European Journal of Human Genetics (2001) 9, 347–354

www.nature.com/ejhg

ARTICLE

Genetic polymorphism of *MUC7*: Allele frequencies and association with asthma

Helen J Kirkbride[1], Jan G Bolscher[2], Kamran Nazmi[2], Lynne E Vinall[1], Matthew W Nash[1], Fiona M Moss[3], David M Mitchell[4] and Dallas M Swallow*[,1]

[1]*MRC Human Biochemical Genetics Unit, Galton Lab, University College London, 4 Stephenson Way, London NW1 2HE, UK;* [2]*Academic Centre for Dentistry, Dept of Oral Biology, Van der Boechorststraat 7, NL-1081 BT Amsterdam, The Netherlands;* [3]*Department of Thoracic Medicine, Central Middlesex Hospital, Acton Lane, London NW10 7NS, UK;* [4]*Chest & Allergy Clinic, St Mary's NHS Trust, Praed Street, London W2 1NY, UK*

***MUC7* encodes a small salivary mucin, previously called MG2, a glycoprotein with a putative role in facilitating the clearance of oral bacteria. The central domain of this glycoprotein was previously shown to comprise five or six tandemly repeated units of 23 amino-acids which carry most of the *O*-linked glycans. The polymorphism of these two allelic forms (*MUC7*5* or *MUC7*6*) has been confirmed in this study in which we have analysed a large cohort of subjects ($n = 375$) of various ethnic origins. We have also identified a novel rare allele with eight tandem repeats (*MUC7*8*). *MUC7*6* was the most common allele (0.78–0.95) in all the populations tested. The tandem repeat arrays of 22 *MUC7*5* alleles and 34 *MUC7*6* alleles were sequenced. No sequence differences were detected in any of the *MUC7*6* alleles. Twenty-one *MUC7*5* alleles sequenced lacked the 4th tandem repeat (structure TR12356), while one showed the structure TR12127. The structure of the *MUC7*8* allele was TR12343456. Because of the known role of MUC7 in bacterial binding, and thus its potential involvement in susceptibility to chest disease we also tested *MUC7* in our previously described series of Northern European atopic individuals with and without associated asthma. The *MUC7*5* allele was rarer in the atopic asthmatics than in the atopic non-asthmatics ($P = 0.014$, OR for no asthma in atopic individuals 3.13, CI 1.01–6.10), and the difference in frequency between all asthmatics and all non-asthmatics was statistically significant ($P = 0.009$) while there was no difference between atopy and non-atopy ($P = 0.199$). In this study we also report the electrophoretic analysis of the MUC7 glycoprotein in saliva from individuals of different *MUC7* genotype.** *European Journal of Human Genetics* (2001) **9,** 347–354.

Keywords: MUC7; mucin; polymorphism; asthma; atopy; MG2

Introduction

The mucus that covers delicate epithelial surfaces of the respiratory, gastrointestinal and reproductive tracts functions in a protective capacity to prevent desiccation and provide lubrication and defence from environmental bacteria and internal enzymes. The major glycoproteins of the mucus are the mucins, a heterogeneous family of highly glycosylated proteins synthesised in epithelial cells. So far 12 genes (*MUC1*, *MUC2*, *MUC3A*, *MUC3B*, *MUC4*, *MUC5AC*, *MUC5B*, *MUC6*, *MUC7*, *MUC8*, *MUC9* and *MUC12*) have been reported to encode these proteins in humans (http://www.gene.ucl.ac.uk/nomenclature). Some of the genes encode glycoproteins that are secreted (MUC2, MUC5AC, MUC5B, MUC6 and MUC7, MUC9)[1–6] and some that are in part membrane bound (MUC1, MUC3A and B, MUC4, MUC12).[7–10] The *MUC* genes share the common feature of having a centrally located region of sequence that encodes tandem repeats (TRs) that may comprise more than 50% of the apomucin.[11] This region contains a large number of potential sites for *O*-linked

***Correspondence: DM Swallow, Galton Lab, University College London, 4 Stephenson Way, London NW1 2HE, UK.**
Tel: +44 20 7679 5040; Fax: +44 20 7387 3496;
E-mail: dswallow@hgmp.mrc.ac.uk
Received 12 December 2000; revised 26 January 2001; accepted 1 February 2001

glycosylation. The repeat units vary in length from 24 nucleotides in *MUC5AC* to 507 nucleotides in *MUC6* and in the extent to which they are conserved within the array.

