CMLS, Cell. Mol. Life Sci. 56 (1999) 729–734
1420-682X/99/100729-06 $ 1.50 + 0.20/0

CMLS Cellular and Molecular Life Sciences

Short sequence repeats in microbial pathogenesis and evolution

A. van Belkum

Erasmus University Medical Centre Rotterdam, Department of Medical Microbiology & Infectious Diseases, Dr. Molewaterplein 40, 3015 GD Rotterdam (The Netherlands), Fax + 31 10 4633875, e-mail: vanbelkum@bacl.azr.nl

Abstract. Repetitive DNA is ubiquitous in microbial genomes. Different classes of short sequence repeats (SSRs) have been identified and demonstrated to be generally heterogeneous in a locus-dependent manner, reflected in variation in the number of repeat units present at a given genomic site or by sequence heterogeneity among individual units. Both types of variability can be used to assess intra-species genetic diversity. Repeat variability often affects the coding potential of the region in which the repetitive element is located. This implies that determination of the primary structure of variable numbers of tandem repeats can be used for epidemiological identification purposes, and also for the analysis of gene function. Precise assessment of SSR structure can also generate insight into the regulation of gene expression. Together, DNA repeat analysis in microbial species provides information on both functional and evolutionary aspects of genetic diversity among microbial isolates.

Key words. Short sequence repeats; variable number of tandem repeat loci; slipped strand mispairing; molecular epidemiology; microbial evolution; *Haemophilus influenzae*; *Helicobacter pylori*; *Candida albicans*.

Introduction

Today, humans are confronted with the rise and spread of several new infectious diseases. Novel pathogens have been identified, while classical diseases with a supposed non-infectious etiology appear to be associated with infectious agents as well. Examples of recently discovered bacterial pathogens include *Legionella pneumophila*, *Borrelia burgdorferi* and *Bartonella henselae*, the agents of Legionnaires' disease, Lyme disease and cat scratch disease, respectively [1]. Well-known clinical syndromes such as peptic ulcer and atherosclerosis are now thought to have, at least in part, an infectious etiology. In these two particular clinical syndromes, *Helicobacter pylori* and *Chlamydia pneumoniae*, respectively, have been identified as putative causative agents [2, 3]. This has simulated interest in infectious disease research, by clinicians and microbiologists in particular, but also by epidemiologists, (molecular) biologists, and pharmacists. Consequently, modern medical microbiology has become a highly multidisciplinary science, where different forms of expertise are integrated in a combined effort to combat a further increase in the number (and spread) of infections and not only those encountered in the medical setting.

Besides the proliferation of novel and well-known microbial species in newly recognised disease entities, knowledge is increasingly being gained concerning the spread of microbial characteristics per se. Features such as virulence or resistance against anti-microbial agents are often encoded by a limited number of genes that seem to find their way throughout the microbial kingdom. In addition to undergoing extensive reshuffling and dissemination within the species context [see for example ref. 4], it has also been established that for many characteristics, species borders do not obstruct the spread of the genes involved. Clearly, the combined spread of genes and genomes requires careful epidemiological monitoring to gain insight into possible and novel modes of disease prevention: halting the dissemination of infectious agents can only be achieved once the routes along which the spreading occurs are known in at least some detail. Molecular markers useful for assessing gene and genome diversity and monitoring microbial evolution in a broad sense are essential com-

ponents of these research efforts. Repetitive DNA can provide the molecular targets required and, moreover, might be helpful additionally in assessing microbial virulence status.

