Population Structure in *Naegleria fowleri* as Revealed by Microsatellite Markers

Bénédicte Coupat-Goutaland[1], Estelle Régoudis[1], Matthieu Besseyrias[2], Angélique Mularoni[3], Marie Binet[4], Pascaline Herbelin[4], Michel Pélandakis[1] *

1 Univ Lyon, Université Lyon 1, CNRS UMR 5240 Microbiology Adaptation and Pathogenesis, Villeurbanne, France, 2 Univ Lyon, Université Lyon 1, Villeurbanne, France, 3 Univ Lyon, Université Lyon 1, ISPB EA 4446 Bioactive Molecules and Medicinal Chemistry, Lyon, France, 4 EDF Research and Development, Laboratoire National d'Hydraulique et Environnement, Chatou, France

* michel.pelandakis@univ-lyon1.fr

Abstract

Naegleria sp. is a free living amoeba belonging to the Heterolobosea class. Over 40 species of *Naegleria* were identified and recovered worldwide in different habitats such as swimming pools, freshwater lakes, soil or dust. Among them, *N. fowleri*, is a human pathogen responsible for primary amoeboic meningoencephalitis (PAM). Around 300 cases were reported in 40 years worldwide but PAM is a fatal disease of the central nervous system with only 5% survival of infected patients. Since both pathogenic and non pathogenic species were encountered in the environment, detection and dispersal mode are crucial points in the fight against this pathogenic agent. Previous studies on identification and genotyping of *N. fowleri* strains were focused on RAPD analysis and on ITS sequencing and identified 5 variants: euro-american, south pacific, widespread, cattenom and chooz. Microsatellites are powerful markers in population genetics with broad spectrum of applications (such as paternity test, fingerprinting, genetic mapping or genetic structure analysis). They are characterized by a high degree of length polymorphism. The aim of this study was to genotype *N. fowleri* strains using microsatellites markers in order to track this population and to better understand its evolution. Six microsatellite loci and 47 strains from different geographical origins were used for this analysis. The microsatellite markers revealed a level of discrimination higher than any other marker used until now, enabling the identification of seven genetic groups, included in the five main genetic groups based on the previous RAPD and ITS analyses. This analysis also allowed us to go further in identifying private alleles highlighting intra-group variability. A better identification of the *N. fowleri* isolates could be done with this type of analysis and could allow a better tracking of the clinical and environmental *N. fowleri* strains.

Introduction

The free living-amoeboflagellate *Naegleria fowleri* is the causative agent of primary amoebic meningoencephalitis (PAM) in humans, a rare but rapidly fatal disease of the central nervous

system. The species is thermophilic and is found in soil and aquatic environments. Most of the fatal infections due to *N. fowleri* happen in young people exposed to warm water in ponds, swimming pools and lakes [1]. Two other *Naegleria* species are thermophilic and share the same habitat as *N. fowleri*: *Naegleria australiensis* which is pathogenic in mice and *Naegleria lovaniensis* which is harmless. Another species, *Naegleria italica* was reported to be pathogenic in mice but was rarely isolated [2,3].

Due to the lack of morphological characters within the genus *Naegleria*, the identification of *N. fowleri* was carried out by various techniques using immunological and molecular tools with specific PCR primers [4–11]. Molecular markers allowed to investigate intra-species variability which was first observed by using RFLP [12,13], and then by using RAPD markers displaying a higher level of diversity [14,15]. Therefore, the RAPD markers evidenced five main variants, namely South Pacific (SP), Chooz (CHO), Cattenom (CAT), Widespread (WP), and Euro-American (EA) which are not always correlated with geographical origin. The existence of RAPD variants was confirmed by the analysis of the 5.8S rRNA gene and the internal transcribed spacers (ITS), except for the CAT variant that could not be differentiated, since it is identical to the SP variant [6,16–19].

ITS analyses remain very useful for typing, but are not sufficiently informative for population genetic studies, or for a thorough tracking of *N. fowleri* strains. In this respect, genetic markers such as microsatellites are more suitable. Microsatellites are tandemly repeated short motifs of 1 to 6 base pairs and are widely distributed in the prokaryotic and eukaryotic genomes. They display variations in the number of repeats and exhibit a high level of polymorphism. Microsatellites are powerful markers for population genetics with broad spectrum of applications and thus can be used for biodiversity, fingerprinting, population studies, genetic mapping within populations, clonal structure and evolutionary history [20,21]. Therefore, microsatellites are of great interest in the study of eukaryotic genetic diversity, in particular for protozoa such as *Plasmodium* sp. [22,23], *Trypanosoma cruzi* [24], *Cryptosporidium parvum* [25] or *Leishmania* sp. [26,27].

In this study, we describe and examine six microsatellites in the different variants of *N. fowleri* which were initially identified by RAPD and ITS. Our objective was to improve the understanding of population structure of this pathogenic species, to enhance the typing for a better tracking, and to better understand their dispersal. The analysis of *N. fowleri* strains with these microsatellites provided valuable new information about their population structure.

Materials and Methods

Sampling

The 47 *N. fowleri* strains and DNA analyzed in this study are shown in Table 1. The French strains, which were isolated between 2009 and 2012, originated from different geographical sites. The strains with identical first letters in the nomenclature came from the same French site. The genotype of the strains were identified and confirmed by sequencing the ITS region (Table 1). Representatives of the variants were cloned.

Culture and DNA extraction

All strains were cultured on non nutrient agar (NNA) plates with *E. coli* cells layers and incubated at 30°C for 4 to 7 days. Isolated strains were recovered from NNA plates and scrapped in order to extract DNA. Extraction was performed with Nucleospin Tissue kit (Macherey Nagel, Hoerdt, France) according to the manufacturer's protocol. DNA quantity and quality, for both RNA and protein contaminations, were controlled with a Nanodrop 1000 spectrophotometer. For PCR analysis, all DNAs were diluted to a final concentration of 10 ng/μL.

Table 1. DNA and strains used with geographical origin.

