

Evolution of the human MUC1 oncoprotein

SEKHAR DURAISAMY, TURNER KUFE, SELVI RAMASAMY and DONALD KUFE

Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA 02115, USA

Received May 24, 2007; Accepted July 3, 2007

Abstract. The mucin (MUC) family consists of secreted and membrane-bound forms. The transmembrane mucin 1 (MUC1) is a heterodimer that is aberrantly overexpressed by diverse human carcinomas and certain hematologic malignancies. The MUC1 N-terminal (MUC1-N) and C-terminal (MUC1-C) subunits are generated by autocleavage within a SEA domain. The MUC1 cytoplasmic domain (MUC1-CD) located downstream of the SEA domain is sufficient for the induction of anchorage-independent growth and tumorigenicity; however, no information is available regarding the origin of these transforming sequences. Previous work demonstrated that, except for the SEA domain, MUC1 has no sequence homology with other membrane-bound mucins. The present results demonstrate that MUC1-CD evolved from repeat regions in the MUC5B secreted mucin. We also show that MUC1 sequences upstream to the SEA domain emerged from MUC5B. These findings indicate that both the MUC1-N and MUC1-C subunits evolved from secreted gel-forming mucins and that the MUC1-CD oncogenic function emerged by diversification after evolution from MUC5B.

Introduction

The mucin-type glycoproteins are classified by the presence of tandem repeat structures that contain a high proportion of prolines, threonines and serines and are extensively modified by *O*-glycosylation (Human Genome Organization Gene Nomenclature Committee; <http://www.hugo-international.org/hugo/>). The human MUC family consists of 20 members that have been subclassified as secreted and membrane-bound forms (1). The secreted mucins (i.e., MUC2, MUC3, MUC5AC, MUC5B and MUC6) form a physical gel barrier that protects epithelial cells lining the respiratory and gastrointestinal tracts and the ductal surfaces of specialized organs such as the liver, pancreas and kidney. The membrane-bound

mucins (i.e., MUC1, MUC3, MUC4, MUC12, MUC13, MUC16 and MUC17) also contribute to formation of the protective mucous gel through ectodomains of *O*-glycosylated tandem repeats that extend from the apical cell surface. The membrane-bound mucins typically have a sea urchin sperm protein, enterokinase and agrin (SEA) domain that resides between the glycosylated ectodomain and the transmembrane domain. Autoproteolysis of the MUC1 SEA domain results in the formation of an N-terminal ectodomain and a C-terminal transmembrane subunit that in turn form a stable non-covalent heterodimer (2-4).

MUC1 localizes to the apical borders of normal secretory epithelial cells (5). However, with transformation and loss of polarity, MUC1 is expressed at high levels over the entire surface and in the cytosol of carcinoma cells (5). Importantly, overexpression of MUC1 induces transformation (6) and resistance to stress-induced apoptosis (7-12). The MUC1 N-terminal ectodomain (MUC1-N) contains variable numbers of highly glycosylated 20 amino acid tandem repeats (13,14). The MUC1 C-terminal subunit (MUC1-C) consists of a 58 amino acid extracellular domain, a 28 amino acid transmembrane domain and a 72 amino acid cytoplasmic tail (15). MUC1-C is targeted to the nucleus (6,10,16-18) and mitochondria (8,19). In addition, MUC1-C stabilizes the Wnt effector, β -catenin, through direct binding of the MUC1-C cytoplasmic domain (MUC1-CD) and the β -catenin Arm repeats (20). Overexpression of MUC1-CD is sufficient for the induction of transformation and the interaction between MUC1-CD and β -catenin is in part responsible for this response (20). Other studies have demonstrated that MUC1-CD functions as a substrate for glycogen synthase 3 β (21), c-Src (22), protein kinase C δ (23) and c-Abl (12). These findings have indicated that, in contrast to the N-terminal mucin-type ectodomain, MUC1-C transduces signals from the cell membrane to the interior of the cell that confer growth and survival responses. The available evidence indicates that MUC1-CD functions like a chaperone that holds client proteins in configurations that dictate activity (10,18,20).

The MUC family members have been grouped according to a biophysical structure rather than having evolved from common ancestral genes (24). In this regard, a recent analysis has shown that MUC1 homologs are restricted to mammalian species (25). In addition, with the exception of a SEA domain, MUC1 has no homology with the other membrane-bound mucins (25). Based on the overexpression of MUC1 in diverse human malignancies and the importance of MUC1-CD to transformation, the present studies have addressed the evolution of these sequences. The results demonstrate that MUC1-CD

Correspondence to: Dr Donald Kufe, Dana-Farber Cancer Institute, Boston, MA 02115, USA
E-mail: donald_kufe@dfci.harvard.edu

Key words: secreted mucins, membrane-bound mucins, MUC1, transformation, MUC5B, MUC2, MUC6

