

## Update on Microhaplotypes for Casework

Kenneth K. Kidd  
for  
The 26+ members of the Microhaplotype Working Group  
An unofficial collaboration

**Green Mountain DNA Conference**  
Burlington, Vermont  
July 26, 2023



## The Microhaplotype Working Group

An informal international collaboration  
Promoting microhaplotypes for forensic casework  
by  
Working to set standards for microhaplotypes  
and  
Coordinating the work in different labs



## Microhaplotypes are Better than the CODIS Markers!



## What is the early history of microhaplotypes?

Microhaplotypes are short DNA sequences, typically consisting of two or more single nucleotide polymorphisms (SNPs) that are inherited together as a block. They have gained significant attention in recent years as a tool for genetic analysis and forensic applications. However, the concept of microhaplotypes and their early history can be traced back to the field of population genetics and the study of genetic variation.


chatGPT, July 14, 2023



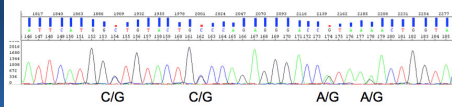

### What is a Microhaplotype?

A SNP-based multiallelic locus  
Small enough amplicon for a single read


Characteristics include:  
All alleles at a locus are the same size  
Low mutation rate  
No stutter



### The Functional Effects of MPS

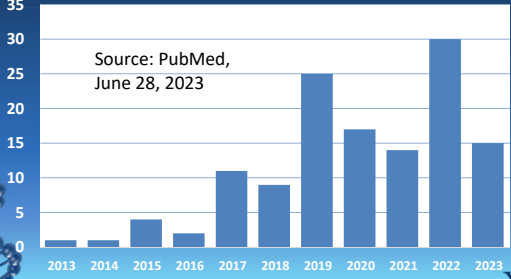



MPS methods allow clonal sequencing of individual strands, thereby distinguishing the parental haplotypes at a locus




### Numbers of Papers published per year on Microhaplotypes

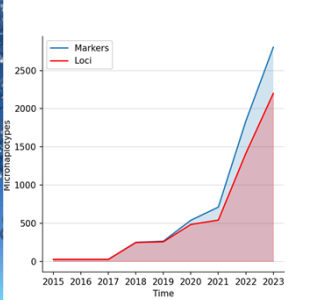
Source: PubMed, June 28, 2023




Year	Number of Papers
2013	1
2014	1
2015	4
2016	2
2017	11
2018	9
2019	25
2020	17
2021	14
2022	30
2023	15



### MicroHapdb contents




•Standage et al. 2023. The MicroHapDB database, <https://osf.io/5xnuf/>.



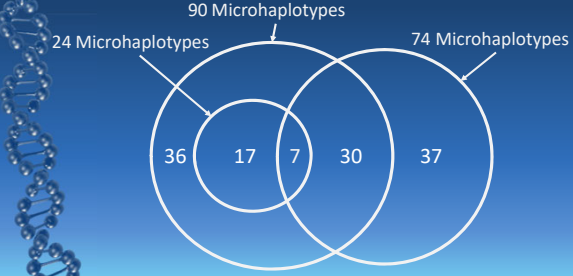
### Forensic Uses of Microhaplotypes

- ❖ Individualization—random match probabilities
- ❖ Ancestry inference—biogeographic origins
- ❖ Familial relationships—identification of biologic relationships; Paternity Index
- ❖ Mixture deconvolution—identifying individual components of a mixture

**Microhaplotypes are better than STRs in all four**



### Evolution of Microhaplotype panels




90 Microhaplotypes

24 Microhaplotypes

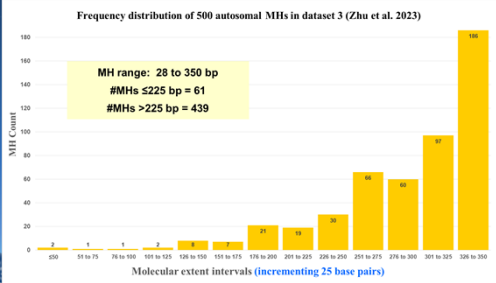
74 Microhaplotypes

36 17 7 30 37




### Searching for new MH with $A_e > 5$

Frequency distribution of 500 autosomal MHs in dataset 3 (Zhu et al. 2023)