The ubiquity of repetitive DNA

DNA repeats can either be dispersed or contiguous. Examples of dispersed prokaryotic repeats are the repetitive extragenic palindrome (REP) elements [5], the enterobacterial repetitive intergenic consensus (ERIC) motifs [6], and the BOX repeats [7]. These moieties share the characteristic that, if they occur in a given genome, they generally do so in single copies per locus, but in multiple copies per genome: the repeat units are spaced by (long stretches of) non-repetitive DNA. However, the number and location of the repeats in a microbial chromosome can differ widely from strain to strain within a given species. Many examples of the ways in which these repetitive elements can be used to gain insight into bacterial genome variability have been presented [8–10]. Although present in many species and apparently associated with a high degree of genetic flexibility, the function of these repeat elements is still largely enigmatic. The results of a limited number of recent studies, however, suggest that, for example the pneumococcal BOX elements might be involved in the regulation of phenomena involved in phase variation or might function as hotspots for homologous recombination [11, 12].

The presence of contiguous stretches of repetitive DNA has been demonstrated for genomes of a diversity of organisms as well [13]. In prokaryotes, many examples have been documented and many different classes of repetitive DNA have been identified. DNA repeats can be categorized on the basis of their location with respect to genes, the size of the repeat unit, and whether or not all repeat units are identical (homogeneous repeat) or sequence variable (degenerate repeat). Many repeat regions are length variable: the number of units per genetic locus can alter drastically and the so-called variable number of tandem repeat (VNTR) or, depending on the length of the repeat unit, short sequence repeat (SSR) regions are thus defined. Frequently, and this is particularly true for the densely 'gene-packed' genomes of microorganisms, repeats are positioned in such a way that variability in repeat number affects gene function. Repeat variability can change the physical integrity of functional DNA domains, whereas open reading frames can be disrupted or otherwise modified [see ref. 13 for a review].

Mechanistic aspects of repeat variability

The mechanisms underlying repeat variability are also diverse. Recombination can be a major mechanism explaining VNTR behavior, but so-called slipped strand mispairing (SSM) is the phenomenon most frequently held responsible for variability in repeat numbers [14]. By a combination of inadequate DNA polymerase action and lack of efficient repair during DNA replication, sequence units can either be inserted or deleted. In humans, expansion of certain repeat regions has been associated with several neurodegenerative diseases [15]. In an elegant model study, the same type of SSR variability was associated with dysfunction and apoptotic death of sensory neurons in *Caenorhabditis elegans* [16]. In both eukaryotic and prokaryotic microorganisms, the structure-function relationship of repetitive DNA has been the subject of many investigations and this review will discuss examples for three medically relevant species: the Gram-negative bacteria *Haemophilus influenzae* and *H. pylori* and the human pathogenic yeast species *Candida albicans*.

SSRs in the *H. influenzae* genome

H. influenzae colonizes the human upper respiratory tract. This bacterial species can also cause serious infections such as life-threatening meningitis. Several research groups have been studying *H. influenzae* pathogenicity, and the involvement of repetitive DNA in modulation of bacterial virulence has been documented on several occasions. In the late 1980s and early 1990s, repeat variability was experimentally associated with modulation of the expression of genes involved in pilus or lipopolysaccharide (LPS) synthesis [17]. The presence of a repeat consisting of a variable number of dinucleotide motifs in a dual promoter controlling mRNA synthesis of genes encoding pilus subunit proteins appeared to be a determining factor in gene expression [18]. Depending on the number of repeat units, the spacing between the -35 and -10 promoter consensus motifs either enabled or obstructed efficient recognition of this site by RNA polymerase. Essentially, this repeat locus provided a regulatory on/off switch at the transcription level, a feature that was identified in many other bacterial species as well [19, 20]. Other examples of *H. influenzae* repeats were clearly influencing translation events. In several genes encoding LPS biosynthesis enzymes, repeats built from tetranucleotide units appeared to be present. These repeats were located within open reading frames and, consequently, variation in the number of repeat units influenced the protein-encoding capacity of a given mRNA molecule [17]. When the whole genome sequence of *H. influenzae* became available, a complete catalogue of SSR loci

could be established [21]. Interestingly, most, if not all, of these repeats were associated with potential virulence genes such as those encoding iron-acquisition factors, adhesion molecules, and the aforementioned LPS biosynthesis enzymes. Overall, *H. influenzae* appears to be a key example of a bacterial species that is using repeat variability as an instrument facilitating specific interactions with a frequently hostile environment [22]. In addition, many other forms of genetic heterogeneity among strains of *H. influenzae*, mostly associated with antigenic diversity, have been documented [23].