Name	Geographical origin	RAPD/ITS variant[a]	Source[b]	Type[c]	Year of isolation	Reference
Kul	Belgium	EA/3	H	S	1973	[14]
Moj200	France	EA/3	E	S	1987	[14]
WM	USA	EA/3	H	D	1969	[15]
SW1	USA	EA/3	E	D	1976	[15]
EPA-1911s	USA	EA/3	E	D	-	[6]
Lovell	USA	EA/3	E	D	1974	[14]
NG060	Australia	SP/5	E	D	-	[6]
PAa	New-Zealand	SP/5	E	D	1972	[14]
Northcott	Australia	SP/5	H	D	1971	[14]
CA66	Australia	SP/5	H	D	1966	[14]
PA117	Australia	SP/5	E	D	1972	[15]
MW4U	Australia	SP/5	E	D	1972	[15]
Mst	Australia	SP/5	H	D	1974	[14]
Na420c	France	CAT/5	E	S	1988	[14]
J26(50)45E	Japan	CAT/5	E	D	1990	[15]
PLC-2	Mexico	WP/2	H	D	1990	[15]
Enterprise	USA	WP/2	E	D	1976	[15]
124	USA	WP/2	E	D	1984	[15]
A1	France	EA/3	E	S	2010	This study
A2	France	EA/3	E	S	2010	This study
A3	France	EA/3	E	S	2010	This study
A4	France	EA/3	E	S	2009	This study
B1	France	EA/3	E	S	2011	This study
B2	France	EA/3	E	S	2011	This study
B3	France	EA/3	E	S	2011	This study
B4	France	EA/3	E	S	2012	This study
B5	France	EA/3	E	S	2012	This study
C1	France	WP/2	E	S	2011	This study
C2	France	WP/2	E	S	2011	This study
C3	France	WP/2	E	S	2011	This study
C4	France	WP/2	E	S	2011	This study
C5	France	WP/2	E	S	2011	This study
D1	France	EA/3	E	S	2010	This study
D2	France	EA/3	E	S	2010	This study
E1	France	CAT/5	E	S	2010	This study
E2	France	CHO/4	E	S	2012	This study
E3	France	CHO/4	E	S	2012	This study
F1	France	WP/2	E	S	2010	This study
F2	France	WP/2	E	S	2012	This study
F3	France	WP/2	E	S	2012	This study
F4	France	WP/2	E	S	2012	This study
F5	France	WP/2	E	S	2012	This study
F6	France	WP/2	E	S	2012	This study
G1	France	WP/2	E	S	2010	This study
G2	France	WP/2	E	S	2011	This study
G3	France	WP/2	E	S	2011	This study

(Continued)

Table 1. (*Continued*)

Name	Geographical origin	RAPD/ITS variant[a]	Source[b]	Type[c]	Year of isolation	Reference
G4	France	WP/2	E	S	2011	This study

[a] EA: Euro-american, SP: South pacific; WP: Widespread; CAT; Cattenom, CHO, chooz variants from [15]; the number of the ITS genotype from [19];

[b] H: human or E: environmental isolates;

[c] D: DNA or S: strains.

doi:10.1371/journal.pone.0152434.t001

Microsatellite isolation

Partial genomic librairies. Total DNA was extracted by phenol-chloroform extraction according to the conventional procedure. To avoid the presence of extrachromosomal DNA, total DNA was loaded onto a 0.8% agarose gel (SeaPlaque Agarose, Lonza-Ozyme, Montigny-Le-Bretonneux, France) and run for 22h in TAE buffer at 20V and 16mA. The selected DNA corresponding to chromosomal DNA was digested with Sau3A (Boehringer Mannheim— Roche Diagnostics, Meylan, France). Restriction fragments between approximately 300 to 900 bp were transferred onto DEAE paper (Sanbrook 1989). The fragments were ligated within PUC18 vector (Amersham biosciences—GE healthcare, Glattbrugg, Switzerland) and amplified after transformation into competent XL blue cells (Amersham biosciences—GE healthcare).

Screening and sequencing. Clones were transferred on solid LB medium plates and transferred onto Hybond-N Nylon membrane (Amersham biosciences—GE healthcare). The screening was performed with an equal mix of oligonucleotides $(TC)_{10}$, $(TG)_{10}$, $(CAC)_5CA$, CT $(CCT)_5$, $CT(ATCT)_6$, and $(TGTA)_6TG$, labelled with the DIG oligonucleotide tailing kit (Boehringer Mannheim—Roche Diagnostics). Positive clones were directly analyzed by sequencing (Amersham, Pharmacia-Biotech T7 sequencing kit—GE healthcare). Five of these clones were retained for the analysis.

Microsatellites PCR amplification and analysis with capillary sequencer

The five clones were amplified with different probes labeled with FAM fluorochrome (Table 2). For further analyses, NG42 and NG141 were amplified into 2 separate fragments (NG42-1, NG42-2 and NG141-3, NG141-5, respectively). Amplifications were performed with EconoTaq

Table 2. Microsatellites loci and primers.

Locus	Microsatellite[a]	Allele size range (bp)	Number of alleles	Primers U/L[b]
NG25	$(GT)_8$	110–116	4	AATATTGCTGATGCGAAGGG/AATTGCACCACCAACTCCG
NG42-1	$(CT)_5…(ACC)_7$	220–226	4	GCAGCACTCACTCCTCCTC/CACCTCTCCCTCTTACAACAG
NG42-2	$(GGT)_7…(GGT)_4$	222–252	5	GATTTCCTGCTGGACGATGA/CCCAAGTCGTCCATGATGAGA
NG69	$(TTC)_7…(CAT)_6G(GGT)_6…(CTT)_6$	238–245	3	CAACCTGTTCCCAAGATTTGTAAG/ TATGGATAAGGAAAGAAGTGATACAAG
NG141-3	$(TG)_6…(TG)_5…(TG)_4…(GT)_6…(GT)_6$	162–216	10	CTCAATGTAGAAAATGCTAAT/TGGAATGAATGGAATATACTC
NG141-5	$(GT)_{11}$	153–163	5	GAGAGTATATTCCATTCATTCCA /TTGATCATCCTCATCATTCCACC

[a] repetitions of microsatellite are indicated for KUL strain.

