ABOUT OMICS GROUP

OMICS Group International is an amalgamation of Open Access publications and worldwide international science conferences and events. Established in the year 2007 with the sole aim of making the information on Sciences and technology 'Open Access', OMICS Group publishes 400 online open access scholarly journals in all aspects of Science, Engineering, Management and Technology journals. OMICS Group has been instrumental in taking the knowledge on Science & technology to the doorsteps of ordinary men and women. Research Scholars, Students, Libraries, Educational Institutions, Research centers and the industry are main stakeholders that benefitted greatly from this knowledge dissemination. OMICS Group also organizes 300 International conferences annually across the globe, where knowledge transfer takes place through debates, round table discussions, poster presentations,

ABOUT OMICS GROUP CONFERENCES

OMICS Group International is a pioneer and leading science event organizer, which publishes around 400 open access journals and conducts over 300 Medical, Clinical, Engineering, Life Sciences, Phrama scientific conferences all over the globe annually with the support of more than 1000 scientific associations and 30,000 editorial board members and 3.5 million followers to its credit.

OMICS Group has organized 500 conferences, workshops and national symposiums across the major cities including San Francisco, Las Vegas, San Antonio, Omaha, Orlando, Raleigh, Santa Clara, Chicago, Philadelphia, Baltimore, United Kingdom, Valencia, Dubai, Beijing, Hyderabad, Bengaluru and Mumbai.

RUSSIAN FEDERATION
Lomonosov Moscow State University
Faculty of Law,
Department of Criminalistics

Development of the standard of the determination of genetic identity on the basis of the predetermined level of reliability of DNA identification

Professor, Doctor of Medicine
IRINA PEREPECHINA

SAN-ANTONIO - 2014

The rapid development of science makes forensic DNA analysis more and more informative. New identification STR-systems allow to reach significant values even in a single multiplex.

Identification values provided by different STR-systems

STR-system	The provided probability
CODIS 13	5.02×10^{-16}
CODIS Panel A (-DYS391)	1.95×10^{-23}
CODIS Panel A&B (-DYS391)	1.60×10^{-27}
PowerPlex 16	2.82×10^{-19}
PowerPlex 18D	3.47×10^{-22}
PowerPlex 21	6.71×10^{-27}
PowerPlex Fusion (-DYS391)	1.36×10^{-28}

Storts D., 2012

In this regard, the problem of the acceptance of the scientific-based threshold of the genetic identity in order to provide the investigation and the court by the reliable criterion to make a decision about the DNA source becomes more and more urgent and needs to be solved.

Let's see what the investigation and the court would actually like to learn from the DNA examination?

☐ Who left this blood?

☐ Who left this sperm?

☐ Who left this hair? etc.

The questions are very clear and require no less clear-cut answers.

However, forensic scientists are not inclined to give the same simple and clear answers. They have good reasons for it as:

- their conclusions are limited by the strict framework of the mathematical models,

- they have no all the available information on the case,

- they consider the possibility of the influence of a number of factors which complicate inference.

This leads to widespread practice to leave factually the final solution about the DNA source to the court.

Of course, it is a very convenient strategy for the experts especially because it is supported by good reasons. Yet, in spite of the well-known arguments, it appears that this strategy is not faultless.

Why?

The view that it is the court, and only the court is empowered to deal with this issue is a delusion:

1) In spite of the fact that the court has all the information relating to the criminal case, judges, as well as forensic scientists, have no absolute criteria for the evaluation of probabilistic values. So, the redirection of the problem of the interpretation to the court does not allow to solve it fundamentally.

2) The evaluation of the DNA data is very specific and requires solid competence in molecular and population genetics as well as in mathematics, which the judges themselves do not have, and so they are dependant on the opinions of the advisers.

3) Following such a strategy DNA data are not independent of the other evidence within the case, that is particularly important in cases where any of the evidence was compromised (recall the "Innocence Project").

The concept

- The proposed concept aims to avoid subjectivity in deciding on genetic identity.

- It is to adopt a conventional scientifically-based threshold which determines the sufficiency of the derived genetic information to establish identity and declare it as a standard common for forensic experts and the court.

The concept

- The ***key point*** of the concept is that the choice of the criterion in question is proposed to be made on the basis of the ***adopted decision*** on what level of reliability of identification ***is acceptable to the domestic judicial system***.

- The standard is to be the same for all departments performing DNA analysis.

The concept

- The standard should be conservative and allow drawing a conclusion ***regardless of the factual background of the case***.

- In cases where such a threshold is not reached, the expert will do the probabilistic inference.

- Also possible the regulation of conditions that allow to reduce the threshold in cases where the circumstances of the case undoubtedly suggest the possibility of using less conservative value.

□ The choice of the standard should also consider the differences in the identification mechanisms in expert examination and in operating databases, especially very large ones ("the phenomenon of birthday").

Who should make a choice?

