MSX1 mutations and associated disease phenotypes: genotype-phenotype relations

Jia Liang[1], Johannes Von den Hoff[2], Joanna Lange[3,6], Yijin Ren[4], Zhuan Bian*[,1] and Carine EL Carels*[,2,5]

The Msx1 transcription factor is involved in multiple epithelial-mesenchymal interactions during vertebrate embryogenesis. It has pleiotropic effects in several tissues. In humans, MSX1 variants have been related to tooth agenesis, orofacial clefting, and nail dysplasia. We correlate all MSX1 disease causing variants to phenotypic features to shed light on this hitherto unclear association. MSX1 truncations cause more severe phenotypes than in-frame variants. Mutations in the homeodomain always cause tooth agenesis with or without other phenotypes while mutations outside the homeodomain are mostly associated with non-syndromic orofacial clefts. Downstream effects can be further explored by the edgetic perturbation model. This information provides new insights for genetic diagnosis and for further functional analysis of MSX1 variants.
European Journal of Human Genetics (2016) 24, 1663–1670; doi:10.1038/ejhg.2016.78; published online 6 July 2016

INTRODUCTION

Vertebrate craniofacial structures develop through multiple interactions between epithelium and the underlying mesenchyme, and a series of transcription factors are involved in these processes.[1] Among these transcription factors are members of the Msx homeobox gene family that are related to the Drosophila msh (muscle segment homeobox) genes.[2] In humans, the MSX gene family consists of two members - MSX1 and MSX2 - that are expressed in partly overlapping patterns in the embryonic craniofacial region.[3] As a transcriptional repressor, Msx1 is crucial in palatogenesis and odontogenesis.[4] In addition, it is also involved in limb formation, development of the nervous system and tumor growth inhibition.[5,6] In humans, MSX1 variants show pleiotropic phenotypes with variable association with non-syndromic cleft lip with or without cleft palate (CL/P,OMIM #608874, Orofacial Cleft 5, OFC5), non-syndromic tooth agenesis (OMIM, #106600 - Tooth Agenesis, Selective, 1; STHAG1), Witkop syndrome (OMIM, #189500) and Wolf-Hirschhorn syndrome (WHS, OMIM, #194190).[7–10] Mice with a homozygous deletion of Msx1 exhibit a complete cleft palate and failure of tooth development.[11]

Affected individuals from the same pedigree carrying the same MSX1 variant also show variable phenotype severity. Hypotheses for this variability include effects of epigenetic and environmental factors, stochastic effects, as well as the effects of modifier genes.[12] The study of genotype-phenotype relations for MSX1 variants therefore seems complex. Previous studies have mainly focused on tooth agenesis, associating the MSX1 variants to the type and number of missing teeth, and it is generally concluded that the most frequently missing teeth in case of MSX1 variants are the second premolars.[13–15] In a previous study we identified a novel MSX1 mutation causing tooth agenesis with cleft lip, further confirming that different MSX1 mutations may cause different phenotypes.[16]

This review reveals a strong correlation between observed phenotypes and the location in the MSX1 protein structure of the disease causing mutations. Mutations in the structural part, which disturb DNA binding of the homeodomain (HD), preferentially cause tooth agenesis with or without other phenotypes, while variants in the natively unfolded N-terminal part of the protein generally cause ns OFC. Interestingly we found that variants associated with ns OFC are all in-frame missense mutations, while syndrome-associated variants all are truncating mutations which do affect the HD. Truncating MSX1 mutations cause more severe tooth agenesis than in-frame MSX1 mutations. Using the known part of the 3D molecular MSX1 structure for locating the different variants in this review, the indel mutations are predicted to have the most deleterious effect on the DNA binding function, and hence cause the most severe tooth agenesis phenotypes.

DISEASE PHENOTYPES CAUSED BY MSX1 MUTATIONS

MSX1 and non-syndromic orofacial clefting

Non-syndromic orofacial clefts (ns OFC) are common birth defects in humans and are generally classified as cleft lip with or without cleft palate or cleft palate only. The etiology of ns OFC is complex involving genetic as well as environmental factors.[17] Msx1 first emerged as a candidate gene for clefts based on a transgenic knock-out (KO) mouse experiment.[11] Thereafter case-control and nuclear family-based studies showed that MSX1 also plays a role in human clefting and more than 10 MSX1 variants have already been related to CL/P[7,18–21] (Table 1; Figures 1a and b). It has been predicted that MSX1 variants account for the etiology of 2% of all OFC cases.[7,21] Although Msx1 was shown to maintain the growth of the primary palate during mammalian palatogenesis through expression regulation of growth factors such as Bmp4,[22] the exact role of Msx1 in these and other regulatory actions is

