



Challenges of DNA Profiling in Mass Disaster Investigations

Antonio Alonso, Pablo Martín, Cristina Albarrán, Pilar García, Lourdes Fernández de Simón, María Jesús Iturralde, Amparo Fernández-Rodríguez, Inmaculada Atienza, Javier Capilla, Julia García-Hirschfeld, Pilar Martínez, Gloria Vallejo, Oscar García¹, Emilio García², Pilar Real², David Álvarez², Antonio León², Manuel Sancho

National Institute of Toxicology and Forensic Science, Biology Service, Madrid; ¹Ertzaintza Laboratory Area, Erandio, Bizkaia; and ²General Commissariat of Scientific Police, Scientific Analysis Service, DNA laboratory, Madrid, Spain

Abstract

In cases of mass disaster, there is often a need for managing, analyzing, and comparing large numbers of biological samples and DNA profiles. This requires the use of laboratory information management systems for large-scale sample logging and tracking, coupled with bioinformatic tools for DNA database searching according to different matching algorithms, and for the evaluation of the significance of each match by likelihood ratio calculations. There are many different interrelated factors and circumstances involved in each specific mass disaster scenario that may challenge the final DNA identification goal, such as: the number of victims, the mechanisms of body destruction, the extent of body fragmentation, the rate of DNA degradation, the body accessibility for sample collection, or the type of DNA reference samples availability. In this paper, we examine the different steps of the DNA identification analysis (DNA sampling, DNA analysis and technology, DNA database searching, and concordance and kinship analysis) reviewing the “lessons learned” and the scientific progress made in some mass disaster cases described in the scientific literature. We will put special emphasis on the valuable scientific feedback that genetic forensic community has received from the collaborative efforts of several public and private USA forensic laboratories in assisting with the more critical areas of the World Trade Center (WTC) mass fatality of September 11, 2001. The main challenges in identifying the victims of the recent South Asian Tsunami disaster, which has produced the steepest death count rise in history, will also be considered. We also present data from two recent mass fatality cases that involved Spanish victims: the Madrid terrorist attack of March 11, 2004, and the Yakolev-42 aircraft accident in Trabzon, Turkey, of May 26, 2003.

A mass disaster is an unexpected event that causes serious injury and death to a number of people. Mass disaster events may be natural disasters (earthquakes, flooding, and tornadoes), accidental disasters (aircraft crashes, train crashes and derailments, and building fires), or intentioned terrorism acts (direct attacks on significant objectives, and bombing of populated areas, including suicide attacks and deployments of chemical and bio-

logical weapons). Forensic identification of victims is essential for humanitarian reasons, but also for civil or criminal investigative needs, and it is essentially based on forensic anthropology, fingerprints, forensic odontology, radiology, and DNA typing (1).

The main task of forensic DNA laboratories faced with mass disaster cases is to help name every anonymous victim, thus bringing closure to

surviving family members and friends. This is done by matching DNA profiles of postmortem tissue samples with those of antemortem DNA samples (personal items or biological specimens) or by kinship analysis with living relatives.

Generally, mass disaster cases require managing, analyzing, and comparing large numbers of biological samples and DNA profiles (mainly autosomal short tandem repeat [STR] profiles but also occasionally mtDNA sequence and Y-Chromosome STR [Y-STR] haplotype data) making necessary the use of electronic laboratory information management systems for large-scale sample logging and tracking, coupled with bioinformatic tools for DNA database searching according to different matching algorithms (ie, complete or partial allele sharing on each locus for autosomal STR markers), and software solutions to evaluate the significance of each match by likelihood ratio (LR) calculations.

There are many different interrelated factors and circumstances involved in each specific mass disaster scenario that may challenge the final DNA identification goal, such as: the number of victims, mechanisms of body destruction, the extent of body fragmentation, rate of DNA degradation, the body accessibility for sample collection, or type of DNA reference samples availability.

In this paper, we shall examine the different steps of the DNA identification analysis (DNA sampling, DNA analysis and technology, DNA database searching, and concordance and kinship analysis) reviewing the "lessons learned" and the scientific progress made in some mass disaster cases described in the scientific literature (2-13). We will put special emphasis on the valuable scientific feedback that the genetic forensic community has received from the collaborative efforts of several public and private USA forensic laboratories that formed part of the advisory Kinship and Data Analysis Panel (KADAP) to advise and assist with the more critical areas of the World Trade Center (WTC) mass fatality of September 11, 2001. This primarily included DNA technology developments to identify the most severely degraded remains (11,14) but also other aspects like mass fatality response (1), administrative matters of sample collection and information management (15), and statistical DNA interpretation issues (12).