SSRs in *H. pylori* whole genome sequences

SSR regions were identified in full detail once the entire genome sequence for *H. pylori* strain HP26695 became publicly available [24]. In this 1997 publication by The Institute for Genomic Research TIGR, contingency genes as described above for *H. influenzae* were again encountered in large numbers. A subsequent computer analysis revealed the presence of more than two dozen potentially SSR-dependent and, consequently, phase-variable genes [25]. The SSRs were associated with genes encoding products involved in the biosynthesis of LPS, cell-surface-associated proteins and restriction modification enzymes. A limited number of genes could not be identified on the basis of homology searches. Early in 1999, the second entire genome sequence for another isolate of the same species became available [26]. In a combined effort, researchers from the Astra Research Center in Boston, the University of Alberta in Edmonton, and Genome Therapeutics Corporation in Waltham, Massachusetts, elucidated the primary structure of the chromosome of strain HPJ99. This facilitated, for the first time, a detailed comparison of the full gene complement of two strains from within the same species and several interesting differences were documented. With respect to SSR polymorphism, 'in silico' evidence for potential phase variation was obtained. Table 1 surveys the SSR sizes for the candidate genes: it is obvious that in several cases, the size of the repeat determines whether or not a gene is switched 'on.' This confirms the putative involvement of many of these candidate contingency genes in antigenic variation or

Table 1. Comparison of the SSR content of the whole genome sequences available for two different *Helicobacter pylori* isolates [adapted from refs 24–26].

Repeat sequence	Gene involved	TIGR HP nr	Number of repeat units	
			HP J99	HP26695
C	α(1–3)fucosyltransferase	0651	5 (+)	13 (+)
C	α(1–3)fucosyltransferase	0379	13/9 (−)	13/6 (+)
C	α(1–2)fucosyltransferase	0093/4	13 (−)	14 (−)
C	Lex2B	0619	12 (−)	13 (−)
C	LPS biosynthesis related	nd	14 (+)	nd
AG	RfaJ glycosyltransferase	0208	8 (−)	11 (−)
C	FliP	0684/5	8 (+)	9 (−)
GA	adherence-protein-associated protein	1417	10 (−)	9 (−)
C	streptococcal M protein	0058	nd	15 (−)
CT	outer membrane protein	0638	9 (+)	6 (+)
GA	alginate O-acetylation protein	0855	nd	5 (+)
T	heme-binding lipoprotein	0298	nd	9 (+)
C	methyl-accepting chemotaxis protein	0103	nd	12 (*)
C	outer membrane protein/adhesin	0896	8 (+)	11 (+)
C	outer membrane protein/adhesin	0722	9 (−)	8 (−)
C	outer membrane protein/adhesin	0725	9 (−)	6 (−)
A	ABC transporter/secretion protein	1206	9 (+)	10 (+)
C	A-specific methyltransferase	1353/4	13/12 (−)	12/15 (−)
C	R/M methyltransferase	1369/70	nd	10 (−)
C	restriction enzyme β-subunit	1471	14 (+)	14 (+)
G	A-specific methyltransferase	1522	11 (−)	12 (−)
G	type I restriction enzyme R protein	0464	10 (−)	15 (+)
CT	C-specific methyltransferase	0051	nd	5 (+)
T	not known	0585/6	nd	8 (−)
G	not known	0217	6/7 (−)	12/6 (+)
AG	not known	0744	7 (−)	9 (−)
AT	not known	0211	nd	5 (+)
G	not known	0335	nd	9 (−)

The TIGR identification number provided corresponds to the HP26695 gene order. nd, not documented; whether this is due to computer search inefficiency or genetic deletion remains to be established; (+) gene involved is switched on; (−) gene involved is switched off; (*) repeat is located in a promoter region.

adaptive evolution. It is interesting to note that in the laboratory setting, the repeats continued to vary; in the case of *H. influenzae*, no such changes were defined [27]. It would be worthwhile investigating the involvement of repeat variability and virulence potential in biological phenomena such as intestinal colonization efficacy or the potential to inflict a pathogenic effect.