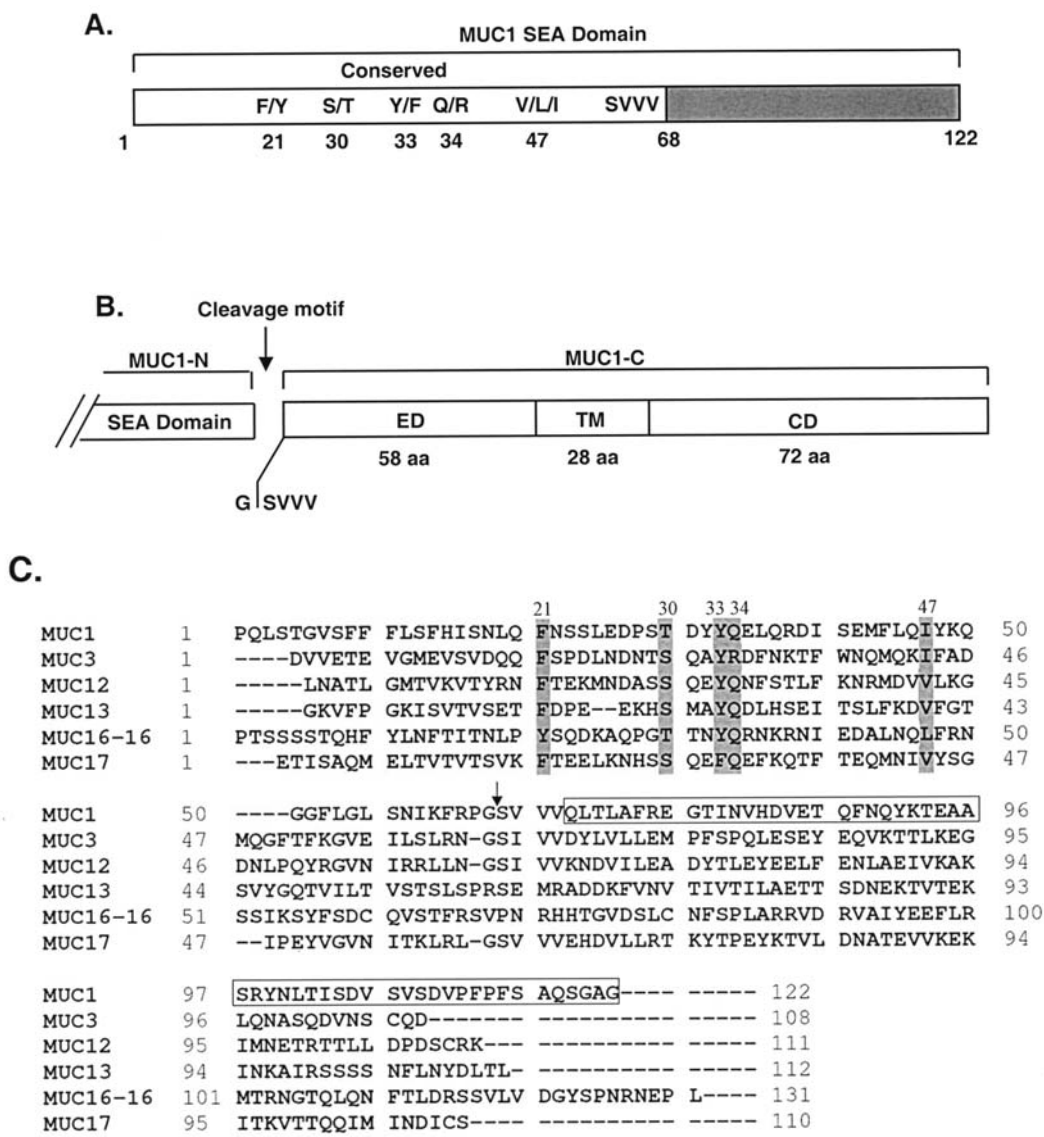


Figure 1. MUC1 SEA domain contains divergent sequences that form the MUC1-C extracellular domain. A, Schema depicting the MUC1 SEA domain with a conserved region of 68 amino acids and a divergent region of 54 amino acids. B, Autoproteolysis of the MUC1 SEA domain between Gly and Ser in the GSVVV motif results in formation of the MUC1-C subunit with an extracellular domain (ED) that includes SVVV and 54 divergent amino acids. TM, transmembrane domain. CD, cytoplasmic domain. C, Alignment of the SEA domains of MUC1 and the indicated transmembrane mucins. The consensus amino acids of the SEA domain are shaded. The arrow denotes the MUC1 autocleavage site. The divergent MUC1 sequences are shaded.

and part of MUC1-N originated from the secreted MUC5B gel-forming mucin.

Materials and methods

Sequence information was retrieved for MUC1, MUC3, MUC12, MUC13, MUC16 and MUC17 mucins from GenBank (26) and the domain annotations were carried out using SMART (Simple Modular Architecture Research Tool) (27). The search for homologs of MUC1 domains was performed using BLAST (Basic Local Alignment Search Tool) (28). The search for homologs of MUC1-CD was also carried out using BLAST against the human non-redundant (NR) protein database. Multiple sequence alignments were created using ClustalW (29) and gaps were removed if present in >50% of species. Visualization and manual adjustments to any mis-

alignments were made using Bioedit (<http://www.mbio.ncsu.edu/BioEdit/bioedit.html>). Phylogenetic trees were constructed using the parsimony method implemented in the protpars program. Bootstrap values were calculated using the seqboot program. Both protpars and seqboot are available in Phylip (30). The trees were visualized using the TreeView program (31).

Results

MUC1 SEA domain contains a divergent carboxy-terminal region. The SEA domain is a ~120 amino acid module that is found in proteins from *Caenorhabditis elegans* to humans (32). SEA modules in all proteins have significant sequence similarity within the initial ~80 amino acids (32). An additional ~40 amino acids that separate the conserved region from

downstream modules have been included in SEA domain alignments (32). For example, the MUC1 SEA domain includes a 68 amino acid conserved region and an additional 54 amino acids (Fig. 1A). The MUC1 protein is cleaved between glycine and serine in a GSVVV motif within the SEA domain (3,33). This autoproteolytic process generates the MUC1 C-terminal subunit (MUC1-C), which consists of a 58 amino acid extracellular domain (MUC1-C/ED), a 28 amino acid transmembrane domain and a 72 amino acid cytoplasmic domain (Fig. 1B). Analysis of the MUC1-C/ED sequences with other proteins containing SEA domains, including the human mucins, demonstrated similarity of consensus sequences highlighted at positions 21, 30, 33-34 and 47 (Fig. 1C). By contrast, there was little similarity after the GSVVV auto-cleavage site (Fig. 1C).