The main challenges in identifying the victims of the recent South Asian Tsunami disaster,

which has produced the steepest death count rise in history (more than 200,000 victims), will also be considered. We also present data from two recent mass fatality cases that involved Spanish victims: the Madrid terrorist attack of March 11, 2004, and the Yakolev-42 aircraft accident in Trabzon, Turkey of May 26, 2003.

DNA Sampling and Information Management

Experiences gained from previous mass fatality incidents reinforce the need to make all necessary steps to guarantee sample preservation for DNA analysis and to use suitable protocols for documenting the chain of custody of DNA sampling and body tracking. To help with this purpose, specialized and trained disaster victim identification (DVI) multidisciplinary teams composed of medical examiners, forensic pathologists, anthropologists, forensic odontologists, fingerprint specialists, radiologists, and experts in search and recovery of physical evidence have been developed worldwide. Some examples of federal resources providing aid to local communities in mass disaster response in the USA are the Disaster Mortuary Operational Response Teams (DMORT) (16), the Federal Bureau of Investigation's Evidence Response Team (ERT) (17), and the Office of the Armed Forces Medical Examiner (OAFME) (18). Different Interpol DVI teams have also been developed worldwide and a standing committee on DVI is responsible for recommending measures for improving identification procedures, by encouraging international co-operation and standardization (19).

Additionally, recent guidelines were published to assist the medical examiner with the whole process of victim identification in mass fatalities, including detailed procedures for DNA sample collection and data management (1).

Victim Sample Recovery

One of the first and major goals in mass disaster cases, with high impact on the real scope of the identification process, is to establish the goal of the DNA analysis: whether to perform DNA analysis on each victim or just from a subset of victims (those not identified by other forensic methods, for instance), or try to identify by DNA analysis all human remains recovered (or just to identify a subset of remains: only the recognizable

body fragments). This obviously depends on the specific circumstances of each mass disaster. In most disasters (Table 1), the standard of care is usually to identify each victim, but under certain conditions, like those after the WTC tragedy, with very high level of body fragmentation and sample disintegration from an open population, the challenge is to identify each remain by DNA testing (10,11).

As a general rule, it would be advantageous to always collect and adequately preserve a DNA sample from each autopsied body, even when the identifications were primarily performed by other forensic methods such as by means of fingerprints or dental records, to allow DNA re-association studies with potential body fragments and future DNA analysis in the case of doubts or contradictions. It is recommended to carry out sample collection of human remains for DNA analysis during autopsy in conjunction with other forensic experts, such as medical examiners and anthropol-

ogists, ensuring photographic documentation of the remains and taking proper precautions to minimize the risk of contamination.

Sample preference for DNA analysis is also determined by each mass disaster scenario. In general, it is recommended to collect samples from the least affected material in a way to avoid both exogenous contamination and body cross-contamination. The preferable human remain sources include: soft tissues (skeletal muscle, organ tissues and skin) and blood. A new system for bar-coded based soft tissue collection and simultaneous body tracking allowing large scale tissue sampling and long-term DNA preservation under desiccation conditions with potential applications in mass fatalities has been recently described (20). Hard tissues (bone and teeth) are the preferable samples when body putrefaction or other environmental insults preclude DNA preservation in soft tissues. Bone has also been sampled to evaluate

Table 1. Representative mass disaster cases investigated by means of DNA technology and described in the scientific literature that were classified chronologically

Mass disaster	Case/location/date	No. of victims	Remains analyzed by DNA	Main challenges	References
Aircraft accident fatality	Airbus A320 air crash/Mount Sainte-Odile (France), January 20, 1992	87	17	reduced number of PCR markers available	(2)
Collective suicide/genocide	Waco disaster/Waco (Texas), April 19, 1993	83	73	remains extensively charred; reduced number of PCR markers available	(3,4)
Terrorist bombing	Argentine-Israeli Association explosion/Buenos Aires (Argentina), July 18, 1994	>100	70		(5)
Aircraft accident fatality	Spitsbergen aircraft accident/Spitsbergen (Norway), August 1996	141	257		(6)
Aircraft accident fatality	Taoyuan Airbus crash accident/Taoyuan (Taiwan), February 16, 1998	202	685	large-scale pair-wise genotype comparisons; large number of families among the victims	(7)
Aircraft accident fatality	air crash accident/Philippines, February, 1998	104	187	low success rate with STR loci	(8)
Aircraft accident fatality	Swissair Flight 111 accident/Atlantic Ocean off Canada's coastline, September 2, 1998	229	1277	large-scale pair-wise genotype comparisons	(9)
Terrorist attack	World Trade Center disaster/New York (USA), September 11, 2001	2749	19,963	large-scale pair-wise genotype comparisons; high DNA degradation, sample disintegration	(10-12)
A tunnel-bound cable car disaster	Kaprun cable car fire disaster/Kaprun (Austria), November 11, 2000	155	155		(13)
Aircraft accident fatality*	Yakolev-42 air crash accident/Trabzon (Turkey), May 26, 2003 (a)	74	85		this study
Terrorist bombing†	Madrid train bombing case/Madrid (Spain), March 11, 2004 (b)	191	220		this study
Natural disaster	South Asian Tsunami, December 26, 2004	>200000	?	large-scale pair-wise genotype comparisons; low rate of body recovery. High DNA degradation; large number of family groups among the victims; lack of reference DNA samples; lack of technical resources?	(31-33)