SSRs in *C. albicans* genomic sequences

No clear-cut examples of repeats involved in the virulence of *C. albicans* have been documented thus far. However, as part of the 'hunt for whole genome sequences,' large amounts of *C. albicans* sequence information have been deposited in the publicly available nucleotide sequence databases. Using computerized searches, many examples of repetitive DNA moieties can be detected in these nucleotide sequences. It is interesting to note that many of the *C. albicans* repeats are polymers of either CAA and/or CAG [28]. These repeats are often located within open reading frames and provide templates for polyglutamine at the protein level, giving rise to a phenomenon similar to that described for the human neurodegenerative diseases as well [15]. Since it has been suggested that these alkaline stretches of glutamine could be involved in specific interactions with, for example, fungal promoter sequences, studies into the involvement of these repeats in the regulation of fungal virulence genes have been initiated. However, to date, no association between repeat length and, for example, the transition from colonization to invasion has been reported [29]. Figure 1 illustrates the nature of the experimental data obtained. Since *C. albicans* is a diploid organism, more than a single product can be anticipated when the SSR loci are amplified by PCR. However, sometimes more than two PCR products are found, probably due to the formation of concatemeric complexes. This is confirmed when the products are analyzed in the presence of a denaturing agent such as urea. Only a single (homozygous) or a double (heterozygous) band is observed (see bottom panel in fig. 1).

In the case of lower eukaryotes, whole genome sequences can also be used for narrowing down SSRs [30]. For *Saccharomyces cerevisiae*, the implications of repeat variability can be deduced not only from experimental data but also on the basis of the chromosomal sequences available.

VNTR regions as evolutionary hotspots

The examples above illustrate that SSRs are certainly not without function in the regulation of gene expression in microorganisms. Besides these functional aspects, repeat loci can also be used to study strain relatedness. Since repeats evolve relatively quickly (frequencies as high as one insertion or deletion event per 1000 replications have been suggested [31]) they may be excellent targets for monitoring short-term bacterial evolution or serve as targets for epidemiological typing. Many repeats, as present in the genome of clinically relevant microorganisms, have been used for tracking outbreaks of infection. In *Staphylococcus aureus*, for example, many repeats have been documented to occur in so-called microbial surface components recognizing adhesive matrix molecules (MSCRAMMs) [32]. From a functional perspective, these repeats seem to provide stereotactic spacing domains at the protein level. The larger the number of repeats the more the active domain of these generally cell-surface-associated proteins will