MUC1 cytoplasmic domain evolved from the MUC5B gel-forming mucin. MUC1-C includes a 72 amino acid cytoplasmic domain (MUC1-CD). A BLAST search demonstrated that MUC1-CD has homologous regions with: i) the secreted MUC5B mucin; ii) the RNA binding motif protein, Y chromosome, family 2, member B isoform (RY2B_HUMAN) (34); iii) v-maf musculoaponeurotic fibrosarcoma oncogene homolog B (MAFB) (35); iv) the ankyrin repeat and sterile alpha motif domain containing 3 protein (ANKS3) (36); v) the kin of IRRE-like 1 protein (KIRREL) (37); vi) neural precursor cell expressed, developmentally down-regulated 9 (NEDD9) (38); and vii) insulin receptor substrate 2 (IRS2) (39). Construction of a phylogenetic tree with the putative homologs revealed that MUC1-CD is most closely related to MUC5B as supported by strong bootstrap values (Fig. 2A). To further ascertain the relationship between MUC1-CD and MUC5B, a BLAST search was performed using MUC1-CD as a query against the mammalian MUC proteins. MUC5B was identified as the closest match. In addition, the MUC2 and MUC6 proteins, which are ancestors of MUC5B, were also identified as matches. Construction of a phylogenetic tree confirmed: i) the close association of MUC1-CD and MUC5B; and ii) evolution of MUC5B from MUC2 and MUC6 (Fig. 2B).

MUC1-CD emerged from tandem repeat regions in MUC5B. MUC5B contains irregular Ser-, Thr- and Pro-rich tandem repeats of 29 amino acids in domains designated RI-RV (40). This mucin-type region is interrupted four times by sequences that are repeated and distinct from the RI-RV tandems (40). MUC5B also contains four von Willebrand factor type D (VWD) domains (Fig. 3A). Sequences related to MUC1-CD were identified in a region located between the third and fourth VWD domains (Fig. 3A). Four repetitive domains beginning at amino acids 2181, 2710, 3408 and 3937 exhibited significant alignments with MUC1-CD (Fig. 3B). Construction of a phylogenetic tree demonstrated that the MUC5B(3408-3447) repeat is most closely related to MUC1-CD as compared to the other three repeats (Fig. 3C). These findings indicate that MUC1-CD evolved from MUC5B.

MUC1-N sequences also evolved from the secreted mucins. The MUC5B locus is clustered on chromosome 11p15.5 with genes encoding the secreted MUC6, MUC2 and MUC5AC

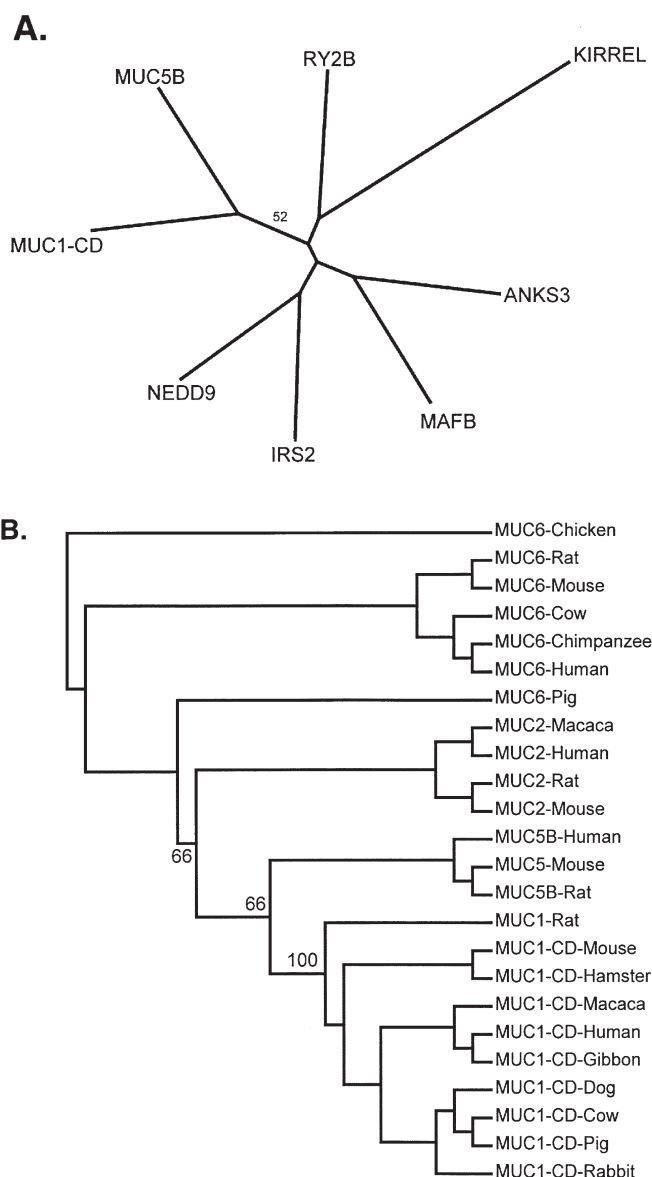


Figure 2. MUC1 cytoplasmic domain is closely related to MUC5B. A, A phylogenetic tree was constructed by the parsimony method using amino acid sequence alignment of MUC1-CD and its putative human homologs, including MUC5B, RNA binding motif protein, Y chromosome, family 2, member B isoform 1 (RY2B_HUMAN), v-maf musculoaponeurotic fibrosarcoma oncogene homolog B (MAFB), the ankyrin repeat and sterile alpha motif domain containing 3 protein (ANKS3), the kin of IRRE like 1 protein (KIRREL), neural precursor cell expressed, developmentally down-regulated 9 (NEDD9) and insulin receptor substrate 2 (IRS2). The number represent the percent bootstrap value that supports branching. B, A phylogenetic tree was constructed by the parsimony method using amino acid sequence alignment of MUC1-CD with MUC5B, MUC2 and MUC6 in different species. The numbers represent the percent bootstrap values that support branching.