The standard is to be adopted by an authorized collegial body comprising:

1) **scientific experts**:

☐ forensic DNA scientists,

☐ molecular and population geneticists,

☐ mathematicians

- calculate the risk of error depending on the probability values and present the scale of risks in such a way that non-expert is able to understand their degree;

2) **specialists in the humanities**:

 - evaluate ethical and social aspects of these risks: the choice of the standard occurs under conditions of the dilemma between I and II type errors which are not of the same value for both the individual and society;
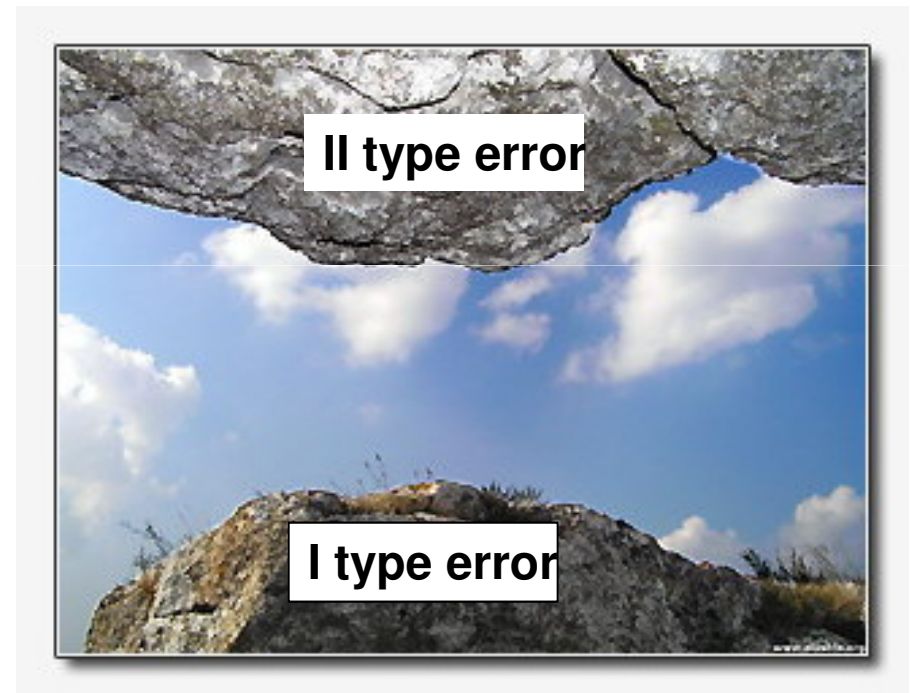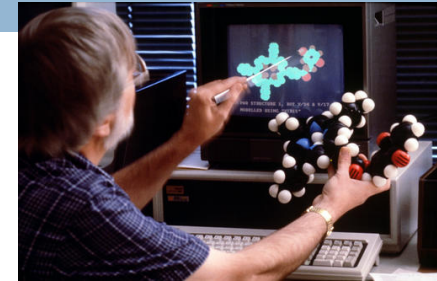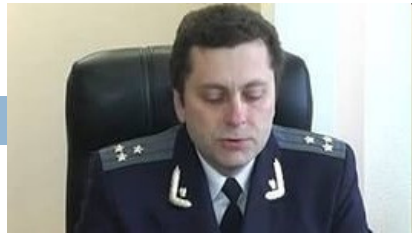
3) **lawyers** - representatives of:

- ☐ law enforcement,
- ☐ judicial community,
- ☐ prosecution,
- ☐ advocacy

- assess the scale of the risks in legal and ethical framework and decide which level is most admissible for the justice system.
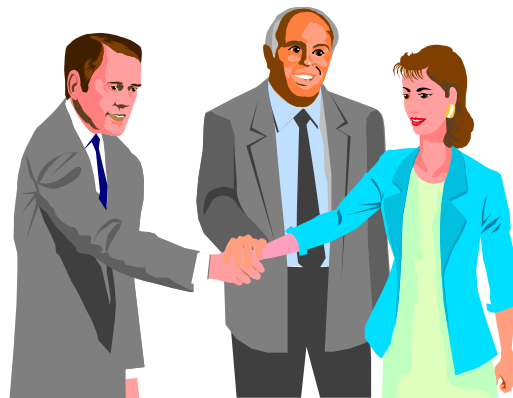
a prosecutor

a judge

a forensic scientist, a geneticist

a lawer

A conventional criterion

a specialist in the humanities

a mathematician

- Such a board will allow to take into account all the significant aspects in the development of the standard, while the empowering it by the regulatory authority will ensure the effective recognition of the standard by the law enforcement system.

- The concept is supported by the elaborated detailed mathematical mechanism for its implementation.

The technology of the choice of the standard of DNA identification

□ The technology of the choice of the standard should enable the participants of the board who are not professionals in probability theory to understand the content of the figures and really appreciate the identification value of the variants of the regulated parameter.

□ It may be based on the evaluation of a series of equivalent probability values which consists in following:

□ The base calculated value - match probability, as well as its derivative value - LR, being considered by itself, carry only limited information on these values in terms of the reliability of establishing the identity.