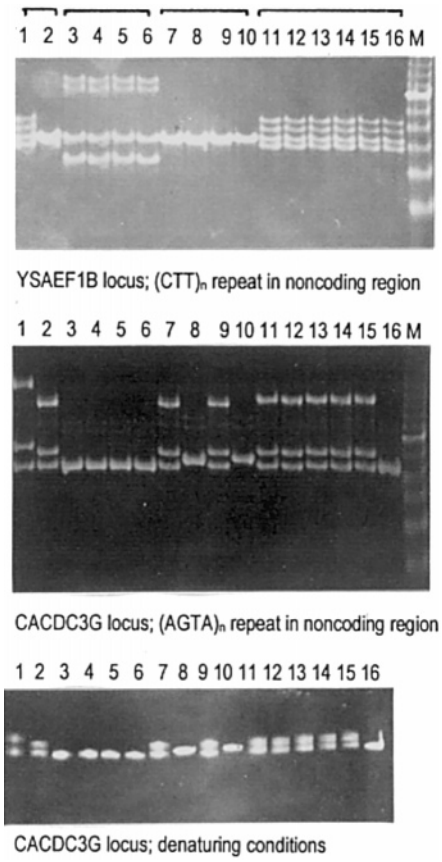

Figure 1. SSR mapping for a number of *Candida albicans* strains. Shown are the DNA fragments obtained upon amplification of a trinucleotide repeat (YSAEF1B) and a tetranucleotide repeat (CACDC3G) both located in non-coding regions [29]. Clear polymorphism is documented among strains, even among those isolated from a single patient (sample numbers overlined). The bottom panel displays the picture obtained upon electrophoresis of the CACDC3G amplicons in a denaturing gel, so that the hetero- or homozygosity of the diploid strains can be deduced.

protrude from the surface of the microbial cell. Both numbers of repeats and repeat unit sequences have been used to perform 'comparative genomics.' Based on identity versus non-identity, epidemiologically linked strains could be discriminated from independent isolates [33]. For *H. influenzae*, similar approaches could be developed based on the repeats encountered by computerized screening of the whole genome sequence. By size assessment using straightforward PCR-mediated amplification, clonally related isolates could be identified [27, 34]. These studies revealed, however, that even among closely related strains, extensive repeat polymorphism could sometimes be documented, probably due to the selective pressures induced by the host which select for the most accurately adapted bacterial cells.

Concluding remarks

The study of repetitive DNA has provided a large wealth of data on the molecular pathogenicity and genomic variability of microorganisms. Many examples of SSRs involved in the modulation of gene expression have been identified recently in a wide variety of microorganisms [25, 30, 35, 36], including viruses [37]. It is anticipated that the functional knowledge obtained for repeats in microorganisms will ultimately help in establishing structure-function relationships for the human repeats as well. As more whole genome sequences become available, they will provide the theoretical framework for the implementation of more detailed studies into the involvement of repeat variation in bacterial pathogenicity and genetic adaptability.