Most of the mucin genes exhibit a high degree of genetically determined polymorphism, due to Variation in Number of Tandem Repeats (VNTR) in the TR domain.[6,11–15] The substantial allelic differences in length of the mucins will not only affect the number of glycosylation sites but is likely to have an effect on other physicochemical properties. Both may have functional consequences and lead to differences in disease susceptibility. An association between short alleles of *MUC1* and gastric cancer has been reported,[16] and our group has recently shown an increase in mean *MUC2* allele length in atopic patients who are free of asthma compared both with atopic patients with asthma and compared with healthy controls.[17]

MUC7 encodes a small salivary mucin, and also shows polymorphism, though to date only two alleles have been reported.[18] *MUC7* is located on chromosome 4q13-q21 and encodes a low M_r secreted mucin (150–220 kDa), found in saliva, previously called MG2.[5] It is thought to function in a protective capacity by promoting the clearance of bacteria in the oral cavity and to aid in mastication, speech and swallowing.[19–21] It has been shown to interact with a variety of bacteria, including four strains of *streptococci*,[21] *Actinobacillus actinomycetemcomitans*,[22] *Staphylococcus aureus* and *Pseudomonas aeruginosa*.[18] *MUC7* has been shown to be expressed in the bronchial tree as well as in the salivary glands.[18,23] It has been detected in fetal trachea and bronchi from the 23rd week of gestation in occasional cells in the submucosal glands and in serous cells in adults.[18,23] The evidence of the role of MUC7 in bacterial binding and its pattern of expression make *MUC7* a candidate gene for involvement in susceptibility to oral and chest disease.

MUC7 contains unique 5′ and 3′ regions (with five potential sites for *N*-glycosylation) and a central tandem repeat domain, which in the most common allele comprises six tandem repeats each of 69 nucleotides (23 amino acids), with many potential sites for *O*-glycosylation and makes up 18% of the coding region. These repeats are very similar but not identical, incorporating between one and seven nucleotide substitutions between them (Figure 1). PCR amplification across the tandem repeats in 14 individuals revealed the presence of a less common allele with five tandem repeats.[18]

The first aim of this study was to determine whether there was any further polymorphism of the *MUC7* TR region. We looked for person-to-person sequence variation within the individual tandem repeats, as well as variation in tandem repeat number in a larger group of Northern Europeans than previously tested, and in other population groups. The second aim was to examine the allele distribution in patients with asthma. We have collected a series of samples from age and sex matched atopic patients with and without asthma, specifically to test the susceptibility genes for development of asthma in individuals already susceptible through atopy. These have previously been tested for polymorphism of the series of other mucin genes known to be expressed in the bronchial tree.[17]

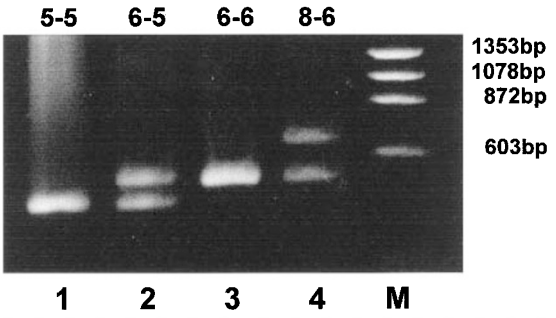

Figure 1 Agarose gel (2%) of PCR products showing the allelic length variation of MUC7. Lane 1: homozygous *MUC7 5-5*; lane 2: heterozygous *MUC7 6-5*; lane 3: homozygous *MUC7 6-6*; lane 4: heterozygous *MUC7 8-6*; M: ϕX174RF DNA/*Hae*III Fragments (Gibco BRL, Paisley, UK).

Previous studies[24] have shown sample to sample differences in electrophoretic mobility on SDS–PAGE of the MUC7 glycoprotein from different salivary glands in the same individuals, which was attributed to heterogeneity of glycosylation, but the effect of the genetically determined polymorphism on mobility was not determined in that study. Our third aim was thus to determine whether the allelic forms of the glycoprotein show the predicted size/mobility differences. Here we show that allelic variation of the MUC7 glycoprotein can be detected but that this is at least partially masked by the heterogeneity of glycosylation.

Materials and methods

Population groups tested

Samples were obtained from people originating from several different population groups. The majority were resident in the UK, in many cases their families originating from elsewhere, and some were obtained from their country of origin. Classification was made on the basis of the country of birth of the immediate ancestors. Those classified as Northern European were native British, French, Dutch, and Danish, Southern European were Greek, Italian, Portuguese, Polish, Spanish; Middle Eastern were Arabic, Iraqi, Israeli, Saudi-Arabian, Iranian; Indian subcontinent were British Asian, Pakistani, Indian and Sri Lankan, and North Indians from Singapore. African were Afro-Caribbean and African resident in the UK and East Asian were Japanese and Chinese.