[1]The State Key Laboratory Breeding Base of Basic Science of Stomatology (Hubei- MOST) and Key Laboratory of Oral Biomedicine Ministry of Education, School and Hospital of Stomatology, Wuhan University, Wuhan, P.R. China; [2]Department of Orthodontics and Craniofacial Biology, Radboud University Medical Center, Nijmegen, The Netherlands; [3]Centre for Molecular and Biomolecular Informatics, Radboud University Medical Center, Nijmegen, The Netherlands; [4]Department of Orthodontics, University of Groningen and University Medical Center of Groningen, Groningen, The Netherlands; [5]Department of Oral Health Sciences, KU Leuven, Leuven, Belgium
*Correspondence: Professor Dr Z Bian, School and Hospital of Stomatology Wuhan University, 237 Luoyu Road, Wuhan 430079, P.R. China. Tel: +86 27 87646337; Fax: +86 27 87873260; E-mail: bianzhuan@whu.edu.cn
or Professor Dr CEL Carels, Department of Orthodontics and Craniofacial Biology Philips van Leydenlaan, 25 6525 Ex Nijmegen, The Netherlands. Tel: +31 24 361 4005; Fax: +31 24 354 0631; E-mail: Carine.Carels@radboudumc.nl
[6]Present address: Bio-Product BV, Nijmegen, The Netherlands.

Table 1 *MSX1* mutations with non-syndromic orofacial clefts (nsOFC)

Type of mutation	HGVS[a] (cDNA)	HGVS[a] (protein)	Homeodomain affected	References
Missense	c.65G>A	p.(G22D)	no	18
Missense	c.127A>C	p.(M43L)	no	19
Missense	c.251A>T	p.(E84V)	no	7
Missense	c.290G>A	p.(G97D)	no	7
Missense	c.311G>A	p.(G104E)	no	20
Missense	c.359T>G	p.(V120G)	no	7
Missense	c.365G>A	p.(G122E)	no	7
Missense	c.458C>A	p.(P153Q)	no	20
Missense	c.471G>T	p.(R157S)	no	7
Missense	c.817G>T	p.(G273C)	no	21
Missense	c.850C>T	p.(P284S)	no	21

[a]In this review, all mutations are annotated using HGVS (Human Genome Variation Society) system.

still largely unknown. By investigating information assembled by KMAD (Knowledge-based Multiple sequence Alignment for intrinsically Disordered proteins),[23] we found the highly conserved amino acid sequence motif Lig_EH1_1 (the Engrailed Homology-1 motif) of which the Eukaryotic Linear Motif (ELM) database[24] reports that it is part of a composite, highly conserved motif (Figure 2) in eukaryotes where it recruits other proteins that subsequently cause down regulation. This motif is observed in the Internally Disordered Part (IDP) recently also called the Natively Unfolded Part (NUP) of the MSX1 protein that lies N-terminal to the homeodomain (see Figure 2). This EH1 motif is known to mediate physical interaction with transcriptional co-repressors of the Groucho/TLE proteins family. The HE1 motif recruits and binds these co-repressors taking part in transcriptional down regulating protein complexes.[25]

MSX1 and non-syndromic tooth agenesis

Non-syndromic tooth agenesis (ns TA) is another common developmental anomaly that can be caused by *MSX1* variants.[14] To date, nearly 20 *MSX1* mutations have been related to ns TA[8,13,26–37] (Table 2). Functional analyses of the mutant proteins suggest that haploinsufficiency of MSX1 underlies this phenotype.[16,38] Mutant proteins were predicted to either exhibit reduced stability or to have reduced DNA binding capacity, or to have reduced capability to interact with their cognate protein factors. As a result, the function of MSX1 as a transcriptional repressor can be greatly impaired.[27,39,40]

This could also be the result of epigenetic silencing of *MSX1* by DNA methylation giving rise to either or combinations of its phenotypes or increase risk for tumoral growth

Syndromes caused by *MSX1* mutations

Mutations in or including the *MSX1* gene can also cause syndromic forms of tooth agenesis, including Wolf-Hirschhorn syndrome, Witkop syndrome, and tooth agenesis combined with orofacial clefting (Table 3). The Wolf-Hirschhorn syndrome is caused by a deletion of the WHS locus - including *MSX1* - on chromosome 4p.[10] The phenotype includes mental and growth retardation, characteristic craniofacial features, seizures, and tooth agenesis. Furthermore, a nonsense mutation in *MSX1* accounts for the genetic etiology of Witkop syndrome characterized by tooth agenesis and nail dysplasia.[9] Up to now, two reports propose *MSX1* as a candidate gene for tooth agenesis combined with orofacial clefting.[16,41]