*In the Yakolev-42 air crash accident case, 62 out of 74 total victims were Spanish military personnel on their way home from a peacekeeping mission in Afghanistan. Thirty out of 62 bodies were documented as unidentified, whereas 32 were positively identified by the Turkish forensic team. All bodies were given by the Turkish authorities to the Spanish military commission in charge to carry out the process of identification and repatriation of the corpses. Corpses were given to families in Spain without further (documented) identification analysis. One year later, comparative DNA analysis among post-mortem body samples taken by the Turkish authorities and reference samples from victims' relatives demonstrated that the thirty unidentified cases were miss-identified and consequently each family received a wrong body. A new DNA analysis from the exhumed bodies was carried out that confirmed all errors and offered concordance with the results obtained by the Istanbul Forensic Science Council from postmortem body samples.

†In this case, a decision was taken only to perform DNA typing from autopsied bodies (62 samples) that were not identified by fingerprint analysis. This decision impeded later DNA re-association studies with more than 158 body fragments that were collected from the different train scenes.

body cross-contamination when high level of soft tissue commingling is suspected (10).

A recommended procedure for reducing errors during data collection is also the use of specific and standardized sample collection forms employing a unique numbering systems to identify each remain in conjunction with the use of Laboratory Information Management Systems (LIMS) which ensure sample information logging on a centralized database (1,15).

Direct and Family DNA References

Two types of reference samples are usually collected for DNA comparison with mass disaster remains: appropriate family references and direct references, such as personal effects, or antemortem biological specimens, such as biopsies and bloodstain cards.

At present, the simplest and most efficient method of DNA identification is to match each STR multilocus victim's profile to a direct antemortem sample of the victim. Personal items, like toothbrushes, and used shavers and razors have been extensively used as direct references in many cases. The main drawbacks of this strategy are the potential source attribution errors, leading to false exclusions, and the presence of exogenous body fluid or cell debris contaminations leading to mixed DNA profiles. Therefore, these samples cannot be used for exclusionary purposes and, whenever possible, a match obtained with a direct reference should be confirmed through kinship analysis or an analysis of a second direct reference sample (1).

The possibility of using antemortem biological specimens (like bloodstain cards) as direct references with an accredited and documented chain of custody will overcome the main drawbacks of personal items. Indeed, the establishment of a DNA repository to store reference bloodstain cards has been recommended to facilitate the identification of military personnel in air crash accidents (21).

Sample preference for family references depends on the type of DNA analysis. The most discriminative power is obtained by using a large number of nuclear STR markers (from 13 to 17 markers) to analyze the following family references: (a) either or both biological parents of the victim, (b) biological mate of the missing person and their child/children, and (c) multiple biologi-

cal full siblings (sharing the same parent as the victim). The analysis of haploid DNA markers with just maternal (mitochondrial DNA) or just paternal (Y-Chromosomal markers) inheritance allow the use of maternally or paternally-related family members as references.

Buccal swabs or blood (collected by venipuncture or by finger stick devices) are the recommended samples for both nuclear and mitochondrial DNA analysis.

The use of standardized collection forms for both direct and family reference samples by trained interviewers, preferably using electronic forms, whenever possible, to avoid handwriting as much as possible, as well as the use of specific sample collection kits (22) will improve the reliability of the sample selection and donor information.

The information data of each reference sample should also be logged, with a unique numbering code using the LIMS system, into the same central database used for victim samples. This allows rapid transference of edited DNA profiling results to each registered sample for subsequent pair-wise comparisons, maintaining and documenting the chain of custody through the whole DNA identification process.

DNA Analysis

Current and New Technologies

Multiplex PCR amplification of a variable number of autosomal short tandem repeats (STRs) loci is at present the preferred technology for DNA identification of mass disaster victims mainly due to its simplicity, adequate sensitivity, and high discrimination power. The continuous development of commercial autosomal STR multiplexes, including an increasing number of STR loci, has also contributed to a worldwide standardization and validation of this technology. Current PCR-multiplex kits for STR profiling like SGM Plus, Identifiler (Applied Biosystems, Foster City, CA, USA) or PowerPlex 16 (Promega Corporation, Madison, WI, USA) can amplify 9-15 STR loci plus the Amelogenin (AMEL) locus for gender determination in a single PCR reaction with very high discrimination power to evaluate direct matches between victims and personal effects (with reciprocal match probabilities that vastly exceed the entire human population) and also with appropriate

discrimination power (with certain limitations) to study potential kinship associations between victims and their close ascendant or descendant relatives.