The hospital patients were from St Mary's Hospital Chest clinic. These were classified into three groups: 102 atopic asthmatics (including 83 Northern Europeans), 14 non-

atopic asthmatics (including 10 Northern Europeans), and 68 atopic non-asthmatic individuals (including 58 Northern Europeans). All patients and 24 of the healthy controls (laboratory personnel) completed detailed questionnaires of past and present medical history and allergies. Eleven of the healthy Northern European volunteers were classified as non-atopic on the basis of their questionnaires. Atopy was defined by positive skin prick tests (two or more positive reactions greater than the histamine control or one very large reaction, greater than 6 mm), and generally confirmed by raised IgE levels (more than 120 IU/l). Full spirometry (Peak expiratory flow rate, PEFR; Forced expiratory volume, FEV1; and Vital capacity, VC) was done on all the atopic and asthmatic individuals. Patients were classified on the basis of their medical history, obtained from their notes, and questionnaire response as well as spirometry and atopy tests. Asthma was diagnosed clinically as follows: on the basis of a clear and characteristic history of episodic breathlessness, chest tightness and wheeze either occurring spontaneously or as a result of exposure to characteristic inciting agents (allergens, irritants, exercise or infection) and spontaneous resolution, or a resolution following bronchodilator therapy; and/or demonstration of a 20% or greater fluctuation in serial peak flow measurements; and/or 20% or greater improvement in FEV1 or peak flow following administration of bronchodilator in the clinic. The non-asthmatics comprised 38 patients referred for allergic rhinitis, 10 for urticaria, seven for eczema and three for food allergies and had no recent or confirmed medical history of asthma. From the patient cohort we previously selected 50 pairs of age and sex matched atopic asthmatic and atopic non-asthmatic individuals of Northern European descent.[17] Ethical committee permission was obtained for this study from Parkside Health Authority (EC no 2893).

DNA extraction

Blood samples or buccal cell swabs[25] were taken, with informed consent, from healthy individuals and patients with asthma/atopy, and the DNA was extracted using a Puregene kit (Flowgen, Leicestershire, UK) or by other standard techniques.

PCR

The *MUC7* polymorphism was examined by PCR using primers designed to amplify the VNTR region, such that an allele containing six tandem repeats (*MUC7*6*) should yield a PCR product of 590 bp while five tandem repeats (*MUC7*5*) should yield a product of 521 bp. The sense primer 5′-GTAGCTACATTAGCACCAGTG-3′ and antisense primer 5′-TTCAGAAGTGTCAGGTGCAAG-3′ are located at positions 547–567 and 1048–1068 on the cDNA (accession number L13283). Each PCR reaction contained 10 μl 10× PCR buffer (500 mM KCl, 100 mM Tris-HCl pH 9.0, 15 mM $MgCl_2$, 0.1% Triton®X-100), 10 μl 2 mM dNTPs, 1 μM sense and antisense primers, 1 μl (0.5 μg) DNA, 0.5 μl (2.5 U) Taq polymerase (Promega, Southampton, UK), and water to a final volume of 100 μl. Following a 2-min denaturation at 94°C, 25 cycles were performed consisting of a 30-s denaturation at 95°C, 30 s annealing at 60°C, 30 s extension at 72°C and a final extension of 6 min at 72°C. The PCR products were subjected to electrophoresis on 2% agarose gels and visualised with ethidium bromide staining under UV light.

DNA sequencing

The bands representing five, six or eight tandem repeats were excised from 2% agarose gels and DNA was extracted using QIAquick gel extraction kit (Qiagen, West Sussex, UK). DNA sequencing was performed using the ABI PRISM BigDye Terminator Cycle Sequencing Ready Reaction Kits (Applied Biosystems, Warrington, UK). For this, 5 μl DNA, 4 μl Ready Reaction mix, 1 μl primer (4 pmol/μl) and 10 μl sterile water were mixed and overlaid with 40 μl oil and subjected to 25 cycles at 96°C for 30 s and 60°C for 4 min 15 s, and electrophoresis on a 377 ABI DNA sequencer (Applied Biosystems,). The sequencing reactions were analysed using Sequencing Analysis and Sequence Navigator (Applied Biosystems).

Saliva collection, SDS–PAGE and Western blotting

Saliva from submandibular glands was collected as described previously.[26] Samples were mixed 1:1 with 2× sample buffer (50 mM Tris-HCl pH 6.8, 2% SDS, 10 mM DTT, 12.5% glycerol, bromophenol blue) and boiled for 5 min. The concentration of MUC7 was estimated by ELISA as described previously and approximately equal amounts were loaded onto the gel.[24] SDS–PAGE was performed on 3–8% Tris-acetate gels (NuPAGE™) using the XCell sure-lock™ and PowerEase500™ apparatus (NOVEX, CA, USA), at 150 V, 120 mA, or on 7.5% Tris-glycine gels using PhastSystem™ (Pharmacia, Buckinghamshire, UK), according to the manufacturer's manual or on 7% gels (Laemmli buffer) using the mini-PROTEANII™ apparatus (BioRad, Hertfordshire, UK). The glycoproteins were stained with periodic acid Schiff's (PAS) reagent followed by silver staining or transferred onto nitrocellulose and probed with a specific polyclonal antibody (CpMG2) raised against a C-terminal peptide of MUC7.[24] The bound antibody was visualised using diaminobenzoate or chloro-1-naphthol as substrate for the detection of the peroxidase linked secondary antibody. Desialylation was performed with 0.5 U/ml sialidase from *Clostridium perfringens* (EC 3.2.1.18, Roche) in 50 mM sodium acetate pH 5.5 at 37°C for 16 h. The efficiency of treatment was verified by observing a shift in migration of MUC7 on SDS–PAGE and loss of reactivity in Western analysis using sialic acid recognising lectins MAA (*Maackia amurensis*) and SNA (*Sambucus nigra*) and mAB E9 recognising a sialidase sensitive epitope on sialyl-Lewis^a.[24,25]