[b] U: Upper primer (5'-3') and L: lower primer 5'-3'.

doi:10.1371/journal.pone.0152434.t002

(Lucigen, Middletown, USA) using an initial denaturation step at 95°C for 5 min, followed by 40 cycles at 95°C for 10s, 57°C (or 50°C) for 20s and 72°c for 20s and a final elongation step at 72°C for 7 min. Fluorescent amplified fragments were analyzed on a ABI 3730 capillary sequencer with a GeneScan 600 LIZ size standard (Applied biosystems—Life Technology, Saint Aubin, France) from the Genomics platform of Nantes (Biogenouest Genomics). Results were analyzed with Peakscanner software (Applied biosystems).

Cloning and sequencing of microsatellite alleles

Different alleles were cloned in pCR2.1 TOPO TA cloning vector (Invitrogen—Life Technologies) and plasmids were extracted with QIAprep Spin Miniprep Kit (Qiagen, Courtaboeuf, France), according to the manufacturer's protocol. Amplifications using M13R and M13F primers were performed with EconoTaq after an initial denaturation step of 95°C for 5 min, followed by 40 cycles of 95°C for 10s, 55°C for 20s and 72°c for 20s and a final elongation of 72°C for 7 min. ABI 3730XL sequencer was used by Genoscreen (Lille, France) to sequence alleles.

Data analyses

Genepop 4.2.2 software [28] was used to calculate allele frequency-based correlation (F_{IS}, F_{ST} and F_{IT}). F_{IS} is the proportion of the total inbreeding within a population due to inbreeding within sub-populations. It ranges between -1 and 1, where a negative value corresponds to an excess of heterozygotes, and a positive value to heterozygote deficiency. $F_{IS} = 0$ indicates Hardy-Weinberg allele proportions. F_{ST} is the proportion of the total inbreeding in a population due to differentiation among sub-populations. Mean F_{ST} estimates over loci in each population were calculated with the FSTAT software (version 2.9.3.2, [29]) using Nei estimators. F_{IT} is the total inbreeding in a population due to both inbreeding within sub-populations, and differentiation among sub-populations. Expected and observed heterozygosity (He and Ho, respectively) were estimated with R software using Adegenet package [30].

Factorial correspondence analysis (FCA) implemented in GENETIX 4.05 software [31] was performed, which places the individuals in a three-dimensional space according to the degree of their allelic state similarities.

Assignment testing was performed by using the GENECLASS 2.0 software package [32]. The Bayesian method of Rannala and Mountain [33] was selected to perform self-assignment tests with algorithm of Paetkau et al. [34] on a simulation with 1000 individuals and alpha = 0.001.

Matrix of genetic distances was performed with the POPULATIONS 1.2.32 software and the Cavalli-Sforza and Edwards model. Phylogenetic networks were inferred from the distance matrix obtained from the microsatellite dataset by using the Neighbor-Net method in SplitsTree 4.13.1 [35].

Ethics Statement

The French sites (A-G) mentioned in Materials & Methods and Table 1 are industrial sites for which the authorizations have been obtained. These sites are secure and their precise location cannot be mentioned. The field studies did not involve endangered or protected species.

Results

Among the partial library, 3500 clones were screened and 30 were positives and were sequenced. A number of clones were excluded because, either no microsatellites were present, or because amplification was not possible, probably due to chimeric alleles. Five clones were

selected for intra-specific analysis and referred to as NG25, NG42, NG69, NG115 and NG141. NG42 and NG141 loci contained different microsatellite repetitions and were divided in two parts for further analysis (NG42-1 and NG42-2 and NG141-3 and NG141-5, respectively). NG115 was found to be monomorphic with one allele at 115 bp for all strains and was not used thereafter. Thus, a total of 6 microsatellites markers were used for this analysis. The locus composition is indicated in Table 2 and shows that most of them are imperfect with repeats.

Microsatellite analyses displayed one or two alleles and can be considered as homozygous or heterozygous. The exception came from NG69 displaying three alleles for most of the strains, due to the presence of a duplication of this locus (Table 3). Therefore, one identical 214 bp allele was present in all samples. It was subsequently cloned and its sequence was identical to the other ones from the NG69 locus. The duplication was confirmed by the blast analysis of the *N. fowleri* genome recently sequenced [36], which indicates that each locus is represented on one contig, except for NG69, which is present on two different contigs.

Genetic diversity

The tested microsatellites revealed a different level of polymorphism in our samples. The NG141-3 was the most variable marker with 10 alleles identified, whereas the other microsatellites produced between 3 to 5 alleles.

The highest number of alleles was 16 and was found for the SP strains whereas strains from CAT and those from EA and WP presented 11, 13 and 12 different alleles, respectively (Table 3).

For the NG141, NG69 and NG42-2 markers, the overall observed heterozygosity per locus was higher than expected and reflected an excess of heterozygosity which was confirmed by the negative F_{IS} values (Table 4). In contrast, the observed heterozygosity per locus was lower than expected for the NG42-1 marker, and displayed a deficiency of heterozygosity. Surprisingly, no heterozygous were found with the NG25 microsatellite, therefore explaining the values of 0 and 1 obtained with H_o and F_{IS}, respectively.

Population structure

The factorial correspondence analysis (FCA) showed that we recovered the 5 variants previously identified, EA, WP, SP, CHO and CAT with two additional ones: NZ and RA, respectively included in the SP and WP groups based on the previous RAPD and ITS analysis (Fig 1). Identified by microsatellite markers, these variants will be named as genetic groups. The GENECLASS program was performed to assess the assignment of the 47 strains to these seven genetic groups. 97.9% of the strains were correctly assigned. Indeed, the SP strain NG060 and the CAT strain JP45E were not assigned. The EA strain SW1 was placed within the WP with a very weak probability (p = 0.016) (S1 Table).

The existence of this structure was also supported by the NeighborNet network (Fig 2). The genetic differentiation between the seven groups was very high with F_{ST} values varying from 0.29 to 0.61 (Table 5).

FCA as well as the NeighborNet network included the three strains SW1, JP45E and NG060 into their respective groups. The FCA showed that the NZ group is very apart from the other ones whereas the SP and EA groups were found to be close. With the NeighborNet network, the SP, NZ and CAT groups are closely related to the EA group. CHO and RA groups are branching with WP (Fig 2). However, the ambiguous splits throughout the network indicate that the relationships within the clusters and groups are not resolved.