As can be seen from Table 2, autosomal STR profiling has been adopted since 1993, when a quadruplex STR system of four simple STRs was validated for use in a mass disaster case in combination with other DNA markers (3,4). In recent years, the use of the CODIS (Combined DNA Index System) core set of 13 STR by using two separate multiplexes (Profiler Plus and Cofiler) (Applied Biosystems), or the application of new commercial STR multiplexes that amplify 15 STR loci plus AMEL like Identifiler (Applied Biosystems) and PowerPlex 16 (Promega) has been the gold standard for DNA identification analysis in mass fatality incidents with very efficient results.

Unfortunately, severely degraded DNA samples could contain only very short DNA template molecules (under 150 bp) making conventional STR typing (150-400 bp) unsuccessful. This has been one of the technical challenges of the WTC disaster, which resulted in the development of new PCR typing strategies by targeting very short DNA sequences. One example is the development of different Mini-STR multiplexes based on redesigned primers to obtain shorter amplicons

(14,23), which effectively increased the success rate to obtain STR typing results from a proportion of WTC remains where conventional STR typing failed (11). A high throughput single nucleotide polymorphism (SNP) typing strategy for very short amplicons (average 69 bp) has also been applied by Orchid-Gene Screen (Dallas, TX, USA) to target a panel of 70 Bi-allelic autosomal SNP markers from the most severely degraded WTC remains. This strategy of SNP typing, known as SNPstream ultra high throughput (UHT) genotyping system, utilizes multiplexed PCR in conjunction with SNP-IT (Orchid) single base extension technology (24). Mitochondrial DNA sequencing of hypervariable regions 1 and 2 (HV1 and HV2) was also used to obtain complementary DNA data from WTC remains (25).

High Throughput Analysis and Laboratory Automation

A general challenge in mass disaster cases is the need to deal with large-scale DNA sample analysis, making it necessary to develop high-throughput strategies to decrease the cost and turnaround time for DNA analysis of both victim and reference samples. Automation and robotic implementation of some steps of the DNA analysis procedure have been applied for the analysis of WTC remains and reference samples.

Table 2. DNA markers and bioinformatic tools used for pair-wise comparisons and likelihood ratio (LR) calculation on mass disaster cases

Mass disaster case	Number of pair-wise comparisons (remains versus references)	DNA markers and technology	Bioinformatic tools (pair-wise comparisons and LR calculation)	References
Waco disaster/Waco (Texas, USA), April 19, 1993	~ 3.000 (73/42)	STR quadruplex system (Additional analysis with: AmpType HLA DQ-Alpha, AmpType Polymarker, D1S80, and mtDNA)	not included	(3,4)
Spitsbergen aircraft accident/ Spitsbergen (Norway), August 1996	~ 36.000 (257/182)	3 STR plus AMEL and 5 VNTR markers	Excel spreadsheet, Pater software	(6)
Taoyuan Airbus crash accident/ Taoyuan (Taiwan), February 16, 1998	~ 140.000 (685/201)	Profiler Kit (9 STR + AMEL), AmpType HLA DQ-Alpha, AmpType Polymarker	Excel spreadsheet, Lab-developed software for PI calculation	(7)
Swissair Flight 111 accident/ Atlantic Ocean off Canada's coastline, September 2, 1998	~ 500.000 (1277/397)	Profiler Plus+Cofiler, (13 STR+AMEL)	Mass Disaster Kinship Analysis Program (MDKAP)	(9)
World Trade Center disaster/New York (USA), September 11, 2001*	~ 220.000.000 (~20.000/~11.000)	Profiler Plus+Cofiler (13 STR+AMEL) (Additional analysis with: miniSTRs, and Short amplicons for the analysis of autosomal and mitochondrial SNPs)	CODIS, DNA-View, MDKAP, MFISys	(10-12,27,28)
Yakovlev-42 air crash accident/ Trabzon (Turkey), May 26, 2003	~ 8.000 (85/98)	Identifiler (15 STR+AMEL) (Additional analysis with: Powerplex16)	LIMS-GPC module PATPCR software	this study
Madrid train bombing case/ Madrid (Spain), March 11, 2004	~ 20.000 (220/98)	Identifiler (15 STR+AMEL) (Additional analysis with: Powerplex16 and mtDNA analysis in one case)	CODIS, LIMS-GPC module, PATPCR software, CEPOL software	this study

*In spite of all tremendous efforts, around 46% of recovered WTC physical remains remain unidentified (9,194 of the 19,963 collected remains). This low success rate reflects the extremely challenging environmental conditions (temperatures of more than 1,000°C) to which an important part of the WTC remains were exposed for very long time (11). At present time, about 58% of victims (1,591 out of 2,749 missing persons) were identified (Prinz M, personal communication). Unidentified remains were dried and vacuum-sealed to be entombed at the future site of the WTC memorial, favoring the possibility of future re-analysis if scientific progress allows it.