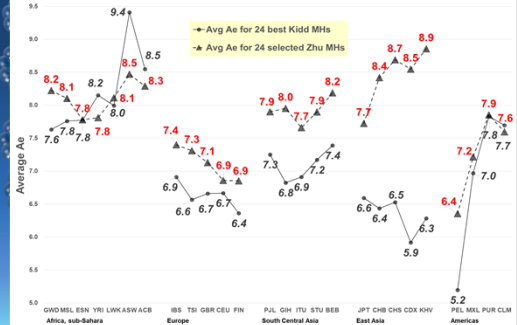


MH range: 28 to 350 bp  
 #MHs  $\leq 225$  bp = 61  
 #MHs  $> 225$  bp = 439

Molecular extent intervals (incrementing 25 base pairs)




### Average $A_e$ of two panels of markers




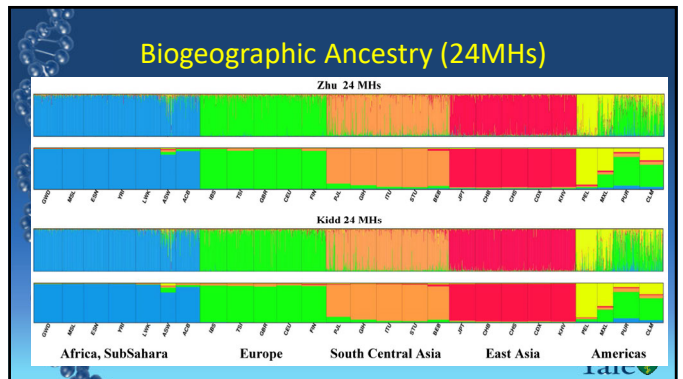
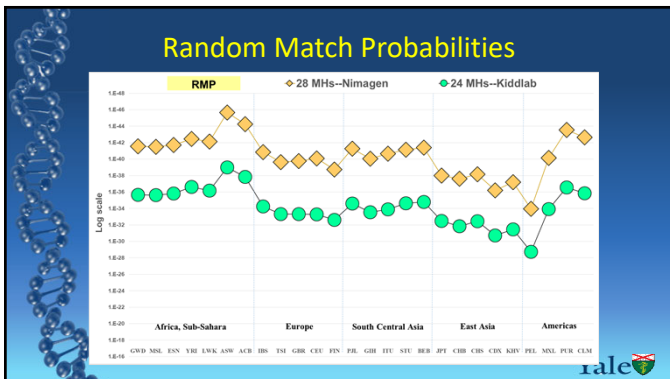
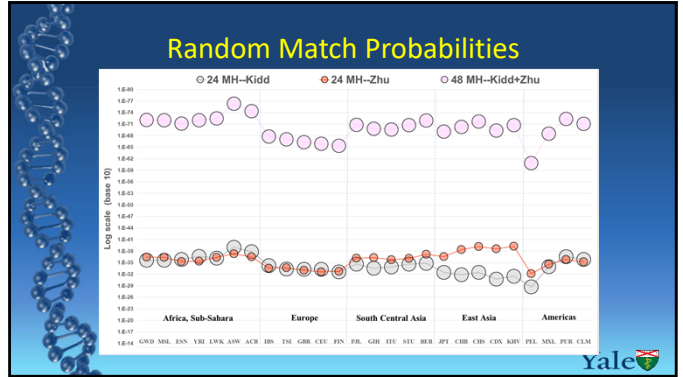
● Avg  $A_e$  for 24 best Kidd MHs  
▲ Avg  $A_e$  for 24 selected Zhu MHs