secreted mucins (41). Based on genomic organization, sequence similarities and their ordering on chromosome 11, the evidence indicates that these four genes evolved from a common progenitor (42,43). Whereas MUC1-CD is related to MUC5B and its ancestors MUC2 and MUC6, we asked if MUC1-N also evolved from these gel-forming mucins. A BLAST search with MUC1-N against the human NR dataset identified MUC5B, MUC2 and MUC6. Significant similarity

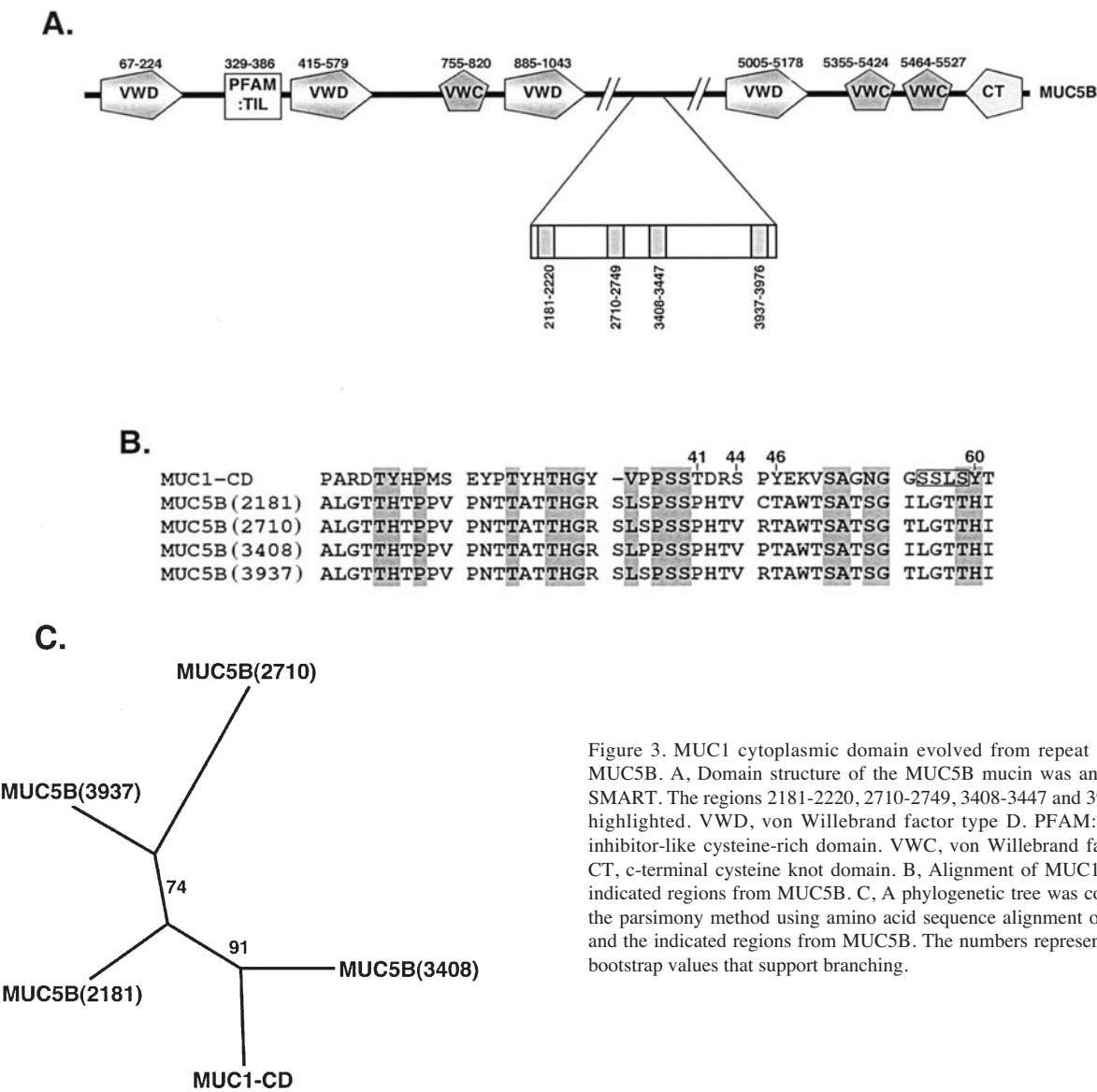


Figure 3. MUC1 cytoplasmic domain evolved from repeat structures in MUC5B. A, Domain structure of the MUC5B mucin was analyzed using SMART. The regions 2181-2220, 2710-2749, 3408-3447 and 3937-3976 are highlighted. VWD, von Willebrand factor type D. PFAM:TIL, trypsin inhibitor-like cysteine-rich domain. VWC, von Willebrand factor type C. CT, c-terminal cysteine knot domain. B, Alignment of MUC1-CD and the indicated regions from MUC5B. C, A phylogenetic tree was constructed by the parsimony method using amino acid sequence alignment of MUC1-CD and the indicated regions from MUC5B. The numbers represent the percent bootstrap values that support branching.

was identified between MUC5B(2941-3212) and the region of MUC1-N between the tandem repeats and the SEA domain (Fig. 4A and B). Construction of a phylogenetic tree demonstrated that the sequences related to MUC1-N are repeated 6 times throughout MUC5B (Fig. 4C). Moreover, the related sequences are shared with repeats in MUC2 and MUC6 (Fig. 4C), supporting their emergence from a common progenitor. The phylogenetic tree demonstrated that the MUC5B(2941-3212) sequences are most closely related to MUC1-N (Fig. 4C). These findings indicate that regions in both MUC1-N and MUC1-C evolved from the secreted gel-forming MUC5B mucin.

