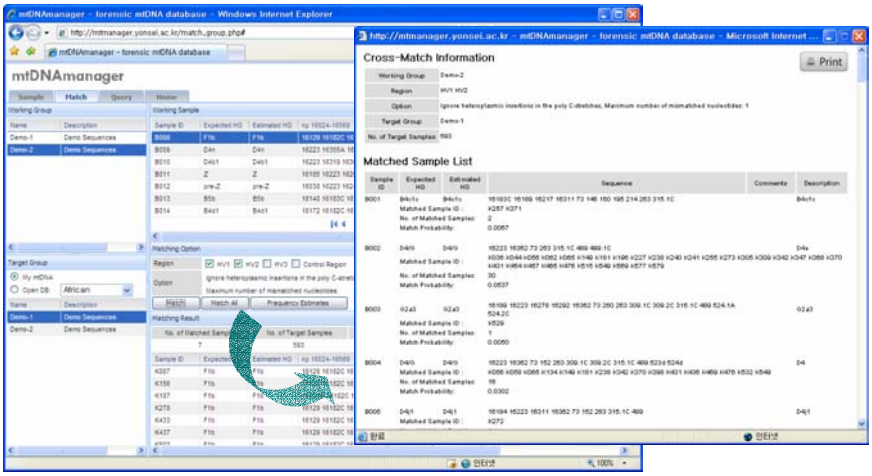
My mtDNA

- **Upon registration**, the mtDNAmanager provides web interfaces where **users can submit and store their own data with batch mode** and search sequences which show a match or include the queried nucleotide polymorphisms from their own database as well as the mtDNAmanager database.



Cross-match between data groups

- **Upon registration**, a match system permits **cross match of all sequence data between two selected groups** as well as retrieval of matched sequences for one of the data of a user-created data group from their own database or from the mtDNAmanager database.



mtDNA haplogroup estimation

Sample List

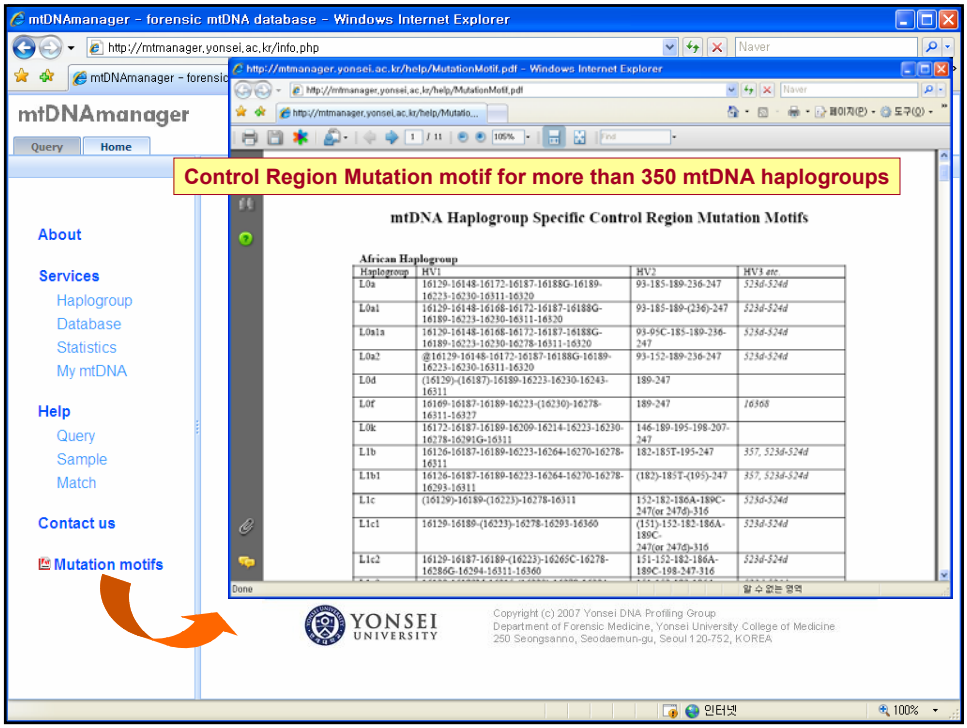
Add | Edit | Delete | Import | Export | Print