Results

Allelic variation in the number of tandem repeats identified by PCR

DNA was tested from 375 individuals from a variety of ethnic backgrounds but crudely classified into six population groups. The two alleles that had been found previously[18] were found in all the groups and an allele containing eight tandem repeats (*MUC7*8*) was identified (in one British atopic asthmatic individual with genotype 8-6). PCR products illustrating the four deduced genotypes are shown in Figure 1.

The allele frequencies of *MUC7* are shown in Table 1. *MUC7*6* is the most common allele in all populations, with *MUC7*5* frequency varying from 0.05 in Africans to 0.22 in the East Asians. Only three individuals were homozygous for the five tandem repeat allele. However this low frequency of detection of homozygotes was compatible with expectations for populations in Hardy Weinberg equilibrium.

Allelic variation detected by sequence analysis

The tandem repeat arrays of 57 alleles (22 *MUC7*5* and 34 *MUC7*6* and one *MUC7*8*) were sequenced (Table 2). The *MUC7*6* sequence was the same as that published, in all cases.[18] We have named these tandem repeats TR1 to TR6 (Figure 2). All but one of the 22 *MUC7*5* alleles were found to lack TR4 and had the structure TR12356. One *MUC7*5* allele (in the African series) showed a different structure, with an arrangement in which the first two TRs are duplicated and the final repeat was like TR2 until position 65 of the repeat, which was adenine as in TR6 (TR12127). *MUC7*8* has the composition TR12343456.

Allelic variation in asthma and atopy

Initially we compared the 50 age- and sex-matched pairs of atopic individuals with and without asthma that were used in our previous study of other *MUC* genes in chest disease.[17] This cohort was carefully selected to include only UK residents of Northern European origin. The frequency of the *MUC7*5* allele was lower in the asthmatic individuals than the non-asthmatic individuals (0.05 *vs* 0.12; P=0.054, Fishers exact test). Since we had actually collected larger numbers of asthmatic patients, the full Northern European series of chest clinic patients (which includes non-atopic asthmatic individuals) and also a few non-atopic, non-asthmatic individuals were tested. The results are shown in Table 3. The difference in *MUC7*5* occurrence between the atopic asthmatic (mean age 39.8) and atopic non-asthmatic (mean age 35.4) groups reached statistical significance (P=0.014, Fishers Exact test). Comparison of all atopic individuals with all non-atopic, and all asthmatic with all non-asthmatic, showed that it was the asthmatic group that had reduced *MUC7*5* frequency (P=0.009) (OR=3.04, CI 1.10–5.61).

Detection of the polymorphism by SDS–PAGE electrophoresis of the mucin glycoprotein

Our previous studies showed heterogeneity of glycosylation, which may vary in the different salivary glands and also from

Table 1 Frequencies of *MUC7**8, *MUC7**6 and *MUC7**5 alleles in populations of different ethnic origin

MUC7 phenotype	*African*	*Indian Subcontinent*	*Middle Eastern*	*S. European*	*N. European*	*East Asian*
5[a] (55)	0	0	0	1	1	1
56	3	14	1	7	40	15
6[a] (66)	26	41	7	34	160	23
68	0	0	0	0	1	0
Total no. individuals	29 (11)	55 (12)	8 (7)	42 (4)	202 (151)	39 (1)
Total no. alleles	58	110	16	84	404	78
Frequency of *MUC7*5* allele ± SE	0.052 ± 0.029	0.127 ± 0.032	0.063 ± 0.061	0.107 ± 0.034	0.104 ± 0.015	0.218 ± 0.047

The total numbers include healthy volunteers and chest clinic patients (numbers of patients shown in parenthesis). [a]Gel phenotype, assumed genotype in parenthesis. SE: standard error (the standard error is calculated assuming allele frequencies (*MUC7*6* and *MUC7*5*) are binomially distributed using the equation $SE=\sqrt{(pq/2n)}$). n: number of individuals tested.