Discussion

Evolution of the MUC family. The MUC family of mucins consists of the secreted and membrane-bound forms. MUC1 is a membrane-bound mucin that is aberrantly overexpressed in diverse malignancies (5). Estimates indicate that ~800,000 of the 1.4 million tumors diagnosed annually in the US

overexpress MUC1. Recent studies have demonstrated that the overexpression of MUC1 induces transformation (6) and that the oncogenic sequences reside in the MUC1 C-terminal subunit (20,44). However, no information is available regarding the origin of the MUC1-C extracellular and cytoplasmic domains. The mucins are classified in the same family based on the presence of an extensively *O*-glycosylated tandem repeat structure, rather than a common genetic origin (24). Thus, with the exception of the SEA domain, MUC1 has no sequence similarity with the other membrane-bound mucins (25). The present results extend these observations by demonstrating that sequences downstream of the MUC1 SEA domain G↓SVVV autocleavage site are unrelated to those found in other SEA domain containing proteins. The SEA domain has been described as a ~120 amino acid module, based on the sea urchin sperm protein, enterokinase and agrin in which it was first identified (32). A phase 1 intron at the end of a preceding module has been used to define the beginning of the SEA domain. In MUC1, the SEA module extends for 122 amino acids (32). After autocleavage of

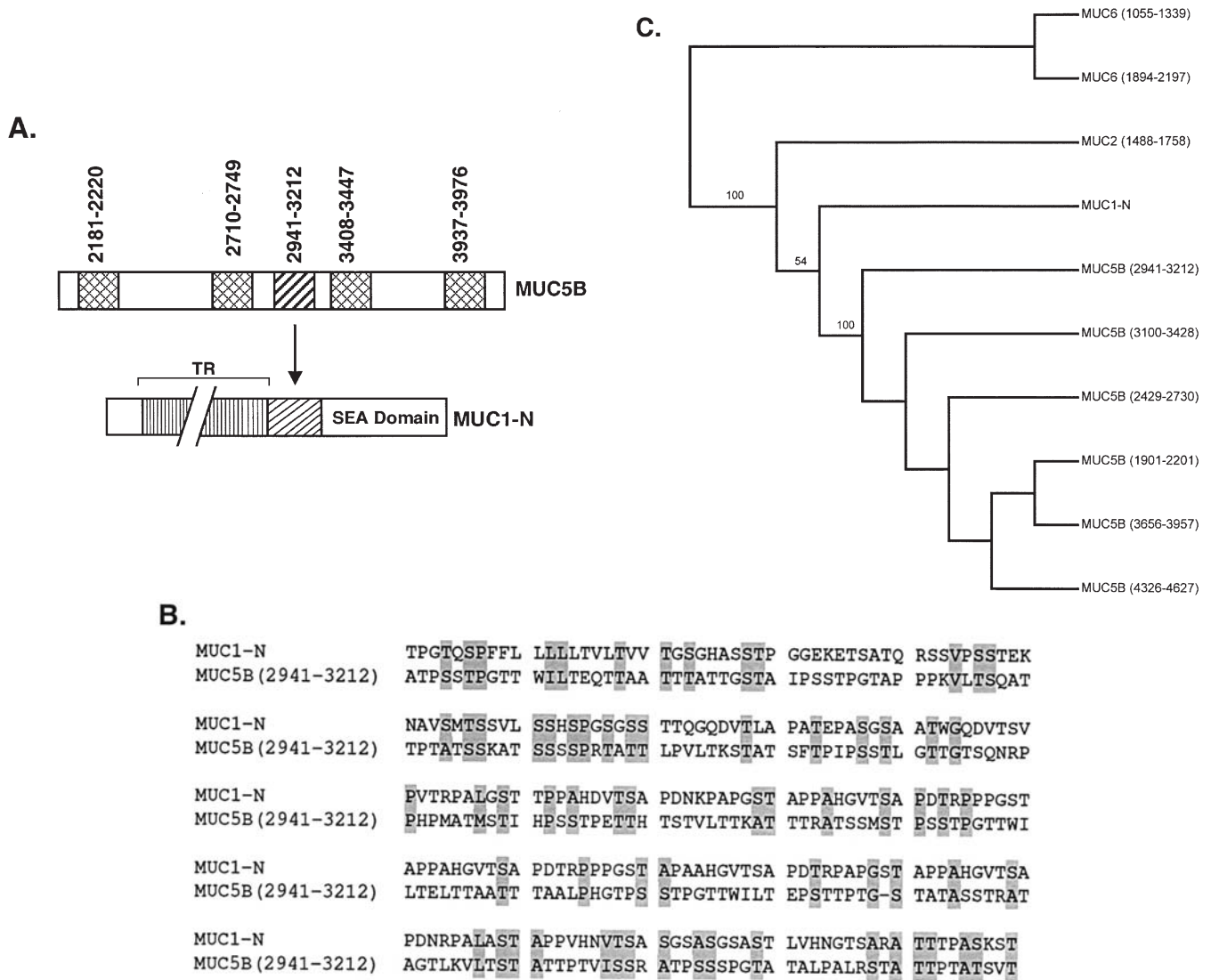


Figure 4. MUC1-N is also related to MUC5B. A, Schema depicting the location of MUC5B(2941-3212) and the related sequence in MUC1-N. TR, tandem repeats. B, Alignment of MUC1-N sequence with MUC5B regions. C, A phylogenetic tree was constructed by the parsimony method using amino acid sequence alignment of MUC1-N, MUC5B and its related proteins, MUC6 and MUC2. The numbers represent the percent bootstrap values that support branching.