FROM GENOTYPE TO PHENOTYPE

Overview of human *MSX1* mutations

Since the first *MSX1* variant was identified in a pedigree with tooth agenesis,[8] several additional pathogenic variants have been discovered. To analyze the *MSX1* mutations, we retrieved all *MSX1* variants from HGMD (Human Gene Mutation Database) Professional up to the 30th of June, 2015. All variants with the variant class 'DM' (Disease causing Mutation) were included.[42] The only exception was a splice site variant[43] that was excluded because no detailed phenotype description was available. We obtained 31 *MSX1* variants including 18 missense mutations, 5 nonsense mutations 5 indel mutations, 1 splice variant, 1 nonstop variant and 1 entire gene deletion (Tables 1–3, Figure 1a). These variants cause either tooth agenesis with and without orofacial clefting, non-syndromic orofacial clefting, Wolf-Hirschhorn syndrome or Witkop syndrome (Tables 1–3, Figure 1b).

### In-frame mutations and truncating mutations

First the *MSX1* variants were separated into two subsets: in-frame mutations comprising only missense mutations, and truncating mutations comprising all nonsense mutations, out-of-frame insertions or deletions, mutations causing defective splicing, nonstop mutations, and deletions of the entire gene. We found that *MSX1* in-frame mutations are more frequent than truncating mutations (18 vs13) (Figure 1a) which is consistent with many other Mendelian genes.[44]

We analysed the phenotypes in terms of the domains where the variants are located and identified a clear segregation. The average number of missing teeth associated with in-frame mutations is lower than with truncating mutations leading to more severe phenotypes (Figure 3). This might be explained by haploinsufficiency, which is thought to play a role in tooth agenesis.[38] Msx1 mutations mostly function in a dose-sensitive manner but because truncating mutations often lead to complete loss of function or to abnormal mRNA and/or protein expression they result in more disabled Msx1 proteins than in case of in-frame mutations.[16]