Table 2 Number of samples from which alleles were sequenced

	African	*Indian Subcontinent*	*Middle Eastern*	*S. European*	*N. European*	*East Asian*
55	0	0	0	0	0	1
56 (5/6*)	2/3[a]	5/4[a]	1/1[a]	4/4[a]	5/5[a]	4/5[a]
66	1	1	1	1	1	1
68 (6/8*)	0	0	0	0	0/1[a]	0

[a]The numbers to the left and right of the slash in each box represent the numbers of the particular allele (5, 6 or 8) sequences – ie both alleles were not always completed in each heterozygote. Individuals of *MUC7* phenotypes 6 and 5 were each assumed to be homozygous for allele length meaning that two alleles were sequenced together.

person to person[24] but the effect of allelic mobility differences due to MUC7 TR variation was not investigated. In this study we have compared the salivary MUC7 from 12 different individuals of known *MUC7* genotype (6 *MUC7 6-6*, 5 *MUC7 6-5* and 1 *MUC7 5-5*) under a variety of different electrophoretic conditions, and using either mixed submandibular/sublingual saliva or pure submandibular saliva. There were clear differences in the mobility of the MUC7

Table 3 Frequencies of *MUC7**8, *MUC7**6 and *MUC7**5 alleles in Northern European healthy volunteers and chest disease patients with different atopy and asthma status. (a) shows numbers of individuals of each inferred genotype and atopy and asthma status. (b) shows Fishers exact test (2 by 2) results and odds ratios

(a)

	Non-asthmatic		*Asthmatic*	
	Non atopic[a]	*Atopic*[a]	*Non atopic*[a]	*Atopic*[a]
Number of individuals	11	58	10	83
55	0	0	0	0
56	4	16	2	9
66	7	42	8	73
68	0	0	0	1

(b)

Comparison	*2 by 2 P values*[b]	*Odds ratio*[c]
Atopic asthmatic *vs* atopic non-asthmatic	0.014	3.13 with a 95% confidence interval of 1.01 to 6.10
Asthma *vs* non asthma	0.009	3.04 with a 95% confidence interval of 1.10 to 5.61
Atopy *vs* no atopy	0.199	1.86 with a 95% confidence interval of 0.88 to 7.02

[a]Average age atopic non-asthmatic (*n*=58)=35.4 years; average age atopic asthmatic (*n*=83)=39.8 years, [b]No correction for multiple tests, [c]The estimated odds ratio for disease status with *MUC7**5 allele carrier status.

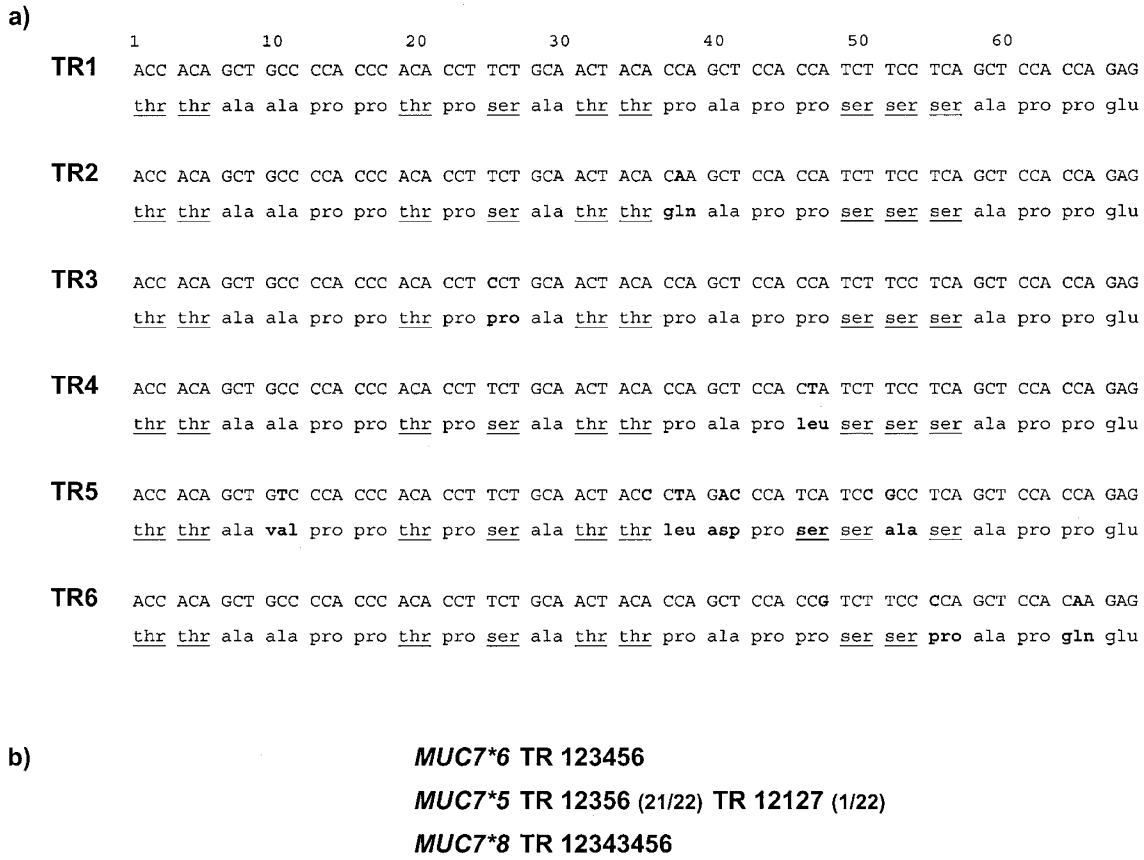

Figure 2 **a** Sequence of the tandem repeats (TRs) in MUC7*6 genomic DNA. **b** Structure of the TR domain of the different *MUC7* alleles. The nucleotide and amino acid differences between TR1 and TRs 2 – 6 are indicated in bold. Potential *O*-glycosylation sites are underlined.