Table 3. Allelic combination of the microsatellite markers for the strains examined.

ITS variant	Strain or Origin	NG25				NG42-1			NG42-2					NG69a[a]							NG141-3							NG141-5					Total number of alleles
		110	114	116	220	223	225	226	222	225	226	228	252	214	238	241	245	162	164	167	179	180	183	185	197	212	216	153	155	157	159	163	
Euro-American	Kul		X			X				X				X	X	X					X								X				13
	Moj		X			X				X				X	X	X						X						X					
	D1		X			X				X				X	X	X					X	X							X				
	D2		X			X				X				X	X	X					X				X				X				
	B1-5		X			X				X			X	X	X	X					X	X							X				
	A1-4		X			X				X				X	X	X					X	X							X				
	SW1		X			X				X				X	X	X		X							X				X				
	WM		X			X	X			X				X	X	X					X				X				X				
	EPA1911		X			X				X				X	X	X					X								X				
	Lovell		X			X				X				X	X	X					X				X				X				
South Pacific	NG060					X				X				X	X	X			X											X			16
	PAa	X				X		X						X	X	X				X						X				X			
	PA117		X			X				X				X	X						X	X					X		X	X			
	Northcott		X			X				X				X	X						X	X					X		X	X			
	CA66		X			X				ND	ND			X			ND				X	X					X		X	X			
	Mst		ND				ND			X	X			ND	ND					X						X		X		X			
	MW7U		X			X				X				X	X	X										X	X			X			
Widespread	G1-4			X		X				X		X		X	X	X			X		X	X						X	X				12
	C1-5			X		X				X		X		X	X	X			X		X	X						X	X				
	F1-6		X			X				X		X		X	X	X			X						X			X	X				
	124			X		ND				X		X		X	X	X			X			X			X			ND		ND			
	Enterprise			X		ND				X		X		X	X	X			X			X						ND	ND				
	PLC2			X		X				X		X		X	X	X			X			X						X	X				
Chooz	E2-3		X			X				X				X	X	X							X	X								X	9
Cattenom	E1		X			X				X				X	X	X			X		X		X					X	X				11
	Na420c		X		X	X				X				X		X			X		X		X					X	X				
	JP45E		X			X			X	X				X		X	X		X		X		X						X				

ND: not determined.

[a] Since the 214 pb allele is present in all strains, it was not consider in the analysis of the data.

doi:10.1371/journal.pone.0152434.t003

**Table 4. Genetic characteristics of the microsatellite markers observed in the populations of *N. fowleri*.**

	H$_o$	H$_e$	F$_{ST}$	F$_{IS}$	F$_{IT}$	D$_{ST}$
NG25	0.000	0.4527	0.5782	1.000	1.000	0.309
NG42-1	0.0454	0.0875	0.4174	0.2302	0.5515	0.060
NG42-2	0.5454	0.4384	0.2183	-0.4883	-0.1633	0.061
NG69	0.8666	0.5108	0.0847	-0.7993	-0.6468	0.068
NG141-3	0.8936	0.8100	0.2915	-0.4113	0.0000	0.223
NG141-5	0.6818	0.6242	0.3462	-0.4923	0.0244	0.186
All			0.312	-0.358	0.066	0.151

H$_o$, observed heterozygosity; H$_e$, expected heterozygosity; F$_{ST}$, inbreeding coefficient among individuals within populations; F$_{IS}$, inbreeding coefficient within the population; F$_{IT}$, inbreeding coefficient as the global population; D$_{ST}$, genetic diversity.

doi:10.1371/journal.pone.0152434.t004

Variation within genetic groups

The different groups displayed specific profiles with NG141-3 marker ([Table 3](#)). Every marker was able to identify some genetic groups, but not all. For example, NG42-2 specifically detected the genetic WP and SP groups but EA, CHO and CAT groups shared the same
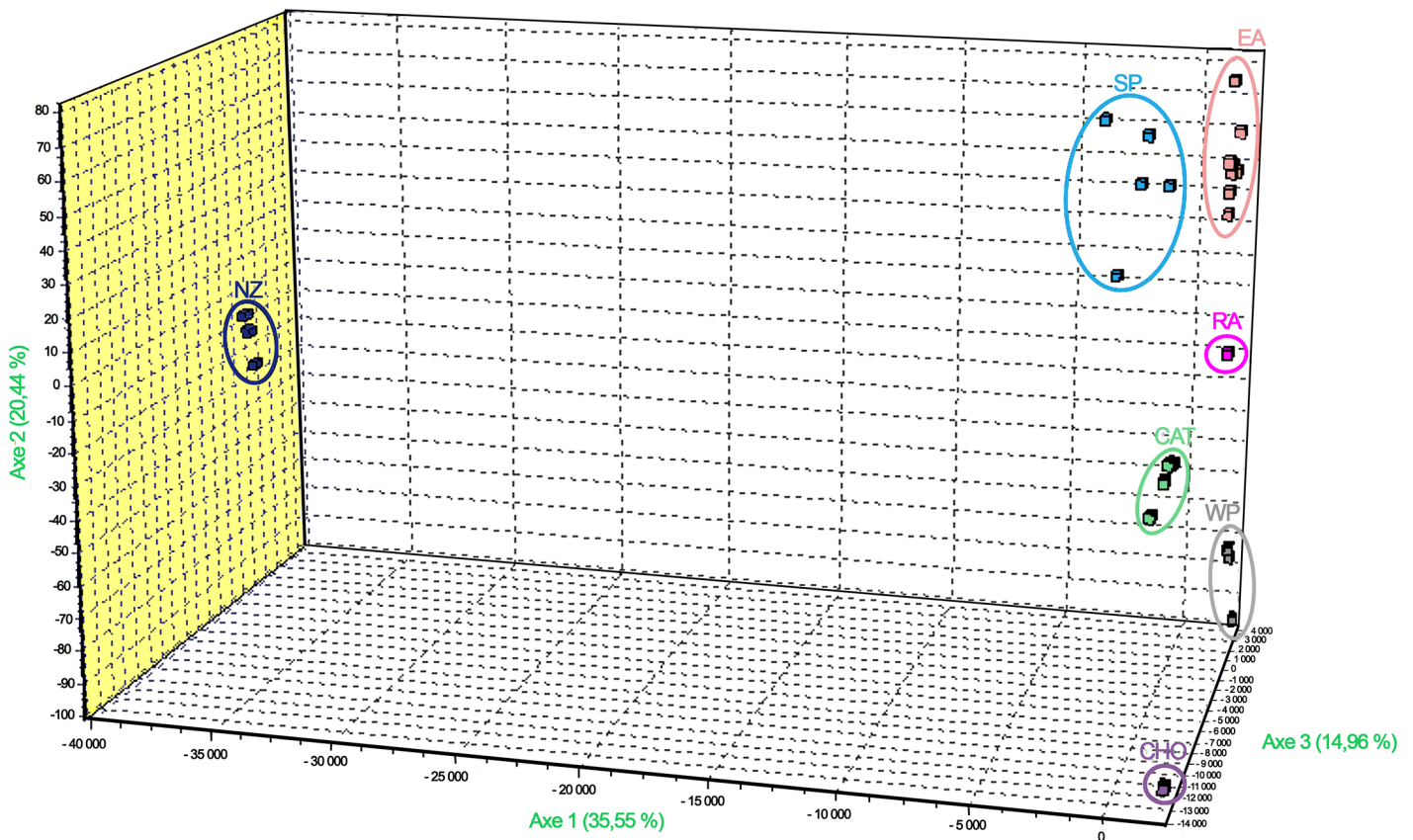
Fig 1. Factorial correspondence analysis (FCA) of *Naegleria fowleri* strains. The strains labelled in blue, orange, pink, green, grey, purple and black correspond to those that were assigned to the SP, EA, RA, CAT, WP, CHO and NZ respectively.