1 Fredricks D. N. and Relman D. A. (1996) Sequence based identification of microbial pathogens: a reconsideration of Koch's postulates. Clin. Microbiol. Rev. **9:** 18–33
2 Blaser M. J. (1997) Not all *Helicobacter pylori* strains are created equal: should all be eliminated? Lancet **349:** 1020–1022
3 Campbell L. A., O'Brien E. R., Cappuccio A. L., Kuo C. C., Wang S. P., Stewart D. et al. (1995) Detection of *Chlamydia pneumoniae* TWAR in human coronary atherectomy tissues. J. Infect. Dis. **172:** 585–588
4 Laible G., Spratt B. G. and Hakenbeck R. (1991) Interspecies recombinational events during the evolution of altered PBP2X genes in penicillin-resistant clinical isolates of *Streptococcus pneumoniae*. Mol. Microbiol. **5:** 1993–2002
5 Versalovic J. and Lupski J. R. (1996) Distinguishing bacterial and fungal pathogens by repetitive-sequence based PCR (rep PCR). Lab. Med. Int. **13:** 7–8
6 Sciacchitano C. J. (1998) DNA fingerprinting of *Listeria monocytogenes* using enterobacterial repetitive intergenic consensus (ERIC) motifs polymerase chain reaction/capillary electrophoresis. Electrophoresis **19:** 66–70
7 Martin B., Humbert O., Camara M., Guenzi E., Walker J., Mitchell T. et al. (1992) A highly conserved repeated DNA element located in the chromosome of *Streptococcus pneumoniae*. Nucleic Acids Res. **20:** 3473–3483
8 Vinuesa P., Rademaker J. L. W., De Bruin F. J. and Werner D. (1998) Genotypic characterization of *Bradyrhizobium* strains nodulating endemic woody legumes of the Canary islands by PCR restriction fragment polymorphism analysis of genes encoding 16S rRNA (16S rDNA) and 16S-23S rDNA intergenic spacers, repetitive extragenic palindrome PCR genomic fingerprinting and partial 16S rDNA sequencing. Appl. Eviron. Microbiol. **64:** 2096–2104
9 Sluijter M., Faden H., Groot R. de, Lemmens N., Goessens W. H. F., Belkum A. van et al. (1998) Molecular characterisation of pneumococcal nasopharynx isolates collected from children during their first two years of life. J. Clin. Microbiol. **36:** 2248–2253
10 Belkum A. van, Bax R., Peerbooms P., Goessens W. H. F., Leeuwen N. van and Quint W. G. V. (1993) Comparison of phage typing and DNA fingerprinting by PCR for discrimination of methicillin resistant *Staphylococcus aureus* strains. J. Clin. Microbiol. **31:** 798–803
11 Saluja S. K. and Weiser J. N. (1995) The genetic basis of colony opacity in *Streptococcus pneumoniae*: evidence for the effect of BOX elements on the frequency of phenotypic variation. Mol. Microbiol. **16:** 215–227
12 Krauss J. and Hakenbeck R. (1997) A mutation in the D,D-carboxypeptidase penicillin binding protein-3 of *Streptococcus pneumoniae* contributes to cefotaxime resistance of the laboratory mutant C604. Antimicrob. Agents Chemother. **41:** 936–942
13 Belkum A. van, Scherer S., Van Alphen L. and Verbrugh H. (1998) Short-sequence repeats in prokaryotic genomes. Microbiol. Mol. Biol. Rev. **62:** 275–293
14 Strand S., Prolla T. A., Liskay R. M. and Petes T. D. (1993) Destabilisation of tracts of simple repetitive DNA in yeast by mutations affecting DNA mismatch repair. Nature **365:** 274–276
15 Caskey C. T., Pizzutti A., Fu Y., Fenwick Jr R. G. and Nelson D. L. (1992) Triplet repeat mutations in human disease. Science **256:** 784–789
16 Faber P. W., Alter J. R., MacDonald M. E. and Hart A. C. (1999) Polyglutamine-mediated dysfunction and apoptotic death of a *Caenorhabditis elegans* sensory neuron. Proc. Natl. Acad. Sci. USA **96:** 179–184
17 Weiser J. N., Love J. M. and Moxon E. R. (1989) The molecular mechanism of phase variation of *H. influenzae* lipopolysaccharide. Cell **59:** 657–665
18 Van Ham M., Van Alphen L., Mooi F. R. and Van Putten J. P. M. (1993) Phase variation of *H. influenzae* fimbriae: transcriptional control of two divergent genes through a variable combined promoter region. Cell **73:** 1187–1196
19 Foster P. L. and Trimarchi. J. M. (1994) Adaptive reversion of a frameshift mutation in *Escherichia coli* by simple base deletions in homopolymeric runs. Science **265:** 407–409
20 Van der Ende A., Hopman A., Zaat S., Essink B. B., Berkhout B. and Dankert J. (1995) Variable expression of class 1 outer membrane protein in *Neisseria meningitidis* is caused by variation in the spacing between the −10 and −35 regions of the promoter. J. Bacteriol. **177:** 2475–2480
21 Fleischmann R. D., Adams M. D., White O., Clayton R. A., Kirkness E. F., Kerlavage E. R. et al. (1995) Whole genome random sequencing and assembly of *Haemophilus influenzae* Rd. Science **269:** 496–512
22 Hood D. W., Deadman M. E., Jennings M. P., Bisercic M., Fleischmann R. D., Venter J. C. et al. (1996) DNA repeats identify novel virulence genes in *Haemophilus influenzae*. Proc. Natl. Acad. Sci. USA **93:** 11121–11125
23 Gilsdorf J. R. (1998) Antigenic diversity and gene polymorphisms in *Haemophilus influenzae*. Infect. Immun. **66:** 5053–5059
24 Tomb J. F., White O., Kerlavage A. R., Clayton A. R., Sutton G. G., Fleischmann R. D. et al. (1997) The complete genome sequence of the gastric pathogen *Helicobacter pylori*. Nature **338:** 539–547
25 Saunders N. J., Peden J. F., Hood D. W. and Moxon E. R. (1998) Simple sequence repeats in the *Helicobacter pylori* genome. Mol. Microbiol. **27:** 1091–1098
26 Alm R. A., Ling L. S. L., Moir D. T., King B. L., Brown E. D., Doig P. C. et al. (1999) Genomic-sequence comparison of