glycoprotein from the different individuals, and the relative migration in general reflected their MUC7 tandem repeat genotype, the *MUC7 6-5* heterozygotes sometimes showing two distinct bands, but samples of the same genotype did not necessarily show identical mobility. A representative example is shown in Figure 3. Note that the MUC7 of the three *MUC7 6-6* individuals differs in mobility. We also explored the effect of sialidase treatment of the saliva, on this person-to-person variation, since it seemed possible that this would remove these differences and improve the distinction between the allelic forms of the protein. The efficacy of the sialidase was confirmed by loss of reactivity towards sialic acid recognising lectins and the monoclonal antibody E9 (data not shown). Removal of the negatively charged sialic acid residues from the MUC7 glycoprotein resulted in a decrease in mobility of the MUC7 glycoprotein of all individuals. As can be seen in Figure 3 the heterozygous sample showed increased separation of the glycoprotein products of the *MUC7*6* and *MUC7*5* alleles. However the effect of sialidase was clearly variable in the different samples (Figure 3). The small shift in mobility of MUC7 in the samples shown in lanes 4 and 9 is particularly noticeable.

Discussion

The previously published allelic variation of *MUC7* within a group of 14 individuals has been confirmed and extended in this study by testing a cohort of 375 subjects of different ancestries. Both the *MUC7*6* and *MUC7*5* alleles were identified in each of the populations we tested, the *MUC7*5* allele being the less frequent in all groups, but varying somewhat in frequency from population to population. A single allele consisting of 8TRs, *MUC7*8*, was found in a British asthmatic individual, but no further repeat number variation was observed. This relatively low level of length variability contrasts with most of the other human mucin genes[11–13,15] with the exception of the invariant *MUC5B*[11] which encodes the large salivary mucin, previously called MG1. The more complex repeat structure of *MUC5B* probably restricts repeat number mutations in this gene. Possibly, in the case of *MUC7*, the lower tandem repeat number is genetically more stable than the larger number of repeat units found in the other mucins.[8,11,27] Whatever the mechanism, it is noteworthy that the salivary mucins show less genetically determined length variability of their polypeptide backbones than mucins in other organs. Sequence analysis also showed that there is very little person-to-person sequence variation within the tandem repeat array, and revealed rearrangement in just one of 22 *MUC7*5* alleles. This was one of only two *MUC7*5* alleles tested in the African series.

Sequence analysis shows that TR4 is missing in the *MUC7*5* allele, meaning that the tandem repeat region is shorter by 16.6% (23 amino-acids) and has nine less potential O-glycosylation sites (five threonine and four serine residues). *MUC7*8* has 17 additional potential *O*-glycosylation sites (10 threonine and seven serine residues) while the number of potential *O*-glycosylation sites in the rare *MUC7*5* allele are not altered. The negative charges of terminal sialic acid residues are important in determining the mobility of mucins on SDS gels, particularly since there is limited binding of SDS to the highly glycosylated domains. However, if the size to charge ratio is constant, the sialic acid residues behave just like SDS in causing a size separation, so that allelic glycoproteins which differ in numbers of tandem repeats show correlated mobility differences.[12,28] In this case, the migration of the mature MUC7 glycoprotein appears to reflect both the size, and charge differences due to the sialic acid residues, better separation being achieved after removal of sialic acid residues. However our results suggest charge