Secondly, in order to analyse the relation between the TA phenotype severity to the mutations in as much detail as possible, we refined the phenotypes by introducing 6 subcategories with increasing (average) number of TA (Table 4). This way we could cover the spectrum of dental phenotype severity from agenesis of 1 - 4 teeth towards agenesis of 17-20 teeth in the permanent dentition (except for the wisdom teeth). Moreover we added the changes in individual amino-acid characteristics (charge, polarity, hydrophilia/hydrophobia and volume) which could contribute or explain the observed phenotype. When the 8 mildest phenotypes (i.e. subcategories 1-4, 5-8 and 9-12) were compared with the 8 most severe phenotypes (subcategories 13-16, 17-20 and 21-24), 3 differences were found (Table 4). First difference concerned the number of hits in the HD: in the group of mildest phenotype 5 of the 8 mutations affected the HD, while in the severest phenotype group 7 of the 8 mutations affected the HD. The second difference relates to the number of frameshift mutations: in the first group only 3 of 8, while in the second group 5 of 8 are frameshift mutations. The third difference is the presence/absence of nucleotide duplications: interestingly, no variants with nucleotide duplications are associated with the 8 mildest phenotypes while they do occur in 4 of the 8 variants with the severest phenotypes. Therefore we can confirm that the TA phenotype severity is not only related to the domain location of the mutation (HD or not) but also to the type of mutation, with frameshift mutations causing more often more severe TA. However, within the frameshifts, mutations due to duplications of nucleotides or dinucleotides only cause severe ns TA (Figure 3 and
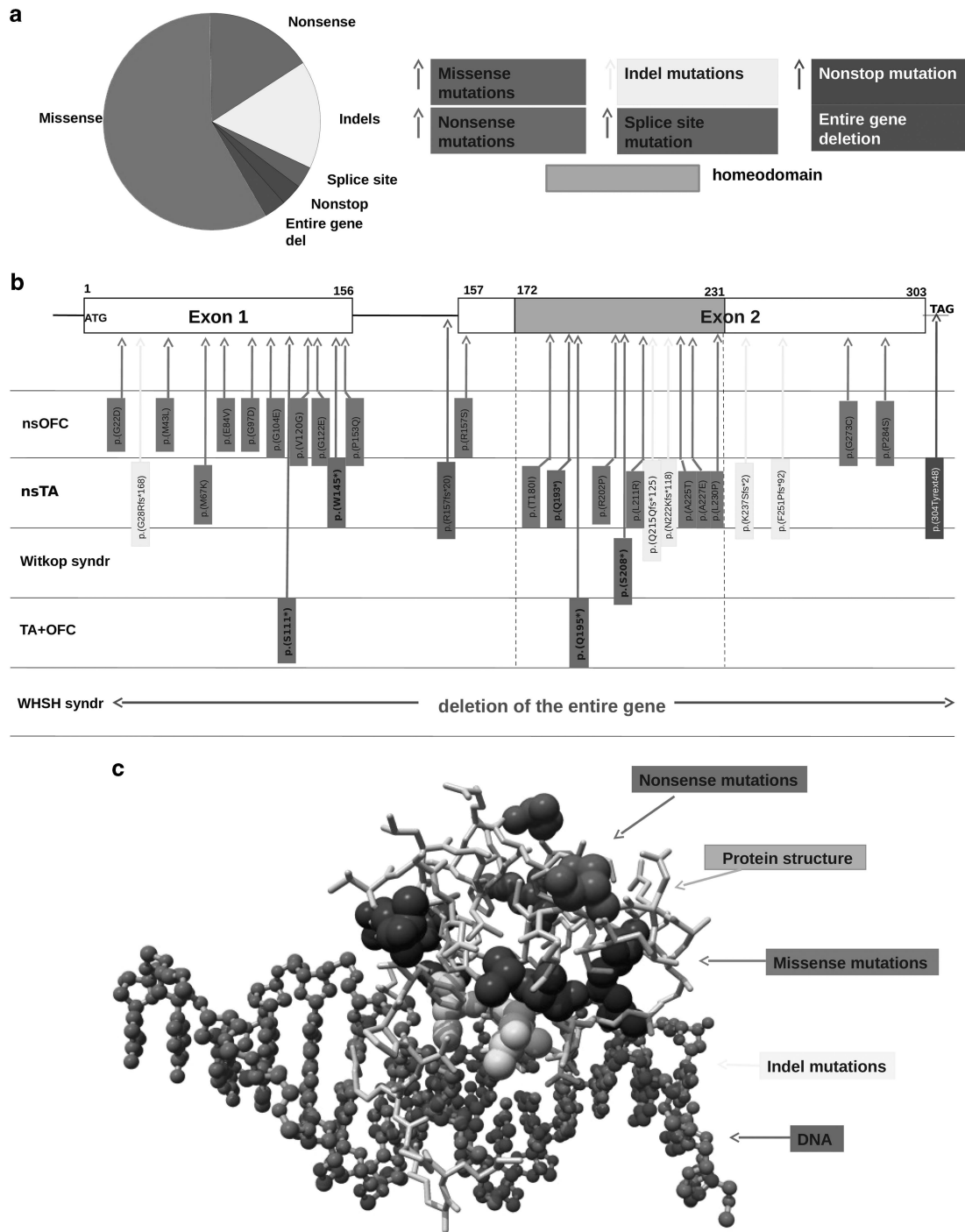
Figure 1 Distribution of all *MSX1* mutations included in this review. Reference sequence NC_000004.12, chrom 4, GRCh38.p2 (GCF_000001405.28); NP 002439.2; NM 002448,3 was used on which the variant description of nucleotides and amino acids are based. (**a**) Pie diagram with the 'in-frame' (blue; *n* = 18) versus 13 'truncating' mutations of which 5 nonsense mutations (*red*), 5 indel mutations (*yellow*), 1 splice site mutation (*green*), 1 nonstop mutation (*purple*), as well as 1 whole gene deletion (*brown*). The in-frame mutations were only missense mutations. (**b**) Mapping of the *MSX1* mutations with associated disease phenotypes to coding and non-coding structures of MSX1. The *horizontal boxes on top* represent exon 1 and exon 2; the *line* between the two exons represents the intron; the *grey box* is the homeodomain (HD) coding area; *ATG*, start codon; *TAG*, stop codon. The mutant *MSX1* variants in the vertical *blue boxes* represent the in-frame missense variants; those in the *red boxes* the nonsense mutations, in the *yellow boxes* the indel mutations, in the *green box* the splice site mutation and in the *purple box* the nonstop mutation. (**c**) In the yet experimentally determined structural part of the MSX1 protein (PSI; Protein Model Portal; 1ig7; Msx-1 Homeodomain/DNA Complex Structure; residues 173-230; http://www.proteinmodelportal.org/query/uniprot/P28360) the mutations are indicated in the same colors as in parts (**a**) and (**b**) of Figure 1: *blue* for the missense mutations; *red* for the nonsense mutations and *yellow* for the indel mutations; DNA is depicted in *green*; and protein structure in *grey*. From the location in the MSX1 structure, the indel mutations can be predicted to severely disturb the DNA binding. Although the same holds true for the missense mutations, they will disturb the DNA binding with variable severity. The red residues become stop codons and thus represent the first absent residues in these mutations; their effect cannot be predicted. A full color version of this figure is available at the *European Journal of Human Genetics* journal online.
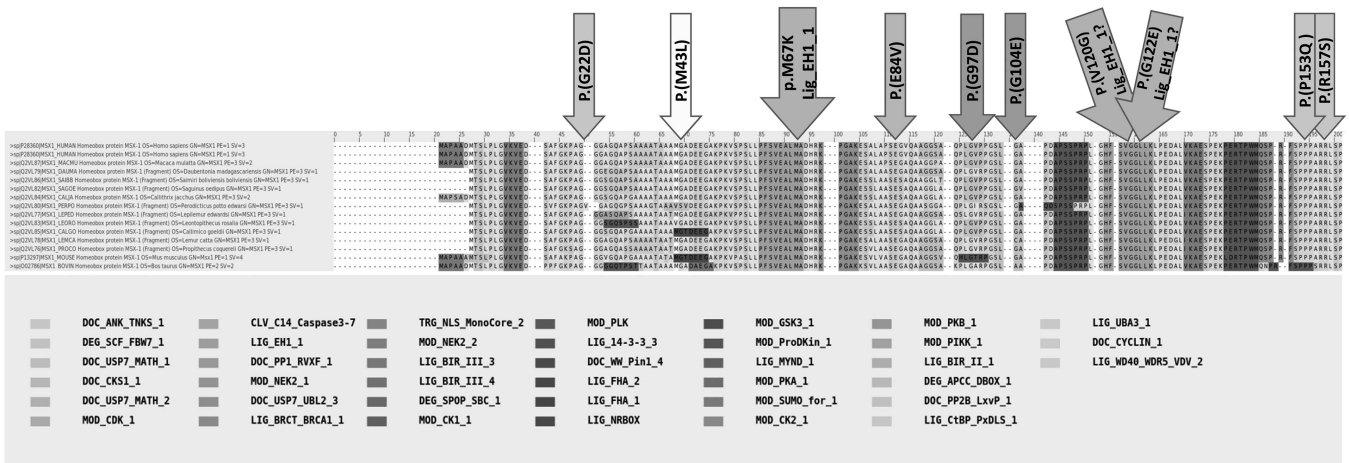
Figure 2 Panel with the amino acid sequences and motifs in the natively unfolded part of the MSX1 protein According to the Knowledge based Multiple sequence Alignment for intrinsically Disordered proteins (KMAD; DOI:10.1093/bioinformatics/btv663; Lange et al, 2016) most of the prediction programs attribute the term 'natively unfolded') to this part of the MSX1 protein lying N-terminally to the homeodomain. Although many of the sequence motifs of this natively unfolded part of the MSX1 protein have not yet been experimentally associated to a function (KMAD-align; DOI:10.1093/bioinformatics/btv663), an important role is obvious when the amino-acid (AA) sequence is highly conserved. This is the case for the AA sequence in the large orange block under the *orange* arrow which refers to the LIG_EH1_1 motif where the M67K mutation causing ns TA is residing. This mutation will dramatically disrupt the function of this LIG_EH1_1 motif conserved sequence. All other arrows point to the MSX1 variants which are all associated with ns OFC. It can be noticed that the motifs/sequences containing these variants show a lesser degree of conservation. *MOD:* stands for variants found in mouse data. Both MOD_SUMO and MOD_CDK1 (see color legend), have previously been mentioned in relation to cleft lip and palate development in mice. A full color version of this figure is available at the *European Journal of Human Genetics* journal online.