Bode Technology, for instance, developed a high throughput DNA extraction and STR analysis for skeletal remains that allowed processing of more than 250 bone samples per day using a 96-well format for DNA extraction, DNA quantification, and STR analysis (11). The procedure is combined with the use of a LIMS system that included bar-coding of samples throughout the whole testing process to maintain the chain of custody.

A review of the implementation of a LIMS system (including a portable LIMS version) at the Armed Forces DNA Identification Laboratory (AFDIL) in conjunction with high throughput DNA analysis strategies and bioinformatic tools for searching and managing large-scale mitochondrial DNA sequence data from degraded skeletal remains has been recently published (26).

DNA Database Searching and Match Significance

Algorithms for Database Searching and Searching Capabilities

Pair-wise comparisons of DNA profiles in mass fatality incidents will require the bioinformatic capability to search (all-against-all) large STR and SNP databases, using at least two different searching algorithms for autosomal markers to look for: (a) perfect match: number of loci at which both alleles were found to match, a number which is expected to be equal to the number of loci analyzed among fragments of the same body or between a victim and a direct reference, and (b) allele sharing by kinship: number of loci at which at least one allele was found to match, a number which is expected to be equal to the number of loci analyzed for parent-child relationships. The software should also have the capability to rank the significance of the DNA match.

In old mass fatality cases, DNA pair-wise comparisons were made by using lab-developed solutions, such as Excel spreadsheets.

A specific software program, known as Mass Disaster Kinship Analysis Program (MDKAP), with the above mentioned features, was developed and first used to assist with the victim identification initiative that followed the Swissair Flight 111 disaster (9). The MDKAP program was rebuilt with enhanced functionality for use in the identification of the missing in the WTC disaster, including the ability to collapse profiles derived from the re-

mains to a reduced number of consensus profiles, ability to assemble overlapping partial profiles, and calculation of likelihood ratios for each pair-wise comparison at various relationships, such as parent-child, sibling, or half-sibling (27).

A new bioinformatic tool known as M-FYSis for Mass Fatality Identification System (28) was also developed by Gene Codes Corporation to assist database searching of DNA profiles in the WTC disaster which, apart from STR data, also manages mtDNA and SNP data, and groups and collapses data from fragmented remains, in order to track samples among collaborating laboratories, and to collect meta-data for administrative review of reference samples. In the Madrid terrorist attack case, the CODIS database was used to compare 220 body remains against 98 reference samples, including 67 samples from relatives, representing 40 family groups and 27 antemortem direct references. In the Yakolev-42 aircraft disaster, we used a GPC (Genetic Profile Comparisons) module integrated in our SQL-LIMS system (Applied Biosystems) to perform "victims against victims" and "victims against references" DNA profile searching. The system allowed us to import different allele data formats on different indexes defined by scenario and sample categories. Each search retrieved a list of matching profiles sorted according to both concordance (number of loci with perfect match) and kinship (number of loci with at least one allele sharing) indexes (Fig. 1). Likelihood ratio calculation on compatible pairs (and trios whenever possible) was performed by a lab-developed software tool (PATPCR-v 2.52) and included the evaluation of three types of parent-child relationships: (a) defective paternity pairs (just one parental), (b) conventional paternity trios, and (c) paternity plus maternity cases.

Match Significance and Interpretation Criteria

A general characteristic of mass disasters is the high number of pair wise comparisons – from thousand to millions, or even more in the Tsunami catastrophe, which have to be carried out to correlate victims to direct and/or family references. Under these circumstances, a significant number of fortuitous hits could occur between non-related samples (12).

The significance of a perfect match for 13 or more STR loci between two body fragments or between a victim and a direct reference (per-