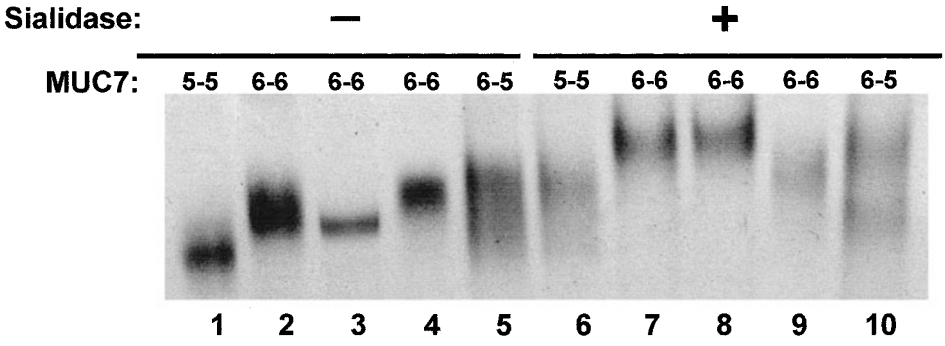

Figure 3 Representative experiments showing the distinction of MUC7 glycoprotein mobility on SDS–PAGE and the effect of sialidase treatment. Submandibular saliva from *MUC7 6-6, 6-5* and *5-5* individuals was subjected to SDS–PAGE. Equal concentrations of MUC7 were applied to a 3–8% gradient Tris-acetate gel, and after electrophoresis stained with Schiff's reagent as described in Materials and methods. Samples 1–5 were untreated. Samples 6–10 were treated with *Clostridium perfringens* sialidase. The removal of sialic acid was verified using sialic acid recognising lectins (MAA and SNA) and antibodies (anti-sialyl Lewis^a) as detailed in Materials and methods (not shown).

heterogeneity which cannot be accounted for by sialic acid alone. Different glycoforms of MUC7 have previously been described which differ in their content of terminal sugars, sialic acid and fucose.[29] It seems possible that sulphate or O-acetylated sialic acid may also sometimes be present.

Here we report a significantly lower frequency of the *MUC7*5* allele in individuals with atopic asthma, in comparison with the other groups. It will be important to determine whether this difference can be substantiated in other cohorts. It seems possible that the association relates to allelic differences in interactions with bacteria, since the glycosylated domain is thought to be responsible, at least in part, for the bacterial binding that allows bacteria to be cleared from the epithelial surfaces.[30] Studies are therefore in progress to examine the binding of bacteria to the different allelic forms of MUC7.

The association could alternatively be the effect of a linked polymorphic locus but this could still be within the MUC7 gene itself. The N-terminal part of the polypeptide chain of MUC7 has also been implicated in the binding to bacteria[21] and this region has been shown to have Candidacidal activity.[31,32] A search for possible functional polymorphism in this region, or elsewhere in the gene, such as a promoter element would be worthwhile.

Acknowledgments

We are grateful to the British Lung Foundation for their support, to EC contract BMH4-CT98-3222 for travel funds, to Ed Hollox, N Saha and Mark Poulter for supplying DNA samples, and to Wendy Pratt for help with the ABI sequencing.