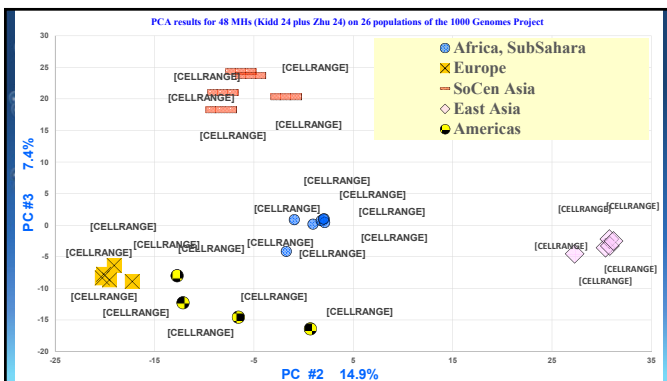
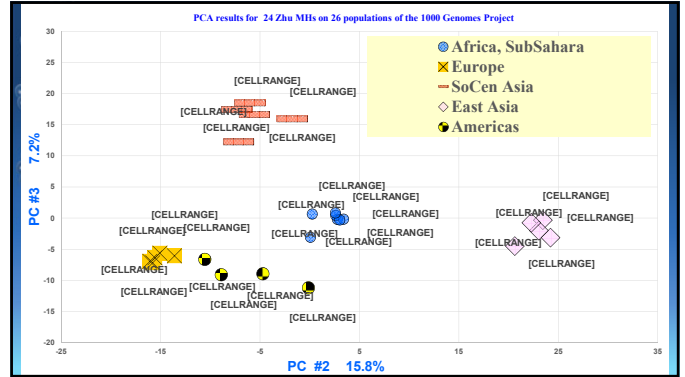
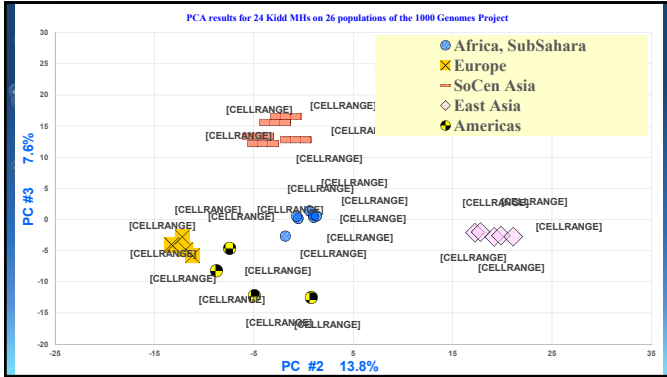
Zhu 24 ▲  
 Kidd 24 ●



## Individualization

Random Match Probability (RMP)  
is one measure of the  
relative uniqueness of individuals



### Adapting an established Ampliseq microhaplotype panel to nanopore sequencing through direct PCR

Direct PCR from swabs      Library construction by ligation of 91 microhaplotypes      Nanopore sequencing in the MinION instrument

Highly scalable workflow based on Flongle flow cells useful for urgent kinship analysis


Good sensitivity -> >90 % profile completeness down to 250 pg input DNA

Direct capture PCR -> applicable in field operations

### Panel of 107 autosomal MHs

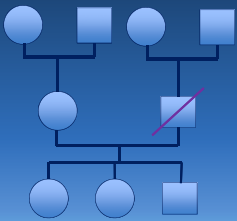

# SNPs	# loci	avg Ae	min	max
3	3.00	2.5	2.4	2.64
4	64.00	3	2.17	3.55
5	28.00	3.47	2.61	4.29
6	11.00	3.87	3.12	5.58
7	1.00	4.03	4.03	4.03

Forensic Genetics Unit, Institute of Forensic Sciences,  
University of Santiago de Compostela, Spain



### Placing Individuals into Family Contexts

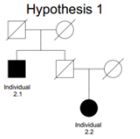
- Unknown remains
- Missing Persons
- Paternity Testing
- Relationship Testing

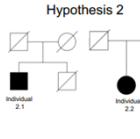
### Forensic Genetics Unit, Institute of Forensic Sciences, University of Santiago de Compostela, Spain

Testing relationship questions using a panel of 100 microhaplotypes tested using the nanopore MinION platform

Hypothesis 1




Hypothesis 2



Markers included	LR obtained (H1 / H2)
GlobalFiler	6.14
GlobalFiler + HDplex	571.98
GlobalFiler + HDplex + Pentagon	28.43
Microhaplotypes	40174.32


De la Puente et al. (2020)  
*Forensic Science International: Genetics*  
45:102213