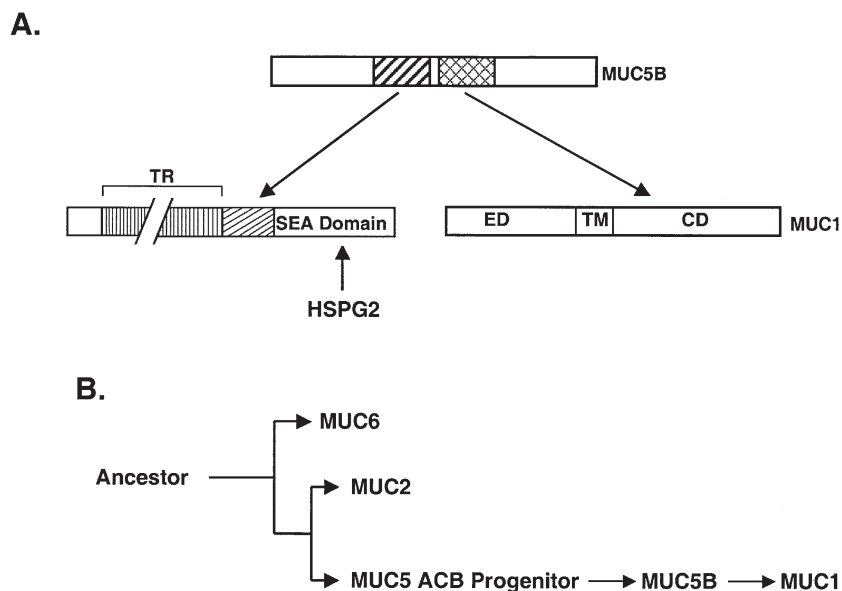


Figure 5. MUC1 evolved from the secreted gel-forming mucins. A, Schema depicting origins of the indicated regions in MUC1-N and MUC1-C. B, Evolution of MUC5AC and MUC5B from a common ancestor and the emergence of MUC1 from MUC5B.

MUC1, a conserved region of the SEA domain (64 amino acids) is located at the carboxy-terminus of MUC1-N (Fig. 5A). By contrast, the divergent sequences following the SVVV motif are located in the extracellular domain of the MUC1-C subunit (Fig. 5A). These findings indicated that the MUC1 region between the conserved SEA domain sequences and the transmembrane domain diverged from corresponding regions in other SEA domain containing proteins.

MUC1-C cytoplasmic domain evolved from MUC5B. The MUC1-CD binds to β -catenin and increases β -catenin levels by blocking its degradation (20). Moreover, MUC1-CD is sufficient to induce transformation by a mechanism that is at least in part dependent on the interaction with β -catenin (20). The present results indicate that MUC1-CD is most closely related to repeat regions found in the secreted MUC5B mucin. Notably, these repeats are not found in the other gel-forming mucins. MUC5B is aberrantly expressed in human gastric and breast cancers (45,46). However, a functional role for MUC5B in transformation is not known. Moreover, previous work has shown that MUC1-CD is phosphorylated on Thr-41, Ser-44, Tyr-46 and Tyr-60 by PKC δ , GSK3 β , c-Src and c-Abl, respectively (12,21-23). Significantly, none of these phosphorylation sites is present in the MUC5B repeats (Fig. 3B). The MUC1-CD SSLS motif (amino acids 56-59), which is important for binding to β -catenin and transformation (20,47), is also not conserved in the MUC5B repeats (Fig. 3B). In addition, the essential tyrosines located in the putative MUC1-CD immunoreceptor tyrosine-based activation motif (ITAM; amino acids 8-23) (48) and the immunoreceptor tyrosine-based inhibitory motif (ITIM; amino acids 44-49) are not present in the MUC5B repeats. These findings indicate that the MUC1-CD oncogenic function arose by diversification after evolution from MUC5B.

MUC1-N also evolved in part from secreted mucins. Our previous work showed that the conserved region of the MUC1 SEA domain emerged from the basement membrane-specific heparin sulphate proteoglycan core protein precursor (HSPG) and clustered with the SEA domains of MUC13, MUC3, MUC12 and MUC17 (25). Otherwise, MUC1 has no homology with the membrane-bound mucins (25). Moreover, recent work has indicated that, in contrast to the secreted gel-forming mucins, the transmembrane mucins are a more recent development in evolution (49). The present results demonstrate that the MUC1-N region between the tandem repeats and the SEA domain evolved from the secreted MUC5B, MUC2 and MUC6 mucins that cluster on chromosome 11. Notably, the MUC1-N sequences are most closely related to those in MUC5B. The present results further demonstrate that MUC5B, and not MUC6 or MUC2, contains repeat structures related to the MUC1 cytoplasmic domain. The MUC5B repeat (amino acids 3408-3447) is most closely related to MUC1-CD and is just downstream to the MUC5B region with homology to MUC1-N (Fig. 5A). Thus, sequences in both MUC1-N and MUC1-C appear to have emerged from adjacent regions in MUC5B (Fig. 5A). Other work has supported the evolution of the *MUC5AC* and *MUC5B* genes from a common *MUC5ACB* ancestor (Fig. 5B), which in turn arose from a progenitor of the *MUC6* and *MUC2* genes (43).