□ However, if we express this values through a series of mathematical equivalents describing its content and evaluate all the data in the complex, it is possible to get extra information.

- The study of their range can be useful for the assessing the risk of the identification error as the equivalent probability values describe different aspects of the same statistical model.

- The equivalents may include new parameters that can be set.

- Some of the possible parameters are shown in the following table.

Equivalent probability values to be assessed

P	Frequency	LR	G (upper) and Q (low), calculated for N				M, %	К
			1 mln	10 mln	200 mln	7 bln		
10⁻⁶	1:1 mln	1,000,000	0,3679	→0	→0	→0	99	100
			0,6321	→1	→1	→1		
10⁻⁷	1:10 mln	10,000,000	0,9048	0,3679	→0	→0	99,9	1,000
			0,0952	0,6321	→1	→1		
10⁻⁸	1:100 mln	100,000,000	0,9900	0,9048	0,1353	→0	99,99	10,000
			0,0100	0,0952	0,8647	→1		
10⁻⁹	1:1 bln	1,000,000,000	0,9990	0,9900	0,8187	0,0009	99,999	100,000
			0,0010	0,0100	0,1813	0,9991		
10⁻¹⁰	1:10 bln	10,000,000,000	0,9999	0,9990	0,9802	0,4966	99,9999	1,000,000
			0,0001	0,0010	0,0198	0,5034		
10⁻¹¹	1:100 bln	100,000,000,000	0,9999	0,9999	0,9980	0,9324	99,99999	10,000,000
			0,0001	0,0001	0,0020	0,0676		
10⁻¹²	1:1,000 bln	1,000,000,000,000	0,9999	0,9999	0,9998	0,9930	99,999999	100,000,000
			0,0001	0,0001	0,0002	0,0070		
10⁻¹³	1:10,000 bln	10,000,000,000,000	0,9999	0,9999	0,9999	0,9993	99,9999999	1,000,000,000
			0,0001	0,0001	0,0001	0,0007		
10⁻¹⁴	1:100,000 bln	100,000,000,000,000	0,9999	0,9999	0,9999	0,9999	99,99999999	10,000,000,000
			0,0001	0,0001	0,0001	0,0001		

Equivalent probability values

- Match probability (P) - is the probability of having matching genotype between two randomly chosen individuals. It is also can be defined as the estimated frequency at which a particular DNA profile would be expected to occur in a population.

- Likelihood ratio (LR) - how likely it is that the observed DNA profile came from the true perpetrator rather than from a random individual.

  LR = 1/P (in the simplest case)

- G – probability of uniqueness of the observed DNA profile in population N:
  $G = (1 - P)^N$

Equivalent probability values

- Q – probability of the revealing in N at least one more object with the same profile:
  $$Q = 1 - (1 - P)^N$$

- The probability (M) that at least in A identifications [in the table A=10,000] implemented with particular P or less there will not be a mistake.

- Minimum number of cases (K) in which with the safety of 99,99% the reliable identification is provided by particular P.

- Other probabilistic values.

G and Q may be calculated in relation to populations of different sizes – e.g.:

- 1 mln (a city),

- 10 mln (a megacity),

- 200 mln (a big country, a region),

- 7 bln (earth).

The example of such calculations applied to one of the values of P (10^{-8}) is shown at the next slide.

The conventionality of such a graduation should, of course, be taken into account.

for $P = 10^{-8}$

N		G	Q
1 million		0,9900	0,0100
10 million		0,9048	0,0952
200 million		0,1353	0,8647
7 billion		$\rightarrow 0$	$\rightarrow 1$

The assessment of the practical reliability of the identification for the particular P

P = 10^{-8} with safety of *99,99%* [should be set] ensures reliability at least 10 000 identifications,

P = 10^{-9} with safety of *99,99%* ensures reliability at least 100 000 identifications,

P = 10^{-10} with safety of *99,99%* ensures reliability at least 1 000 000 identifications,

and so on.

The number of identifications may be converted into the number of years - the period of the reliable implementation of the identification with the particular or less P.

Conclusion

□ It is proposed to develop the standard of DNA identification on the basis of the predetermined level of reliability which to be adopted by both the scientific community and domestic justice system. This will allow to avoid subjectivity in deciding on the threshold of the genetic identity and determination of the DNA source.

□ As a technology of the choice of the standard an analysis of the equivalent probability values which describe different aspects of the same statistical model is proposed. Such a strategy allows to set any parameters.

- The evaluation of the range of equivalent values can be useful for the investigator and the court also in application to a particular case.

- The outlined approach may be applicable not only to DNA analysis, but also to other kinds of identification.

Acknowledgments

To Prof. Sergey A. Grishechkin for the collaboration.

Thank you for your attention!

smi-100@mail.ru

Let Us Meet Again

We welcome you all to our future conferences of OMICS Group International

Please Visit:

www.omicsgroup.com

www.conferenceseries.com