doi:10.1371/journal.pone.0152434.g001
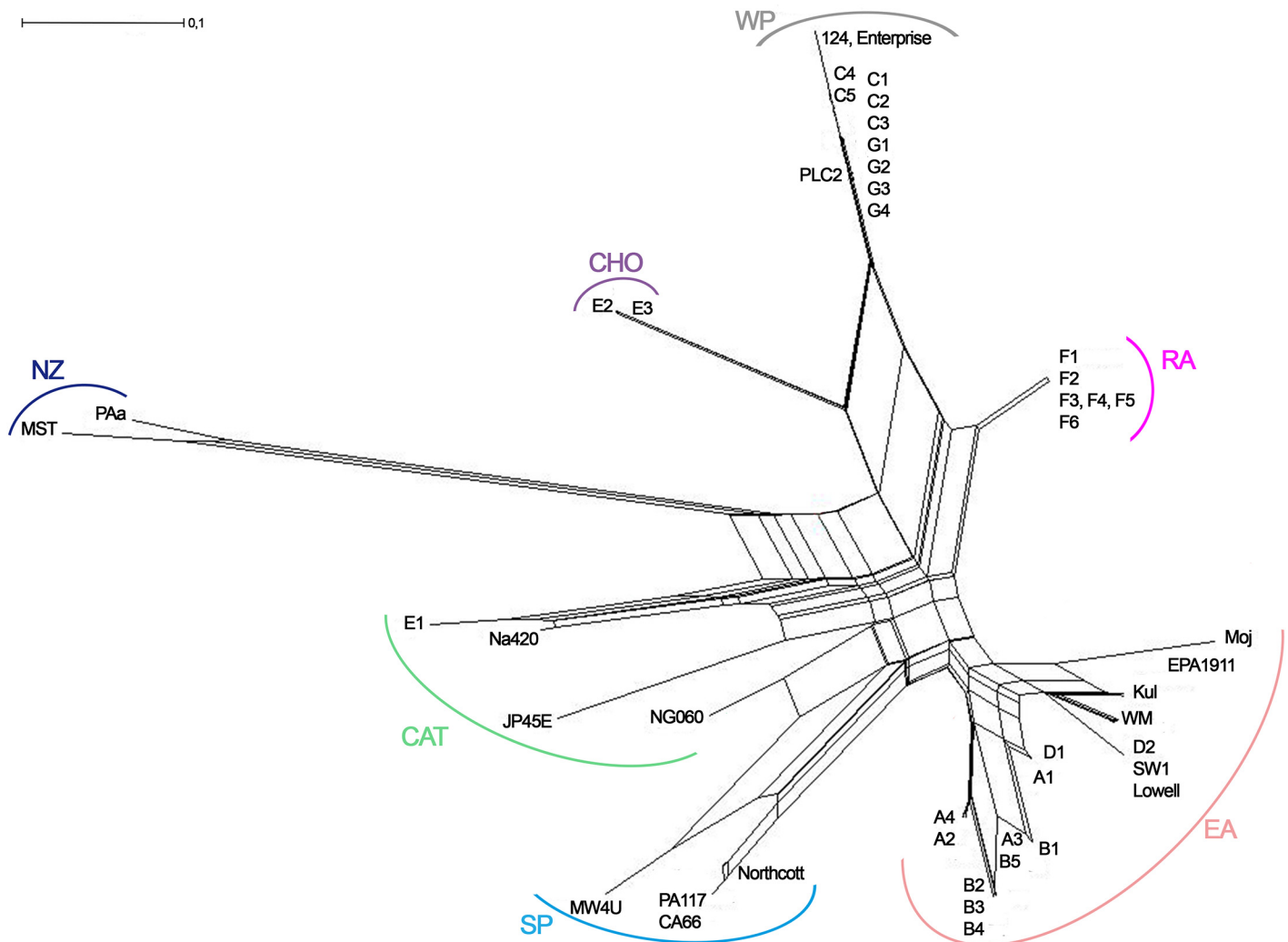
Fig 2. The NeighborNet network of the seven genetic groups of *N. fowleri* identified by the microsatellite markers. The strains labelled in blue, orange, pink, green, grey, purple and black correspond to those that were assigned to the SP, EA, RA, CAT, WP, CHO and NZ respectively.

doi:10.1371/journal.pone.0152434.g002

Table 5. Pairwise comparison of Fst between the seven genetic groups.