The present findings indicate that MUC1 evolved in part from MUC5B (Fig. 5B) and thereby this lineage of the secreted gel-forming mucins.

Acknowledgements

This work was supported by Grant CA97098 awarded by the National Cancer Institute. The authors acknowledge Kamal Chauhan for technical support.

References

1. Taylor-Papadimitriou J: Report on the first international workshop on carcinoma-associated mucins. *Int J Cancer* 49: 1-5, 1991.
2. Ligtenberg MJ, Kruijschaar L, Buijs F, van Meijer M, Litvinov SV and Hilken J: Cell-associated episialin is a complex containing two proteins derived from a common precursor. *J Biol Chem* 267: 6171-6177, 1992.
3. Macao B, Johansson DG, Hansson GC and Hard T: Auto-proteolysis coupled to protein folding in the SEA domain of the membrane-bound MUC1 mucin. *Nat Struct Mol Biol* 13: 71-76, 2006.
4. Levitin F, Stern O, Weiss M, *et al*: The MUC1 SEA module is a self-cleaving domain. *J Biol Chem* 280: 33374-33386, 2005.
5. Kufe D, Inghirami G, Abe M, Hayes D, Justi-Wheeler H and Schlom J: Differential reactivity of a novel monoclonal antibody (DF3) with human malignant versus benign breast tumors. *Hybridoma* 3: 223-232, 1984.
6. Li Y, Liu D, Chen D, Kharbanda S and Kufe D: Human DF3/MUC1 carcinoma-associated protein functions as an oncogene. *Oncogene* 22: 6107-6110, 2003.
7. Yin L and Kufe D: Human MUC1 carcinoma antigen regulates intracellular oxidant levels and the apoptotic response to oxidative stress. *J Biol Chem* 278: 35458-35464, 2003.
8. Ren J, Agata N, Chen D, *et al*: Human MUC1 carcinoma-associated protein confers resistance to genotoxic anti-cancer agents. *Cancer Cell* 5: 163-175, 2004.
9. Yin L, Huang L and Kufe D: MUC1 oncoprotein activates the FOXO3a transcription factor in a survival response to oxidative stress. *J Biol Chem* 279: 45721-45727, 2004.
10. Wei X, Xu H and Kufe D: Human MUC1 oncoprotein regulates p53-responsive gene transcription in the genotoxic stress response. *Cancer Cell* 7: 167-178, 2005.
11. Yin L, Kharbanda S and Kufe D: Mucin 1 oncoprotein blocks hypoxia-inducible factor 1 α activation in a survival response to hypoxia. *J Biol Chem* 282: 257-266, 2007.
12. Raina D, Ahmad R, Kumar S, *et al*: MUC1 oncoprotein blocks nuclear targeting of c-Abl in the apoptotic response to DNA damage. *EMBO J* 25: 3774-3783, 2006.
13. Gendler S, Taylor-Papadimitriou J, Duhig T, Rothbard J and Burchell JA: A highly immunogenic region of a human polymorphic epithelial mucin expressed by carcinomas is made up of tandem repeats. *J Biol Chem* 263: 12820-12823, 1988.
14. Siddiqui J, Abe M, Hayes D, Shani E, Yunis E and Kufe D: Isolation and sequencing of a cDNA coding for the human DF3 breast carcinoma-associated antigen. *Proc Natl Acad Sci USA* 85: 2320-2323, 1988.
15. Merlo G, Siddiqui J, Cropp C, *et al*: DF3 tumor-associated antigen gene is located in a region on chromosome 1q frequently altered in primary human breast cancer. *Cancer Res* 49: 6966-6971, 1989.
16. Li Y, Yu W-H, Ren J, *et al*: Heregulin targets γ -catenin to the nucleolus by a mechanism dependent on the DF3/MUC1 protein. *Mol Cancer Res* 1: 765-775, 2003.
17. Li Y, Chen W, Ren J, *et al*: DF3/MUC1 signaling in multiple myeloma cells is regulated by interleukin-7. *Cancer Biol Ther* 2: 187-193, 2003.
18. Wei X, Xu H and Kufe D: MUC1 oncoprotein stabilizes and activates estrogen receptor α . *Mol Cell* 21: 295-305, 2006.
19. Ren J, Bharti A, Raina D, Chen W, Ahmad R and Kufe D: MUC1 oncoprotein is targeted to mitochondria by heregulin-induced activation of c-Src and the molecular chaperone HSP90. *Oncogene* 25: 20-31, 2006.
20. Huang L, Chen D, Liu D, Yin L, Kharbanda S and Kufe D: MUC1 oncoprotein blocks GSK3 β -mediated phosphorylation and degradation of β -catenin. *Cancer Res* 65: 10413-10422, 2005.