Table 2 *MSX1* mutations in individuals with non-syndromic tooth agenesis

Type of mutation	HGVS (cDNA)	HGVS (protein)	Homeodomain affected	Number of missing teeth[a]	Number of patients[b]	References
Missense	c.200T>A	p.(M67K)	No	6-13	4	[26]
Missense	c.539C>T	p.(T180I)	Yes	5	1	[27]
Missense	c.605G>C	p.(R202P)	Yes	4-11	9	[8]
Missense	c.632T>G	p.(L211R)	Yes	8-18	3	[27]
Missense	c.673G>A[c]	p.(A225T)	Yes	11-19	2	[28]
Missense	c.680C>A	p.(A227E)	Yes	5-13	4	[29]
Missense	c.689T>C	p.(L230P)	Yes	2-4	5	[30]
Non-stop	c.910_911dup TA	p.(304Tyrext48)	No	15	1	[31]
Nonsense	c.434G>A	p.(W145*)	Yes	5-14	3	[32]
Nonsense	c.577C>T	p.(Q193*)	Yes	12-15	2	[33]
Deletion	c.708delG	p.(K237Sfs*2)	No	8	1	[34]
Insertion	c.80dupG	p.(G28Rfs*168)	Yes	6-12	2	[13]
Insertion	c.644dupA	p.(Q215Qfs*125)	Yes	18	1	[35]
Insertion	c.665dupA	p.(N222Kfs*118)	Yes	11	1	[34]
Insertion	c.750_751insAGGGGGTGGG	p. (F251Pfs*92)	No	4-6	7	[36]
Splice	c.470-9G>A	p.(R157fs*20)	Yes	4-14	4	[37]

[a]Numbers represent the number or range of missing teeth, third molars excluded.
[b]Number of patients indicates the total number of patients with confirmed phenotype.
[c]Autosomal recessive mutation.

Table 4). When we additionally analyzed the changes in amino acid parameters in the 2 phenotype severity subgroups (Table 4), not much could be concluded extra, as only 2 missense variants were present in the severest phenotype group versus 5 in the mildest phenotype group. Nevertheless, the volume of the mutant amino acid seems important. The volume of the AA was the only parameter that was oppositely different from the wild type (WT) amino acid. In the mildest phenotype with an average of 3,6 tooth ageneses, the volume of the mutant amino acid was smaller than the WT (L>P) (Table 4) while the charge, polarity and hydrophilia did not change. In the severest phenotype with 15 tooth ageneses, all other parameters being equal, the mutant amino acid was bigger than the WT (A<T) (Table 4).

Thirdly, the mutations in the HD were mapped on the 3D structure of this domain and visualized with YSARA[45] (Figure 1c). From the location in the MSX1 structure, the indel mutations can be predicted to severely disturb the DNA binding. Although the same holds true for the missense mutations, they will disturb the DNA binding with variable severity (Table 4; Figure 1c). The red residues become stop codons and thus represent the first absent residues in these variants; therefore their effect cannot be predicted.

Table 3 *MSX1* mutations with combined phenotypes or syndrome

Type of mutation	HGVS (cDNA)	HGVS (protein)	Homeodomain affected	Phenotype[a]	Number of patients	References
Nonsense	c.332C>A	p.(S111*)	Yes	4-12;CL/P	8	43
Nonsense	c.583C>T	p.(Q195*)	Yes	14;CL/P	2	16
Nonsense	c.623C>A	p.(S208*)	Yes	7-24;Witkop	8	9
Entire gene deletion	/	/	Yes	11-21;Wolf-Hirschhorn	5	10

[a]Numbers represent the range of missing teeth, third molars excluded; CL/P, cleft lip with or without palate.

Mutations (not) affecting the homeodomain

Mutations located in different protein domains tend to cause different disease phenotypes,[46] as a result of different protein functions being affected. For example in case of *TP63*, mutations in different domains cause either ectrodactyly ectodermal cleft lip/palate (EEC) syndrome or ankyloblepharon-ectodermal dysplasia-cleft lip/palate (AEC).[47]

MSX1 consists of two exons, the second of which includes the highly conserved homeodomain (HD). The homeodomain is essential for protein stability, DNA binding, transcriptional repression and interactions with other odontogenic molecules like PAX9, TATA binding protein (TBP) and DLX family members.[38] When mapping the 31 identified *MSX1* variants onto the gene and protein, some appear to affect the HD, including all mutations in the HD itself and truncating mutations upstream of the HD, and others did not (Figure 1b). The latter included all in-frame mutations outside the HD and truncating mutations downstream of the HD. This yielded 16 mutations affecting vs 15 mutations not affecting the HD (Figure 3). Mutations affecting the HD preferentially cause tooth agenesis with or without other phenotypes, while mutations not affecting the HD preferentially cause orofacial clefting (Figure 3). Only 4 out of 15 mutations not affecting the HD cause ns TA including three truncating mutations which affect the MSX1 protein structure[31] and one missense mutation – the p.(M67K) - which is predicted to reside in the highly conserved EH1 sequence motif in the natively unfolded part of the MSX1 protein N-terminal of the HD.