References

1 Wickstrom C, Davies JR, Eriksen GV, Veerman EC, Carlstedt I: MUC5B is a major gel-forming, oligomeric mucin from human salivary gland, respiratory tract and endocervix: identification of glycoforms and C-terminal cleavage. *Biochem J* 1998; **334**: 685–693.
2 Hovenberg HW, Davies JR, Herrmann A, Linden CJ, Carlstedt I: MUC5AC, but not MUC2, is a prominent mucin in respiratory secretions. *Glycoconj J* 1996; **13**: 839–847.
3 Herrmann A, Davies JR, Lindell G *et al*: Studies on the 'insoluble' glycoprotein complex from human colon. Identification of reduction-insensitive MUC2 subunit oligomers and C-terminal cleavage. *J Biol Chem* 1999; **274**: 15828–15836.
4 Toribara NW, Roberton AM, Ho SB *et al*: Human Gastric Mucin. Identification of a unique species by expression cloning. *J Biol Chem* 1993; **268**: 5879–5885.
5 Bobek LA, Tsai H, Biesbrock AR, Levine MJ: Molecular cloning, sequence, and specificity of expression of the gene encoding the low molecular weight human salivary mucin (MUC7). *J Biol Chem* 1993; **268**: 20563–20569.
6 Lapensee L, Paquette Y, Bleau G: Allelic polymorphism and chromosomal localization of the human oviductin gene (MUC9). *Fertil Steril* 1997; **68**: 702–708.
7 Crawley SC, Gum Jr JR, Hicks JW *et al*: Genomic organization and structure of the 3′ region of human MUC3: alternative splicing predicts membrane-bound and soluble forms of the mucin [In Process Citation]. *Biochem Biophys Res Commun* 1999; **263**: 728–736.
8 Moniaux N, Nollet S, Porchet N, Degand P, Laine A, JP: A Complete sequence of the human mucin MUC4: a putative cell membrane-associated mucin. *Biochem J* 1999; **338**: 325–333.
9 Williams SJ, McGuckin MA, Gotley DC, Eyre HJ, Sutherland GR, Antalis TM: Two novel mucin genes down-regulated in colorectal cancer identified by differential display. *Cancer Res* 1999; **59**: 4083–4089.
10 Williams SJ, Munster DJ, Quin RJ, Gotley DC, McGuckin MA: The MUC3 gene encodes a transmembrane mucin and is alternatively spliced. *Biochem Biophys Res Commun* 1999; **261**: 83–89.
11 Vinall LE, Hill AS, Pigny P *et al*: Variable number tandem repeat polymorphism of the mucin genes located in the complex on 11p15.5. *Hum Genet* 1998; **102**: 357–366.
12 Swallow DM, Gendler S, Griffiths B *et al*: The hypervariable gene locus PUM, which codes for the tumour associated epithelial mucins, is located on chromosome 1, within the region 1q21-24. *Ann Hum Genet* 1987; **51**: 289–294.
13 Fox M, Lahbib F, Pratt W *et al*: Regional localization of the intestinal mucin gene MUC3 to chromosome 7q22. *Ann Hum Genet* 1992; **56**: 281–287.
14 Toribara NW, Gum JJ, Culhane PJ *et al*: MUC-2 human small intestinal mucin gene structure. Repeated arrays and polymorphism. *J Clin Invest* 1991; **88**: 1005–1013.
15 Gross MS, Guyonnet-Duperat V, Porchet N, Bernheim A, Aubert JP, Nguyen VC: Mucin 4 (MUC4) gene: regional assignment (3q29) and RFLP analysis. *Ann Genet* 1992; **35**: 21–26.
16 Carvalho F, Seruca R, David L *et al*: MUC1 gene polymorphism and gastric cancer–an epidemiological study. *Glycoconjugate J* 1997; **14**: 107–111.
17 Vinall LE, Fowler JC, Jones AL *et al*: Polymorphism of human mucin genes in chest disease: possible significance of MUC2. *Am J Resp Cell Mol Biol* 2000; **23**: 678–686.
18 Biesbrock AR, Bobek LA, Levine MJ: *MUC7* gene expression and genetic polymorphism. *Glycoconj J* 1997: **14**: 415–422.
19 Amerongen AV, Bolscher JG, Veerman EC: Salivary mucins: protective functions in relation to their diversity. *Glycobiology* 1995; **5**: 733–740.
20 Levine MJ, Reddy MS, Tabak LA *et al*: Structural aspects of salivary glycoproteins. *J Dent Res* 1987; **66**: 436–441.
21 Liu B, Rayment S, Oppenheim FG, Troxler RF: Isolation of human salivary mucin MG2 by a novel method and characterization of its interactions with oral bacteria. *Arch Biochem Biophys* 1999; **364**: 286–293.
22 Groenink J, Lightenberg AJ, Veerman EC, Bolscher JG, Nieuw Amerongen AV: Interaction of the salivary low-molecular weight mucin (MG2) with Actinobacillus actinomycetemcomitans. *Antonie van Leeuwenhoek* 1996; **70**: 79–87.
23 Buisine M-P, Devisme L, Copin M-C *et al*: Developmental mucin gene expression in the human respiratory tract. *Am J Respir Cell Mol Biol* 1999; **20**: 209–221.
24 Bolscher JG, Groenink J, van der Kwaak JS *et al*: Detection and quantification of MUC7 in submandibular, sublingual, palatine, and labial saliva by anti-peptide antiserum. *J Dent Res* 1999; **78**: 1362–1369.
25 Freeman B, Powell J, Ball D, Hill L, Craig I, Plomin R: DNA by mail: an inexpensive and noninvasive method for collecting DNA samples from widely dispersed populations. *Behav Genet* 1997; **27**: 251–257.
26 Veerman EC, van den Keybus PA, Vissink A, Nieuw Amerongen AV: Human glandular salivas: their separate collection and analysis. *Eur J Oral Sci* 1996; **104**: 346–352.
27 Gum JR, Ho JL, Pratt W *et al*: MUC3 Human intestinal mucin: Analysis of gene structure, the carboxyl terminus and a novel upstream repetitive region. *J Biol Chem* 1997; **272**: 26678–26686.
28 Tytgat KM, Swallow DM, Van Klinken BJ, Buller HA, Einerhand AW, Dekker J: Unpredictable behaviour of mucins in SDS/polyacrylamide-gel electrophoresis [letter]. *Biochem J* 1995; **310**: 1053–1054.

29 Ramasubbu N, Reddy MS, Bergey EJ, Haraszthy GG, Soni SD, Levine MJ: Large-scale purification and characterization of the major phosphoproteins and mucins of human submandibular-sublingual saliva. *Biochem J* 1991; **280**: 341 – 352.
30 Wu AW, Csako G, Herp A: Structure, biosynthesis, and function of salivary mucins. *Mol Cell Biochem* 1994; **137**: 39 – 55.
31 Gururaja TL, Levine JH, Tran DT *et al*: Candidacidal activity prompted by N-terminus histatin-like domain of human salivary mucin (MUC7)1. *Biochim Biophys Acta* 1999; **1431**: 107 – 119.
32 Liu B, Rayment SA, Gyurko C, Oppenheim FG, Offner GD, Troxler RF: The recombinant N-terminal region of human salivary mucin MG2 (MUC7) contains a binding domain for oral Streptococci and exhibits candidacidal activity. *Biochem J* 2000; **345**: 557 – 564.