Data from EMPOP

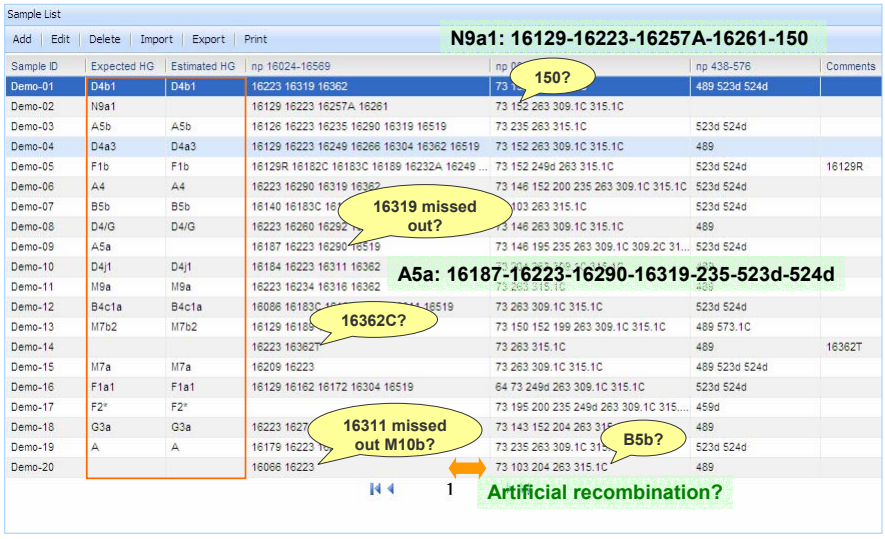
Sample ID	Expected HG	Estimated HG	np 16024-16569	np 001-437	np 438-576	Comments	Description
f1D5	T1	T1	16126 16163 16186 16189 16294 16519	73 195 263 309.1C 315.1C			T1
f1D6	T2	T2	16126 16294 16296 16519	73 150 200 263 315.1C			T2
f1E1	H11a	H11a	16293 16311	143 195 263 309.1C 315.1C			H*
f1E2	H	H	16291 16519	263 315.1C			H*
f1E3	H	H	16093 16271 16519	263 309.1C 315.1C			H*
f1E4	H1b	H1b	16189 16356 16362 16519	263 309.1C 315.1C	523d 524d		H1b
f1E5	X2	X2	16134 16183C 16189 16223 16278 16519	73 195 263 309.1C 309.2C 315.1C			X2d
f1E6	J1c	J1c	16069 16126 16261 16265 16319	73 185 228 263 295 309.1C 315.1C	462 489 524.1A 524.2C...		J1c
f1F1	H	H	16234 16293 16519	263 315.1C 315.2C			H*
f1F2	W	W	16223 16282 16295 16519	73 119 189 195 204 207 263 315.1C			W
f1F3	K1a	K1a	16222 16224 16311 16362 16519	73 114 263 309.1C 315.1C	497		K1a
f1F4	T1a	T1a	16126 16163 16186 16189 16294 16519	73 152 195 263 309.1C 315.1C			T1a
f1F5	H5	H5	16260 16304	263 309.1C 315.1C	456		H5
f1F6	U5a	U5a	16192 16256 16270 16399	73 143 263 315.1C			U5a1
f1G1	HV0	HV0	16153 16298	72 93 195 263 315.1C			HV0
f1G2	H	H	16311	263 309.1C 315.1C	523d 524d		H*
f1G3	H	H	16325 16519	263 309.1C 315.1C			H*
f1G4	H5	H5	16304	263 309.1C 309.2C 315.1C	456		H5
f1G5	T1	T1	16126 16163 16186 16189 16294 16519	73 195 263 315.1C			T1
f1G6	C	C	16093 16223 16234 16288 16298 16327 16518...	73 249d 263 309.1C 315.1C	489		C

Control region haplogroups | 1 2 3 4 5 6 7 8 9 10 | Coding region haplogroups

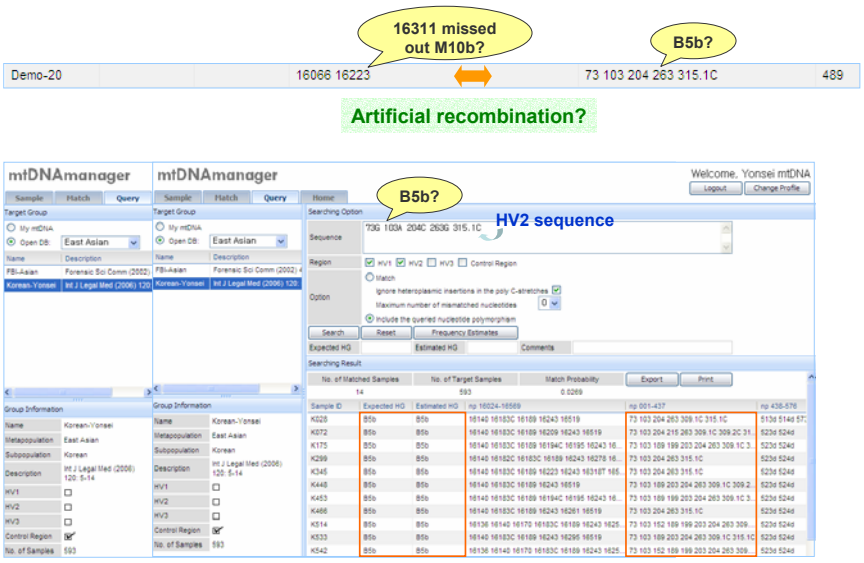
Using the bioinformatics tools of mtDNAmanager, **more than 98% of mtDNAs** of high quality datasets can be allocated to the same mtDNA haplogroups as those confirmed with the coding region SNP information.