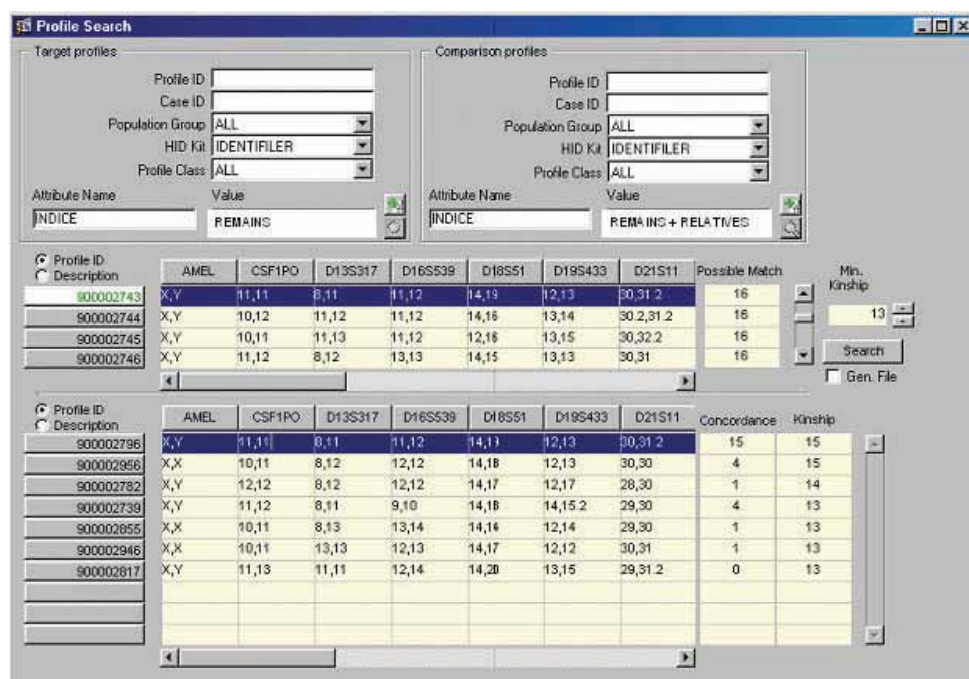


Figure 1. Screen shot of the GPC (Genetic Profile Comparison) module of the SQL-LIMS system (Applied Biosystems). In the central part of the panel a set of victim's target profiles (Identifiler) is shown, logged by selecting the appropriated index (remains) that were searched one by one against all the comparison profiles (Remains + References) from the same disaster. A minimum of 13 markers was chosen (Min. Kinship) for pair-wise comparisons. In this example the highlighted victim profile retrieved a list matching profiles sorted according to the concordance and the kinship indices.

sonal item or antemortem biological specimen) is generally enough to declare identification, offering LR results above 10^9 . This corresponded to a posterior probability higher than 99.9% of correctly identifying all victims in a mass disaster scenario involving 1,000 victims (12). The possibility of reaching this threshold of probability when perfect matches are considered depends on the state of DNA degradation, which can give rise to partial DNA profiles with decreased discrimination power.

On the other hand, the significance of a genetic compatibility (allele sharing) between a sample of the victim's index and a sample of the relative's index can be challenged in mass disaster cases by the incidence of fortuitous hits (false positives) even in the case of parent-child relationships. Indeed, the incidence of false positives is proportional to the number of pair-wise comparisons performed. It can be a problem to distinguish the true hits from the false ones (12,29), even in the case of disasters of lower intensity, and contrary to theoretical predictions (30). In Table 3, we

present empirical data on the number of fortuitous hits obtained after 8,000 pair-wise comparisons performed between 85 remains, representing 74 bodies, and 98 relatives, representing 56 family groups, of the Yakolev-42 aircraft accident. As can be seen, a significant number of false positive pair-wise comparisons (80 false hits, representing 1% of all pair-wise comparisons) were observed when only 13 STR loci (not necessarily the CODIS set) were considered. This number of fortuitous hits was reduced when the number of markers was increased and just one false positive hit was observed when 15 STR loci (Identifiler) were considered. LR calculations for these fortuitous hits ranged from 7 to 23,758 (Table 3). A similar incidence of fortuitous hits was observed in the Madrid terrorist attack case (data not shown) where profiling of pentameric STRs (Penta D and Penta E) and mtDNA sequencing was necessary to confirm the identification based on Identifiler STR data of a body compatible with two members of different family groups. These results clearly showed the need to increase both the number of STR loci and

Table 3. Incidence of fortuitous hits during DNA pair-wise comparisons of the Yakolev-42 air crash accident

Kinship Index*	Number of fortuitous hits	Likelihood ratio (mean, range)
15	1	1448
14	16	2288 (27-13421)
13	80	1109 (7-23758)

*Number of loci at which at least one allele is shared.

the threshold of interpretation (minimum LR or posterior odd), when only one parental relative was available for testing, to distinguish true from false identifications.

A higher incidence of fortuitous hits could be expected when other genetic relationships are considered (full sibling, half sibling). Enlargement of population data sets are needed to evaluate the real significance of mtDNA and Y-STR haploid data. The development of interpretation guidelines for biostatistical evaluation of joint autosomal and haploid DNA data is also desirable.