21. Li Y, Bharti A, Chen D, Gong J and Kufe D: Interaction of glycogen synthase kinase 3 β with the DF3/MUC1 carcinoma-associated antigen and β -catenin. *Mol Cell Biol* 18: 7216-7224, 1998.
22. Li Y, Kuwahara H, Ren J, Wen G and Kufe D: The c-Src tyrosine kinase regulates signaling of the human DF3/MUC1 carcinoma-associated antigen with GSK3 β and β -catenin. *J Biol Chem* 276: 6061-6064, 2001.
23. Ren J, Li Y and Kufe D: Protein kinase C δ regulates function of the DF3/MUC1 carcinoma antigen in β -catenin signaling. *J Biol Chem* 277: 17616-17622, 2002.
24. Dekker J, Rossen JW, Buller HA and Einerhand AW: The MUC family: an obituary. *Trends Biochem Sci* 27: 126-131, 2002.
25. Duraisamy S, Ramasamy S, Kharbanda S and Kufe D: Distinct evolution of the human carcinoma-associated transmembrane mucins, MUC1, MUC4 AND MUC16. *Gene* 373: 28-34, 2006.
26. Benson DA, Boguski MS, Lipman DJ, *et al*: GenBank. *Nucleic Acids Res* 27: 12-17, 1999.
27. Schultz J, Milpetz F, Bork P and Ponting CP: SMART, a simple modular architecture research tool: identification of signaling domains. *Proc Natl Acad Sci USA* 95: 5857-5864, 1998.
28. Altschul SF, Gish W, Miller W, Myers EW and Lipman DJ: Basic local alignment search tool. *J Mol Biol* 215: 403-410, 1990.
29. Thompson JD, Higgins DG and Gibson TJ: CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22: 4673-4680, 1994.
30. Felsenstein J: PHYLIP-Phylogeny Inference Package (Version 3.2). *Cladistics* 5: 164-166, 1989.
31. Page R: TreeView: an application to display phylogenetic trees on personal computers. *Comput Appl Biosci* 12: 357-358, 1996.
32. Bork P and Patthy L: The SEA module: a new extracellular domain associated with O-glycosylation. *Protein Sci* 4: 1421-1425, 1995.
33. Parry S, Silverman HS, McDermott K, Willis A, Hollingsworth MA and Harris A: Identification of MUC1 proteolytic cleavage sites *in vivo*. *Biochem Biophys Res Commun* 283: 715-720, 2001.
34. Chai NN, Zhou H, Hernandez J, Najmabadi H, Bhasin S and Yen PH: Structure and organization of the RBMY genes on the human Y chromosome: transposition and amplification of an ancestral autosomal hnRNPG gene. *Genomics* 49: 283-289, 1998.
35. Wang PW, Eisenbart JD, Cordes SP, Barsh GS, Stoffel M and Le Beau MM: Human KRML (MAFB): cDNA cloning, genomic structure, and evaluation as a candidate tumor suppressor gene in myeloid leukemias. *Genomics* 59: 275-281, 1999.
36. Nagase T, Nakayama M, Nakajima D, Kikuno R and Ohara O: Prediction of the coding sequences of unidentified human genes. XX. The complete sequences of 100 new cDNA clones from brain which code for large proteins *in vitro*. *DNA Res* 8: 85-95, 2001.
37. Donoviel DB, Freed DD, Vogel H, *et al*: Proteinuria and perinatal lethality in mice lacking NEPH1, a novel protein with homology to NEPHRIN. *Mol Cell Biol* 21: 4829-4836, 2001.
38. Law SF, Estojak J, Wang B, Mysliwiec T, Kruh G and Golemis EA: Human enhancer of filamentation 1, a novel p130cas-like docking protein, associates with focal adhesion kinase and induces pseudohyphal growth in *Saccharomyces cerevisiae*. *Mol Cell Biol* 16: 3327-3337, 1996.
39. Sun XJ, Wang LM, Zhang Y, *et al*: Role of IRS-2 in insulin and cytokine signalling. *Nature* 377: 173-177, 1995.
40. Desseyn JL, Guyonnet-Duperat V, Porchet N, Aubert JP and Laine A: Human mucin gene MUC5B, the 10.7-kb large central exon encodes various alternate subdomains resulting in a super-repeat. Structural evidence for a 11p15.5 gene family. *J Biol Chem* 272: 3168-3178, 1997.
41. Pigny P, Guyonnet-Duperat V, Hill AS, *et al*: Human mucin genes assigned to 11p15.5: identification and organization of a cluster of genes. *Genomics* 38: 340-352, 1996.
42. Desseyn JL, Buisine MP, Porchet N, Aubert JP, Degand P and Laine A: Evolutionary history of the 11p15 human mucin gene family. *J Mol Evol* 46: 102-106, 1998.
43. Desseyn JL, Aubert JP, Porchet N and Laine A: Evolution of the large secreted gel-forming mucins. *Mol Biol Evol* 17: 1175-1184, 2000.
44. Huang L, Ren J, Chen D, Li Y, Kharbanda S and Kufe D: MUC1 cytoplasmic domain coactivates Wnt target gene transcription and confers transformation. *Cancer Biol Ther* 2: 702-706, 2003.
45. Perrais M, Pigny P, Buisine MP, Porchet N, Aubert JP and van Seuning-Lempire I: Aberrant expression of human mucin gene MUC5B in gastric carcinoma and cancer cells. Identification and regulation of a distal promoter. *J Biol Chem* 276: 15386-15396, 2001.
46. Sonora C, Mazal D, Berois N, *et al*: Immunohistochemical analysis of MUC5B apomucin expression in breast cancer and non-malignant breast tissues. *J Histochem Cytochem* 54: 289-299, 2006.
47. Yamamoto M, Bharti A, Li Y and Kufe D: Interaction of the DF3/MUC1 breast carcinoma-associated antigen and β -catenin in cell adhesion. *J Biol Chem* 272: 12492-12494, 1997.
48. Li Q, Ren J and Kufe D: Interaction of human MUC1 and β -catenin is regulated by Lck and ZAP-70 in activated Jurkat T cells. *Biochem Biophys Res Commun* 315: 471-476, 2004.
49. Lang T, Hansson GC and Samuelsson T: An inventory of mucin genes in the chicken genome shows that the mucin domain of Muc13 is encoded by multiple exons and that ovomucin is part of a locus of related gel-forming mucins. *BMC Genomics* 7: 197, 2006.