	EA	SP	NZ	WP	RA	CHO	CAT
EA	0.0000*						
SP	0.3145*	0.0000					
NZ	0.5956*	0.6117	0.0000				
WP	0.4740*	0.5144*	0.5738	0.0000			
RA	0.2883*	0.3773	0.5746	0.3846*	0.0000		
CHO	0.4674	0.5432	0.6111	0.3415	0.5161	0.0000	
CAT	0.3334*	0.4455*	0.4380	0.4797*	0.2906	0.4423	0.0000

*indicates significant difference between group pairs (p value < 0.05).

doi:10.1371/journal.pone.0152434.t005

alleles. Additionally, differentiation within a given group could be observed. Therefore, the NG141-3 locus generated four different profiles within the EA group. The heterozygous profile (179 bp-180 bp) was only observed in France, the heterozygous profile (162 bp-197 bp) was only observed in the USA, and the two remaining ones (179 pb and 179 bp-197 bp) were both found in USA and France (Table 3). Note that the heterozygous profile 179 bp-197 bp was found only once in a French site among the twenty French strains examined, whereas the heterozygous profile 179 bp-180bp was recovered on the majority of them.

A number of strains also had a specific microsatellite profile with one or several loci. Therefore, the B strains displayed a singular profile with the 252 bp allele at the NG42-2 locus. This private allele was not found elsewhere. The difference of 27 bp between this allele and the other one (225 bp) led us to sequence them. As expected, the result revealed an insertion of 27 nucleotides that is not due to microsatellite variations. This insertion, located at the beginning of the sequence could suggest the mismatch of the upper primer that could therefore explain the weaker intensity peak for the 252 bp allele compared to the 225 bp allele (S1 Fig). Other private alleles could be identified, particularly in the SP group (Table 3). The two strains KUL and Moj 200 shared a same homozygous profile of 179 bp for the NG141-3 locus that was not observed elsewhere either. For the same DNA strains, similar profile results were observed, even though some parameters were changed, such as Taq DNA polymerases (Econotaq, Ozyme and Hot-Master Taq), PCR programs (temperature and number of cycles), or sequencers (MEGABACE 1000 or 96-capillary ABI 3730xl DNA Analyzer) (S2 Table).

Discussion

Microsatellite markers revealed a higher level of discrimination than any other markers used until now, and were used in this study to further explore the genetic diversity of *N. fowleri* and establish its population structure. To observe the same microsatellite profiles within a few years apart with different material and procedure i.e. different DNA extraction kits and sequencers, demonstrates the reproducibility of these markers. Another important point is the stability of the allelic profiles. Indeed, the DNA strains isolated a long time ago show the same profiles as recently isolated. Particulary, identical specific profiles with NG42 and NG141 were observed for the B or F strains.

In agreement with previous studies, numerous heterozygous were observed with most of the microsatellite markers suggesting that this species is diploid [37–39]. This was confirmed by the genome of *N. fowleri* recently published [36]. All loci are present in one copy in the *N. fowleri* genome except for NG69 that is present on two contigs. This confirms our duplication hypothesis that explains the third allele of 214 bp present in all samples.

The previous analyses based on the RAPD and ITS data identified five main genetic groups EA, WP, SP, CAT, and CHO. The two last groups encompass predominantly French strains and exhibited similarities with SP [6,15]. Specifically, the strains from the CHO group are exclusively French, originating in eastern France, and they belong to type 4 [15, 19]. While the CAT group contains both French and Japanese strains. The microsatellite markers confirmed these five groups and within two groups (SP and WP) enabled the identification of two new groups, i.e., the NZ group which was considered in the previous analysis based on the RAPD and ITS data as a member of SP and RA which belonged to the WP group. Each of these groups emerged from the different genetic analyses such as factorial correspondence analysis and NeighborNet network in SplitsTree. Furthermore, the different groups were found to be well differentiated as given by the very high F_{ST} value.

The examination of the NeighborNet network could provide some information about the evolutionary history of this species. The branching pattern of SP, NZ and CAT as well as EA

was in agreement with the results obtained with RAPD and ITS data [6,15]. In contrast, the branching position of CHO with the WP was unexpected since CHO displays ITS and RAPD similarities with SP, NZ and CAT [6,15]. According to the microsatellite data, the CHO and RA groups shared the same specific profile with the NG25 locus (116 bp), and could explain this grouping. Another unexpected result as to do with the emergence of the additional group RA. On the basis of the RAPD and ITS data, this group belonged to the WP group. However, in this study, it can be considered as a distinct group according to the F_{ST} value (Table 5 and Figs 1 and 2). In any cases, the RA group remains very closely related to the WP group but more extensive sampling is needed to confirm the branching pattern obtained in this study.

As previously found with RAPD and ITS analyses, ubiquist genotypes were also observed with the microsatellite markers, yet found to be more discriminating. This confirmed that a significant variability was not necessarily correlated with the geographical origin of the strains. For example, within the EA group, the two French strains, D1 and D2 from the same geographical site displayed two different allele profiles, and one of which was identical to the USA strains (WM and Lovell). Similar cases were observed within the WP group. Conversely, within the CAT group, the Japanese and French strains are distinguishable by using microsatellites in contrast to the RAPD and ITS markers. Additionally, the South Pacific strains exhibited allelic profiles that were not found elsewhere, and showed higher variability than the other groups.

The ubiquity of the genetic groups suggested clonal dispersion of the species. Another point which underlined this type of propagation is the excess of heterozygosity produced from most markers reflected by the negative F_{IS} value. However, no heterozygosity was observed with NG25 although variability was present (four alleles described).

Most of the variants EA, WP and SP were observed both from clinical and environmental isolates (Table 1). So far, the other variants CAT and CHO are only of environmental origin. As already mentioned, the fact that some variants are either less prevalent in the environment or endemic could explain why they are not currently found in clinical cases [40]. So it is very likely that all variants of N. fowleri are pathogens. Additionally, no degree of virulence between them has been reported so far.