New Challenges: South Asian Tsunami Perspective

At the moment of writing this review, the information on the Tsunami identification efforts is very scarce and perhaps not very trustworthy. However, it has been anticipated that even if we consider the Tsunami as a group of smaller disasters, affecting 12 different countries (with Indonesia, Sri Lanka, India and Thailand as the worst affected) and producing victims from about 30 countries, the sheer scale of this unpredicted natural disaster, with more than 200,000 victims, is one of the principal challenges to face (31). Furthermore, among the victims, a high number of relatives is expected to be found, as well as entire families that died without any family reference to be compared with. The additional complication of reduced availability of direct reference samples (personal items destroyed by the Tsunami) should also be considered. The situation can also be further complicated by the rate and speed of body recovery from the sea, affecting DNA integrity in some cases. An undetermined proportion of bodies has not yet been discovered and may never be (32).

All these challenges require an approach to the identification process of the Tsunami victims as an integral forensic science identification effort, based not only upon DNA data but also

on forensic anthropology, fingerprints, odontology, radiology. Experience from the previous cases also teaches us about the importance of scientific collaboration and coordination using international standardized DNA technology. This should be facilitated by the use of recent developments in laboratory information management systems and DNA database searching bioinformatic tools.

The massive efforts needed to identify the victims are poised to become the most intensive forensics investigations in history. Indeed, apart from local resources, Interpol DVI teams from more than 20 countries are engaged to provide qualified practical assistance and a Crisis Management Support Group (CMSG) was formed for the co-ordination of the international response to the Tsunami (19). Also, many countries have offered technical resources for DNA typing (33). The amount of economic and technological resources, including DNA technology, which can be used for body identification in the context of the whole Tsunami relief operation is a crucial factor to guarantee an acceptable success with this identification challenge.

Acknowledgements

We thank Juan Antonio Luque from the Institute of Toxicology and Forensic Science in Barcelona for allowing us to use PATPCR software and Mecki Prinz from the NY City Office of the Chief Medical Examiner for last statistical data on the WTC disaster and very helpful comments. We also acknowledge Juan Miguel Monje, medical examiner of the Audiencia Nacional in Madrid, for the coordination of Yakolev-42 body exhumations for DNA analysis.

References

- 1 Mass Fatality Incidents: A Guide for Human Forensic Identification. NCJ 199758, June 2005, Special Report, by National Institute of Justice. Available from: <http://www.ncjrs.org/pdffiles1/nij/199758.pdf>. Accessed: April 30, 2005.
- 2 Ludes B, Tracqui A, Pfitzinger H, Kintz P, Levy F, Disteldorf M, et al. Medico-legal investigations of the Airbus, A320 crash upon Mount Ste-Odile, France. *J Forensic Sci.* 1994;39:1147-52.
- 3 Clayton TM, Whitaker JP, Maguire CN. Identification of bodies from the scene of a mass disaster using DNA amplification of short tandem repeat (STR) loci. *Forensic Sci Int.* 1995;76:7-15.
- 4 Clayton TM, Whitaker JP, Fisher DL, Lee DA, Holland MM, Weedn VW, et al. Further validation of a quadruplex STR DNA typing system: a collaborative effort to identify victims of a mass disaster. *Forensic Sci Int.* 1995;76:17-25.
- 5 Corach D, Sala A, Penacino G, Sotelo A. Mass disasters: rapid molecular screening of human remains by means