two unrelated isolates of the human gastric pathogen *Helicobacter pylori*. Nature **397:** 176–180

27 Belkum A. van, Scherer S., Van Leeuwen W., Willemse D., Van Alphen L. and Verbrugh H. A. (1997) Variable number of tandem repeats in *Haemophilus influenzae*. Infect. Immun. **65:** 5017–5027

28 Field D. and Wills C. (1996) Long polymorphic microsatellites in simple organisms. Proc. R. Soc. Lond. B **263:** 209–215

29 Verduyn Lunel F., Licciardello L., Stefani S., Verbrugh H. A., Melchers W. J. G., Meis J. F. G. M. et al. (1998) Lack of consistent short sequence repeat polymorphisms in genetically homologous colonising and invasive *Candida albicans* strains. J. Bacteriol. **180:** 3771–3778

30 Field D. and Wills. C. (1998) Abundant microsatellite polymorphism in *Saccharomyces cerevisiae*, and the different distributions of microsatellites in eight prokaryotes and *S. cerevisiae*, result from mutation pressures and a variety of selective forces. Proc. Natl. Acad. Sci. USA **95:** 1647–1652

31 Levinson G. and Gutman. G. A. (1987) High frequency of short frameshifts in poly-CA/GT tandem repeats borne by bacteriophage M13 in *Escherichia coli* K-12. Nucleic Acids Res. **15:** 5323–5338

32 Patti J. M., Allen B. L., McGavin M. J. and Hook M. (1994) MSCRAMM-mediated adherence of microorganisms to host tissue. Annu. Rev. Microbiol. **48:** 585–617

33 Belkum A. van, Riewerts Eriksen N., Sijmons M., Van Leeuwen W. B., VandenBergh M., Kluytmans J. A. J. W. et al. (1997) Coagulase and protein A polymorphisms do not contribute to persistence of nasal colonisation by *Staphylococcus aureus*. J. Med. Microbiol. **46:** 222–232

34 Belkum A. van, Melchers W., Ijsseldijk C., Nohlmans L., Verbrugh H and Meis J. F. G. M. (1997) Outbreak of amoxicillin resistant *Haemophilus influenzae*: variable number of tandem repeats as novel molecular markers. J. Clin. Microbiol. **35:** 1517–1520

35 Jackson P. J., Walthers E. A., Kalif A. S., Richmond K. L., Adair D. M., Hill K. K. et al. (1997) Characterisation of the variable-number tandem repeats in vrrA from different *Bacillus anthracis* isolates. Appl. Environ. Microbiol. **63:** 1400–1405

36 Glew M. D., Baseggio N., Markham P. F., Browning G. F. and Walker. I. D. (1998) Expression of the pMGA genes of *Mycoplasma gallisepticum* is controlled by variation in the GAA trinucleotide repeat lengths within the 5′ noncoding regions. Infect. Immun. **66:** 5833–5841

37 Zaia J. A., Gallez-Hawkins G., Churchill M. A., Morton-Blackshire A., Pande H., Adler S. P. et al. (1990) Comparative analysis of human cytomegalovirus *a* sequence in multiple clinical isolates by using polymerase chain reaction and restriction fragment length polymorphism assays. J. Clin. Micobiol. **28:** 2602–2607

.