All *MSX1* missense variants associated with ns OFC also map to the natively unfolded part (NUP) of the MSX1 protein and when locating them onto these sequences and motifs they all map to mildly less conserved sequences and motifs, than the LIG_EH_1 motif (Figures 1b and 2). Although the function of these motifs all are not yet known or experimentally validated, we can conclude that contrary to ns TA, ns OFC is mainly associated with MSX1 variants involved in regulatory cell processes, which are located in a bit less conserved area's than those causing ns TA. Moreover all the mutations associated with nsOFC are mapping outside the HD, which suggests that the HD is less important in orofacial clefting. For the syndromic phenotypes, all mutations are truncating mutations affecting the HD and also the sequence on the C-terminal side of the HD (Figures 1b and c).

EDGETIC PERTURBATION MODEL

Several difficulties hinder the prediction of an exact phenotype from the type and location of the *MSX1* mutation. These include gene pleiotropy, incomplete penetrance, and also variable expressivity. Only few examples exist in which the corresponding phenotype could reliably be predicted from the mutations.[48]

Recently, network modeling has been introduced in order to explain how specific mutations may lead to distinct phenotypes.[44,46,49–51] In a specific network model called the edgetic perturbation model, a mutation is considered to alter molecular interactions either due to
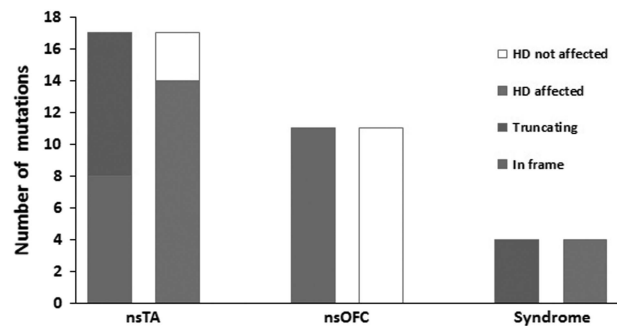
Figure 3 Comparison of different types of *MSX1* variants. In-frame versus truncating mutations and their accordingly phenotypes: non-syndromic tooth agenesis (ns TA) and non-syndromic orofacial clefts (ns OFC). In-frame (*blue*) versus truncating (*red*) variants and their accordingly number of missing teeth. All mutations in this histogram cause tooth agenesis with or without other phenotypes. The bars indicate the average number of missing teeth with error bars representing the maximum and minimum number. The arrows above the bars indicate the presence of other phenotypes. MSX1 variants affecting homeodomain (HD) versus not affecting HD with their accordingly phenotypes. Non-syndromic tooth agenesis (ns TA) and non-syndromic orofacial clefts (ns OFC). A full color version of this figure is available at the *European Journal of Human Genetics* journal online.

edgetic perturbations or due to node removal (Figure 4). Edgetic perturbation leads to the removal or addition of specific interactions while all other interactions (or edges) remain equal. In case of node removal all the interactions with other molecules are deleted. The perturbation of specific interactions arising from individual genetic variants can give rise to distinct phenotypes (Figure 4).

In line with this edgetic perturbation model, we hypothesize that in-frame *MSX1* mutations are likely to cause edgetic perturbations involving only one edge leading to ns TA (Figure 4c). Mutations not affecting the HD would perturb a different edge, leading to ns OFC (Figure 4b). Finally truncating mutations might cause node removal perturbing all edges leading to a combination of several phenotypes or a syndrome (Figure 4d).

CONCLUSIONS

MSX1 mutations cause different phenotypes depending on their location in the gene. Variants affecting the HD mainly cause tooth agenesis with or without other phenotypes, while mutations not affecting the HD preferentially cause nsOFC. Mutations causing ns OFC are all in-frame mutations not affecting the HD while syndrome-associated mutations are all truncating mutations which do affect the HD.