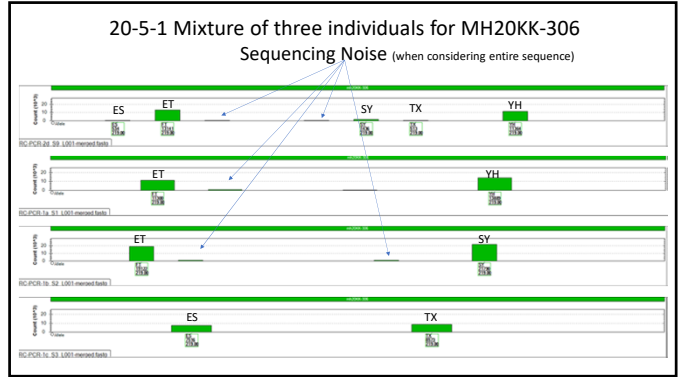
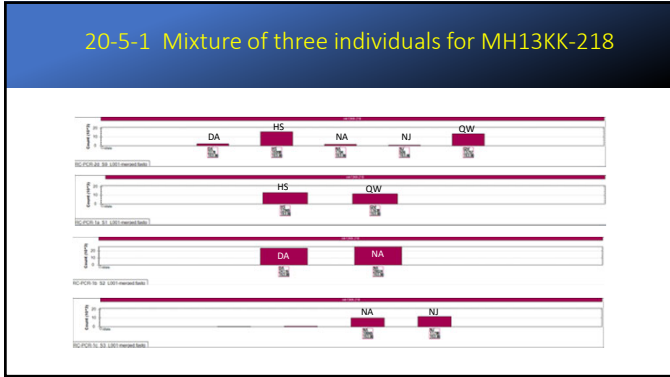


### Deconvoluting Mixtures

Work by the Podini Lab at The George Washington University

They are using the NimaGen 24 MH panel mentioned before





### Key to Nomenclature in MicroHapDB

lab ID (2 to 4 characters)

preface | alt. allele definitions

**mh04KKxx-###.v#**

chromosome number | locus ID # specific to lab (3 to 6 characters/numerals)

The – and . are necessary delimiters

Standage et al. 2023. the MicroHapDB database, <https://osf.io/5xnuf/>.

### MicroHapDB Contents

- Microhaplotypes (see attached graphic for growth over time)
  - 2804 distinct marker (SNP) definitions
  - 2200 distinct loci
- Frequencies
  - 56505 haplotypes
  - 125 population groups
  - 894905 total microhap frequencies
  - 22 paper sources (see <https://microhapdb.readthedocs.io/en/stable/citations.html>)
  - 13 distinct labs/PIs (see table)

Standage et al. 2023. the MicroHapDB database, <https://osf.io/5xnuf/>.

## Laboratories represented in MicroHapdb

lab code	Principal Investigator	Institution
CP	Peng Chen	Nanjing Medical University
FHL	Haoliang Fan	Southern Medical University
HYP	Yiping Hou	Sichuan University
KK	Kenneth Kidd	Yale University
LV	Lev Voskoboinik	Hebrew University of Jerusalem
NH	Nakahara Hiroaki	Juntendo University
PK	Peter de Knijff	Leiden University
SCUZJ	Ji Zhang	Sichuan University
SHY	Hongyu Sun	Sun Yat-sen University
USC	Maria de la Puente / Chris Phillips	University of Santiago de Compostela
WL	Le Wang	Institute of Forensic Science, Chinese Ministry of Public Security
ZBF	Bo-Feng Zhu	Southern Medical University
ZHA	Lagabaiyila Zha	Central South University



## Conclusions

- ❖ Microhaplotypes (MHs) *are becoming* “The Next Generation Forensic DNA Marker” lacking only reference databases of “criminals”
- ❖ MHs fulfill all of the requirements for a multi-locus forensic panel: individualization, ancestry inference, relationship identification, mixture deconvolution
- ❖ MHs are better than STRs for several reasons



## The Next Steps

- ❖ The forensic community needs to expand the pool of vetted microhaplotypes
  - ❖ Vetting includes checking sequence, population data
- ❖ The forensic community needs to decide on how many MH should be in a panel.
  - ❖ 24 is not enough; 90 is probably too many
  - ❖ The number depends on the average  $A_e$
- ❖ There may be different panels for different purposes
- ❖ Decisions on which MHs to include in panels need to be made.
- ❖ Probabilistic Genotyping for MHs needs to be developed.



My colleagues and I thank you  
for your attention