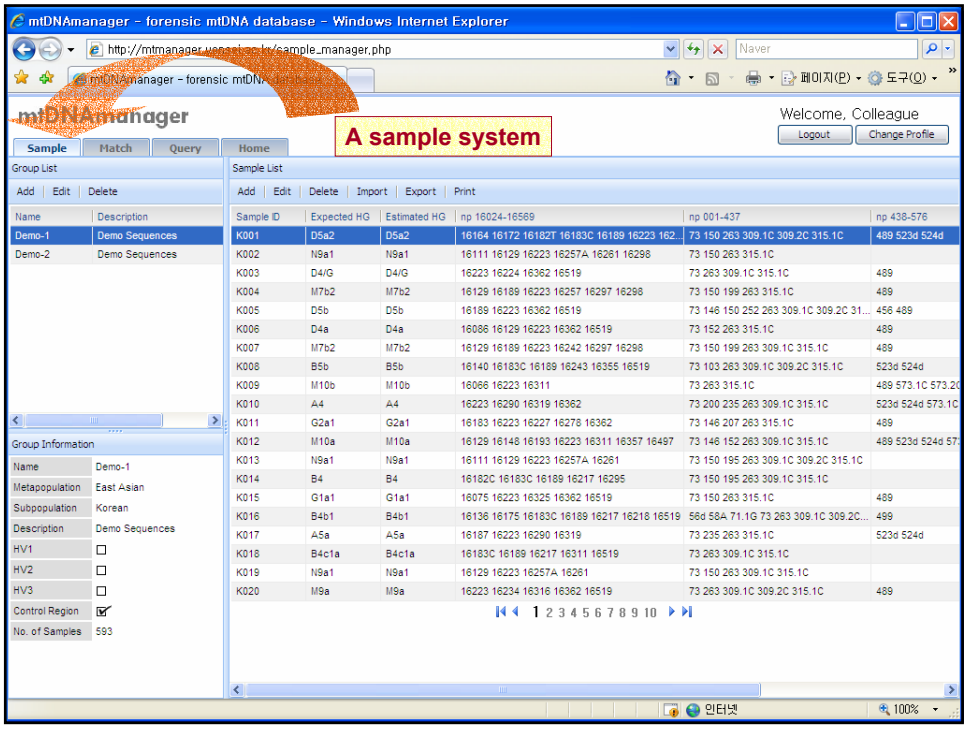


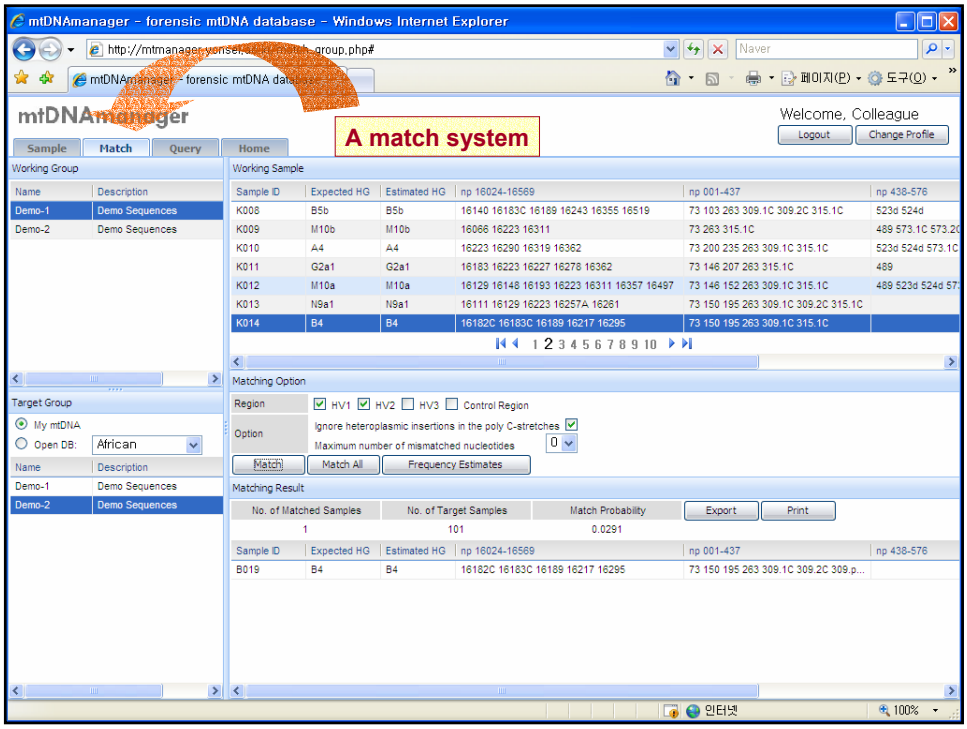
mtDNA data quality control

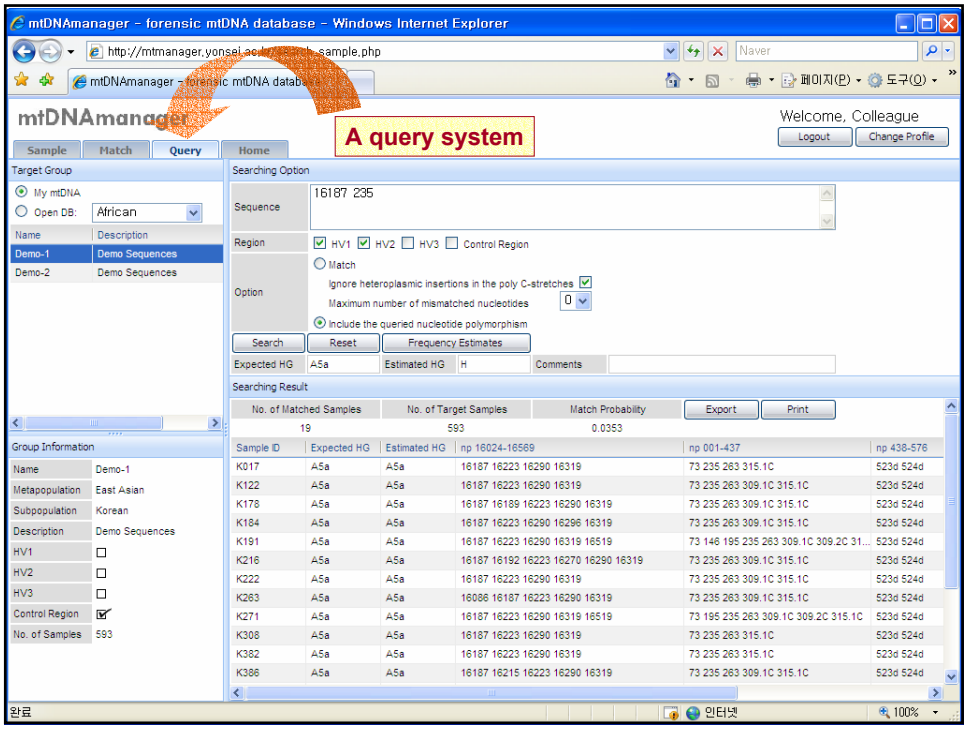


Detection of errors in mtDNA sequence




mtDNAmanager – forensic mtDNA database – Windows Internet Explorer
A sample system
mtDNAmanager
Welcome, Colleague
Logout
Change Profile
Sample
Match
Query
Home
Group List
Add | Edit | Delete
Sample List
Add | Edit | Delete | Import | Export | Print
Group Information


mtDNAmanager – forensic mtDNA database – Windows Internet Explorer
A match system
mtDNAmanager
Welcome, Colleague
Logout
Change Profile
Sample
Match
Query
Home
Working Group
Working Sample
Target Group
Matching Option
Matching Result
Match
Match All
Frequency Estimates
Export
Print



Concluding remarks

- The mtDNAmanager is a forensic mitochondrial DNA database aimed at **supporting data quality control** and **generating reliable frequency estimates** using an approach based on haplogroup estimation and data comparison.
- The **mtDNAmanager** provides **a convenient web interface** for analysis, query and storage of human mtDNA control region sequences.
- For comments, bug reports, suggestions for improvement, please contact us through the website **(http://mtmanager.yonsei.ac.kr)**.

ISFG

YONSEI UNIVERSITY COLLEGE OF MEDICINE