As a potential diagnostic tool, tracking N. fowleri is important to understand the dispersal mode and the colonization of this species in the various sites. In a previous article, we underlined the interest to fingerprint the pathogenic species [6]. Discriminating markers allow better exploration of the environment and to further establish the level of genetic diversity in the sites. In addition, since more and more PAM cases are reported worldwide [41], these markers could allow a better identification of the N. fowleri isolates in patients. Until now, the ITS markers were currently used for detecting the variability of the N. fowleri species. However the genotypes which were identified are not sufficiently informative. For example, the EA group, which was only represented by the ITS genotype 3 is widely distributed in the European and American continents. In contrast, at least four different profiles, were obtained within the EA group with our microsatellite markers. This enables to better understand and identify the clinical and environmental N. fowleri strains.

Conclusion

Microsatellite markers confirmed the genetic heterogeneity within the species and can lead to a better tracking of the N. fowleri isolates. These markers were able to differentiate the genetic groups and within these, some strains by highlighting the private alleles.

Additionally, microsatellite markers allowed to better understand the evolutionary history of this species as well as its dispersal mode by determining the prevalence of the genetic groups in the sites. Our results exhibited the four similar genetic groups previously identified with ITS

and RAPD studies. However, the different clustering methods as well as all the tests used in this study produced two supplementary genetic groups, RA and NZ, that were previously included in the WP and SP groups, respectively. Interestingly, the RA group is very closely related to WP. Previous results based on ITS data showed that all members of WP and RA were identical. One might then suggest that WP and RA have diverged from a same lineage. However, additional data should be collected to confirm if it is indeed an isolated group. Similarly, other geographical strains such as the African and Asian ones are necessary to confirm the genetic groups, and additional microsatellite markers could be identified from the recent sequencing of the *N. fowleri* genome in order to reinforce this study.

Supporting Information

S1 Fig. Alignment sequence of the 2 alleles (225bp and 252bp) of the NG42 locus.
(DOCX)

S1 Table. Assignment of the 47 *N. fowleri* strains with GENECLASS.
(XLSX)

S2 Table. N. fowleri analysis. *N. fowleri* strains examined before this study.
(XLSX)

Acknowledgments

Author Contributions

References

1. Visvesvara GS, Roy SL, Maguire JH. Pathogenic and Opportunistic Free-Living Amebae: *Acanthamoeba* spp., *Balamuthia mandrillaris*, *Naegleria fowleri*, and *Sappinia pedata*. Tropical Infectious Diseases. Elsevier Inc.; 2007. pp. 707–713.

2. De Jonckheere JF, Pernin P, Scaglia M, Michel R. A comparative study of 14 strains of *Naegleria australiensis* demonstrates the existence of a highly virulent subspecies: *N. australiensis italica* n. spp. J Protozool. 1984; 31: 324–31. PMID: 6470990

3. Robinson BS, Monis PT, Henderson M, Gelonese S, Ferrante A. Detection and significance of the potentially pathogenic amoeboflagellate *Naegleria italica* in Australia. Parasitol Int. 2004; 53: 23–27. PMID: 14984832

4. Kilvington S, Beeching J. Development of a PCR for identification of *Naegleria fowleri* from the environment. Appl Environ Microbiol. 1995; 61: 3764–3767. PMID: 7487014

5. Reveiller FL, Varenne MP, Pougnard C, Cabanes P-A, Pringuez E, Pourima B, et al. An Enzyme-Linked ImmunoSorbent Assay (ELISA) for the Identification of Naegleria fowleri in Environmental Water Samples. J Eukaryot Microbiol. 2003; 50: 109–113. PMID: 12744523

6. Pélandakis M, Serre S, Pernin P. Analysis of the 5.8S rRNA Gene and the Internal Transcribed Spacers in *Naegleria* spp. and in *N. fowleri*. J Eukaryot Microbiol. 2000; 47: 116–121. PMID: 10750838

7. Pélandakis M, Pernin P. Use of multiplex PCR and PCR restriction enzyme analysis for detection and exploration of the variability in the free-living amoeba *Naegleria* in the environment. Appl Environ Microbiol. 2002; 68: 2061–2065. PMID: 11916734

8. Qvarnstrom Y, Visvesvara GS, Sriram R, da Silva AJ. Multiplex real-time PCR assay for simultaneous detection of *Acanthamoeba* spp., *Balamuthia mandrillaris*, and *Naegleria fowleri*. J Clin Microbiol. 2006; 44: 3589–95. PMID: 17021087

9. Behets J, Declerck P, Delaedt Y, Verelst L, Ollevier F. A duplex real-time PCR assay for the quantitative detection of *Naegleria fowleri* in water samples. Water Res. 2007; 41: 118–126. PMID: 17097714

10. Puzon GJ, Lancaster JA, Wylie JT, Plumb JJ. Rapid Detection of *Naegleria fowleri* in Water Distribution Pipeline Biofilms and Drinking Water Samples. Environ Sci Technol. American Chemical Society; 2009; 43: 6691–6696.

11. Madarová L, Trnková K, Feiková S, Klement C, Obernauerová M. A real-time PCR diagnostic method for detection of *Naegleria fowleri*. Exp Parasitol. 2010; 126: 37–41. doi: 10.1016/j.exppara.2009.11.001 PMID: 19919836

12. De Jonckheere JF. Characterization of *Naegleria* species by restriction endonuclease digestion of whole-cell DNA. Mol Biochem Parasitol. 1987; 24: 55–66. PMID: 3039368

13. De Jonckheere JF. Geographic origin and spread of pathogenic *Naegleria fowleri* deduced from restriction enzyme patterns of repeated DNA. Biosystems. 1988; 21: 269–275. PMID: 2840136

14. Pélandakis M, Kaundun SS, Jonckheere JF, Pernin P. DNA diversity among the free-living amoeba, *Naegleria fowleri*, detected by the random amplified polymorphic DNA method. FEMS Microbiol Lett. 1997; 151: 31–39.

15. Pélandakis M, De Jonckheere JF, Pernin P. Genetic variation in the free-living amoeba *Naegleria fowleri*. Appl Environ Microbiol. 1998; 64: 2977–2981. PMID: 9687460

16. Tiewcharoen, Junnu V, Sassa-deepaeng T, Waiyawuth W, Ongrothchanakul J, Wankhom S. Analysis of the 5.8S rRNA and internal transcribed spacers regions of the variant *Naegleria fowleri* Thai strain. Parasitol Res. 2007; 101: 139–43. PMID: 17279395

17. De Jonckheere JF. Sequence Variation in the Ribosomal Internal Transcribed Spacers, Including the 5.8S rDNA, of *Naegleria* spp. Protist. Elsevier; 1998; 149: 221–8.