- of short tandem repeats typing. Electrophoresis. 1995; 16:1617-23.
- 6 Olaisen B, Stenersen M, Mevag B. Identification by DNA analysis of the victims of the August 1996 Spitsbergen civil aircraft disaster. *Nat Genet.* 1997;15: 402-5.
 - 7 Hsu CM, Huang NE, Tsai LC, Kao LG, Chao CH, Linacre A, et al. Identification of victims of the 1998 Taoyuan Airbus crash accident using DNA analysis. *Int J Legal Med.* 1999;113:43-6.
 - 8 Goodwin W, Linacre A, Vanezis P. The use of mitochondrial DNA and short tandem repeat typing in the identification of air crash victims. *Electrophoresis.* 1999;20:1707-11.
 - 9 Leclair B, Fregeau CJ, Bowen KL, Fournay RM. Enhanced kinship analysis and STR-based DNA typing for human identification in mass fatality incidents: the Swissair flight 111 disaster. *J Forensic Sci.* 2004;49: 939-53.
 - 10 Budimlija ZM, Prinz MK, Zelson-Mundorff A, Wiersema J, Bartelink E, MacKinnon G, et al. World Trade Center human identification project: experiences with individual body identification cases. *Croat Med J.* 2003;44:259-63.
 - 11 Holland MM, Cave CA, Holland CA, Bille TW. Development of a quality, high throughput DNA analysis procedure for skeletal samples to assist with the identification of victims from the World Trade Center attacks. *Croat Med J.* 2003;44:264-72.
 - 12 Brenner CH, Weir BS. Issues and strategies in the DNA identification of World Trade Center victims. *Theor Popul Biol.* 2003;63:173-8.
 - 13 Meyer HJ. The Kaprun cable car fire disaster – aspects of forensic organisation following a mass fatality with 155 victims. *Forensic Sci Int.* 2003;138:1-7.
 - 14 Butler JM, Shen Y, McCord BR. The development of reduced size STR amplicons as tools for analysis of degraded DNA. *J Forensic Sci.* 2003;48:1054-64.
 - 15 Hennessey M. World trade center DNA identifications: the administrative review process. Thirteenth International Symposium on Human Identification – 2002. Available from: <http://www.promega.com/geneticidpr oclussymp13proc/content/hennesseyrev1.pdf>. Accessed: April 30, 2005.
 - 16 Disaster Mortuary Operational Response Team (DMORT), National Disaster Medical System, U. S. Department of Health and Human Services Office of Emergency Preparedness. Available from: <http://www.dmort.org>. Accessed: April 30, 2005.
 - 17 Federal Bureau of Investigation. Evidence Response Team. Available from: <http://www.fbi.gov/hq/lab/ert/ ertmain.htm>. Accessed: April 30, 2005.
 - 18 Department of Defense. Armed Forces Institute of Pathology. The Office of the Armed Forces Medical Examiner. Available from: <http://www.afip.org/Departments/ oafme/oafme.html>. Accessed: April 30, 2005.
 - 19 International Police Criminal Organization (Interpol), 200 quai Charles de Gaulle. Available from: <http:// www.interpol.com>. Accessed: April 30, 2005.
 - 20 Grassberger M, Stein C, Hanslik S, Hochmeister M. Evaluation of a novel tagging and tissue preservation system for potential use in forensic sample collection. *Forensic Sci Int.* In press 2005.
 - 21 Smith A. Disaster victim identification of military aircrew, 1945-2002. *Aviat Space Environ Med.* 2003; 74: 1198-200.
 - 22 Schumm JW, Song EY, Burger M, Sangha J. Collecting and processing buccal cell samples. *Progress in Forensic Genetics.* 2004;10:550-552.
 - 23 Coble MD, Butler JM. Characterization of new miniSTR loci to aid analysis of degraded DNA. *J Forensic Sci.* 2005;50:43-53.
 - 24 Budowle B, Planz J, Campbell RS, Eisenberg A. Budowle B, Planz JV, Campbell RS, Eisenberg AJ. Single nucleotide polymorphisms and microarray technology in forensic genetics – development and application to mitochondrial DNA. *Forensic Science Review.* 2004; 16:21-36.
 - 25 Bieschke E, Mar E, Prinz M, Shaler R. Relevance of mtDNA sequence data for the victim identification after the WTC terrorist attack. IV. International Forensic Y–User Workshop – Haploid DNA markers in forensic genetics. Berlin, November 18-20, 2004. Available from: <http://www.ystr.org/index.html>. Accessed: April 30, 2005.
 - 26 Edson SM, Ross JP, Coble MD, Parson TJ, Barritt SM: Naming the dead – confronting the realities of rapid identification of degraded skeletal remains. *Forensic Science Review.* 2004;16:63-90.
 - 27 Leclair B. Large-scale comparative genotyping and kinship analysis: evolution in its use for human identification in mass fatality incidents and missing persons databasing. *Progress in Forensic Genetics.* 2004;10:42-4.
 - 28 Cash HD, Hoyle JW, Sutton AJ. Development under extreme conditions: forensic bioinformatics in the wake of the World Trade Center disaster. *Pac Symp Biocomput.* 2003;638-53.
 - 29 Biruš I, Marcikić M, Lauc D, Džijan S, Lauc G. How high should paternity index be for reliable identification of war victims by DNA typing? *Croat Med J.* 2003;44: 322-6.
 - 30 Presciuttini S, Toni C, Tempestini E, Verdiani S, Casarino L, Spinetti I, et al. Inferring relationships between pairs of individuals from locus heterozygosities. *BMC Genet.* 2002;3:23.
 - 31 Brenner C. Tsunami victim identification considerations. Available from: <http://dna-view.com/tsunami.htm>. Accessed: April 30, 2005.
 - 32 Fleck F. Tsunami body count is not a ghoulish numbers game. *Bull World Health Organ.* 2005;83:88-9.
 - 33 Ladika S. South Asia tsunami. DNA helps identify missing in the tsunami zone. *Science.* 2005;307:504.

Received: May 2, 2005

Accepted: June 7, 2005

Correspondence to:

Antonio Alonso

National Institute of Toxicology and Forensic Science, Biology Service

Luis Cabrera 9

28002 Madrid, Spain

a.alonso@mju.es