18. Zhou L, Sriram R, Visvesvara GS, Xiao L. Genetic Variations in the Internal Transcribed Spacer and Mitochondrial Small Subunit rRNA Gene of *Naegleria* spp. J Eukaryot Microbiol. 2003; 50: 522–526. PMID: 14736150

19. De Jonckheere JF. Origin and evolution of the worldwide distributed pathogenic amoeboflagellate *Naegleria fowleri*. Infection, Genetics and Evolution. 2011. pp. 1520–1528. doi: 10.1016/j.meegid.2011.07.023 PMID: 21843657

20. Gur-Arie R, Cohen CJ, Eitan Y, Shelef L, Hallerman EM, Kashi Y. Simple sequence repeats in *Escherichia coli*: abundance, distribution, composition, and polymorphism. Genome Res. 2000; 10: 62–71. PMID: 10645951

21. Tóth G, Gáspári Z, Jurka J. Microsatellites in different eukaryotic genomes: survey and analysis. Genome Res. 2000; 10: 967–981. PMID: 10899146

22. Anderson TJ, Haubold B, Williams JT, Estrada-Franco JG, Richardson L, Mollinedo R, et al. Microsatellite markers reveal a spectrum of population structures in the malaria parasite *Plasmodium falciparum*. Mol Biol Evol. 2000; 17: 1467–1482. PMID: 11018154

23. Karunaweera ND, Ferreira MU, Hartl DL, Wirth DF. Fourteen polymorphic microsatellite DNA markers for the human malaria parasite *Plasmodium vivax*. Mol Ecol Notes. 2006; 7: 172–175.

24. Oliveira RP, Broude NE, Macedo AM, Cantor CR, Smith CL, Pena SD. Probing the genetic population structure of *Trypanosoma cruzi* with polymorphic microsatellites. Proc Natl Acad Sci U S A. 1998; 95: 3776–3780. PMID: 9520443

25. Mallon M, MacLeod A, Wastling J, Smith H, Reilly B, Tait A. Population structures and the role of genetic exchange in the zoonotic pathogen *Cryptosporidium parvum*. J Mol Evol. 2003; 56: 407–417. PMID: 12664161

26. Schwenkenbecher JM, Wirth T, Schnur LF, Jaffe CL, Schallig H, Al-Jawabreh A, et al. Microsatellite analysis reveals genetic structure of *Leishmania tropica*. Int J Parasitol. 2006; 36: 237–246. PMID: 16307745

27. Kebede N, Oghumu S, Worku A, Hailu A, Varikuti S, Satoskar AR. Multilocus microsatellite signature and identification of specific molecular markers for *Leishmania aethiopica*. Parasit Vectors. 2013; 6: 160. doi: 10.1186/1756-3305-6-160 PMID: 23734874

28. Rousset F. genepop'007: a complete re-implementation of the genepop software for Windows and Linux. Mol Ecol Resour. 2008; 8: 103–106. doi: 10.1111/j.1471-8286.2007.01931.x PMID: 21585727

29. Goudet J. FSTAT (Version 1.2): A Computer Program to Calculate F-Statistics. J Hered. 1995; 86: 485–486.

30. Jombart T. adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics. 2008; 24: 1403–1405. doi: 10.1093/bioinformatics/btn129 PMID: 18397895

31. Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F. GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations. Lab génome, Popul Interact CNRS Umr. 1996; 5000: 1996–2004.

32. Piry S, Alapetite A, Cornuet J-M, Paetkau D, Baudouin L, Estoup A. GENECLASS2: a software for genetic assignment and first-generation migrant detection. J Hered. 2004; 95: 536–9. PMID: 15475402

33. Rannala B, Mountain JL. Detecting immigration by using multilocus genotypes. Proc Natl Acad Sci. 1997; 94: 9197–9201. PMID: 9256459

34. Paetkau D, Slade R, Burden M, Estoup A. Genetic assignment methods for the direct, real-time estimation of migration rate: a simulation-based exploration of accuracy and power. Mol Ecol. 2004; 13: 55–65. PMID: 14653788

35. Huson DH, Bryant D. Application of phylogenetic networks in evolutionary studies. Mol Biol Evol. 2006; 23: 254–67. PMID: 16221896

36. Zysset-Burri DC, Müller N, Beuret C, Heller M, Schürch N, Gottstein B, et al. Genome-wide identification of pathogenicity factors of the free-living amoeba Naegleria fowleri. BMC Genomics. 2014; 15: 496. doi: 10.1186/1471-2164-15-496 PMID: 24950717

37. Cariou ML, Pernin P. First Evidence for Diploidy and Genetic Recombination in Free-Living Amoebae of the Genus Naegleria on the Basis of Electrophoretic Variation. Genetics. 1987; 115: 265–270. PMID: 17246363

38. Adams M, Andrews RH, Robinson B, Christy P, Baverstock PR, Dobson PJ, et al. A genetic approach to species criteria in the amoeba genus Naegleria using allozyme electrophoresis. Int J Parasitol. 1989; 19: 823–834. PMID: 2635158

39. Fritz-Laylin LK, Ginger ML, Walsh C, Dawson SC, Fulton C. The Naegleria genome: a free-living microbial eukaryote lends unique insights into core eukaryotic cell biology. Res Microbiol. 2011; 162: 607–18. doi: 10.1016/j.resmic.2011.03.003 PMID: 21392573

40. De Jonckheere JF. What do we know by now about the genus Naegleria? Exp Parasitol. 2014; 145: S2–S9. doi: 10.1016/j.exppara.2014.07.011 PMID: 25108159

41. Siddiqui R, Khan NA. Primary Amoebic Meningoencephalitis Caused by Naegleria fowleri: An Old Enemy Presenting New Challenges. PLoS Negl Trop Dis. Public Library of Science; 2014; 8: e3017.