Truncating MSX1 mutations cause more severe tooth agenesis phenotypes than in-frame MSX1 mutations. Our findings and our hypotheses based on the edgetic perturbation model can help explain

Table 4 _MSX1_ variants in subphenotypes of increasing severity of non-syndromic tooth agenesis (nsTA) and changes in charge, polarity and hydrophilia/-phobia

Type of mutation > No. of tooth ageneses V	_MSX1_ variants (HGVS;cDNA)	MSX1 variants (HGVS; predicted protein)	No. of patients/average no. of missing teeth/(ref)	Variants in or affecting DNA-binding domain ?	Charge, polarity and hydrophilic/ phobic changes due to mutations in amino acid	Predicted volumetric effect on protein?
1 – 4 TA	c.689T>C	p.(L230P)	5 / 3.6 / (30)	Yes	L: neutral, apolar, hydrophobic> P: neutral, apolar, hydrophobic	L > P
	c.750_751insAGGGGGTGGG	p. (F251Pfs*92)	7 / 4.4 / (36)	No	[c]	[c]
5 – 8 TA	c.539C>T	p.(T180I)	1 / 5 / (27)	Yes	T: neutral, polar, hydrophilic > I: neutral, apolar, hydrophobic	T < I
	c.605G>C	p.(R202P)	9 / 6.8 / (8)	Yes	R: positive, polar, hydrophilic > P: neutral, apolar, hydrophobic	R >> P
	c.680C>A	p.(A227E)	4 / 8 / (29)	Yes	A: neutral, apolar, hydrophilic > E: neutral, polar, hydrophilic	A << E
	c.708delG	p.(K237Sfs*2)	1 / 8 / (34)	No	[c]	[c]
	c.200T>A	p.(M67K)	4 / 8.3 / (26)	No	M: neutral, apolar, hydrophobic > K: positive, polar, hydrophilic	M < K
	c.470-9G>A	p.(R157fs*20)	4 / 8.3 / (37)	Yes, HD truncated	[c]	[c]
9 – 12 TA	c.80dupG	p.(G28Rfs*168)	2 / 9 / (13)	Yes, HD truncated	[c]	[c]
	c.434G>A	p.(W145*)	3 / 9 / (32)	Yes, HD truncated	[c]	[c]
	c.665dupA	p.(N222Kfs*118)	1 / 11 / (34)	Yes	[c]	[c]
	c.632T>G	p.(L211R)	3 / 11.7 / (27)	Yes	L: neutral, apolar, hydrophobic> R: positive, polar, hydrophilic	L << R
13 – 16 TA	c.577C>T	p.(Q193*)	2 / 13.5 / (33)	Yes	[c]	[c]
	c.673G>A[a]	p.(A225T)	2 / 15 / (28)	Yes	A: neutral, polar, hydrophobic >> T: neutral, polar, hydrophilic >	A <T
	c.910_911dup TA	p.(304Tyrext48)	1 / 15 / (31)	No	[c]	[c]
17 – 20 TA	c.644dupA	p.(Q215Qfs*125)	1 / 18 / (35)	Yes	[c]	[c]
21 – 24 TA	NA[b]	[c]	[d]	[d]	[d]	[d]
25 – 28 TA	NA[b]	[c]	[d]	[d]	[d]	[d]

Legend:
Missense mutation
Indel mutation
Splice site mutation
Nonsense mutation
Nonstop mutation

< smaller than
<< much smaller than
> bigger than
>> much bigger than
WT = wild type
M= mutant

[a] Autosomal recessive mutation.
[b] NA: no MSX1 mutation found for this phenotype
[c] No protein prediction possible
[d] No data available
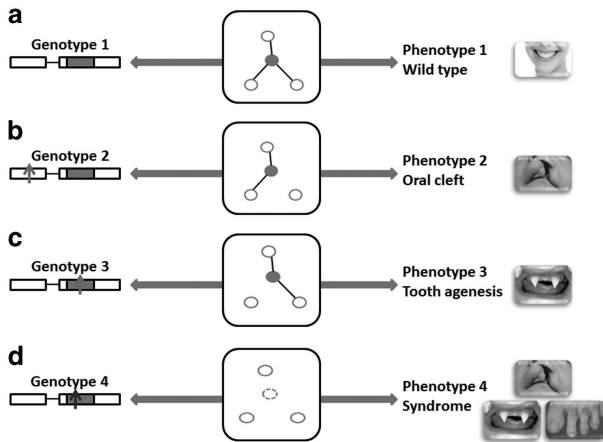
Figure 4 The edgetic perturbation model. Different *MSX1* variants might have different effects on the MSX1 network. Compared to the wild type genotype with full interaction (**a**), loss of some interaction (**b** and **c**), or loss of all interaction (node removal, **d**), leads to different phenotypes. Blue arrows represent the in frame mutations; red arrows are the truncating mutations. Nodes represent molecules, and edges represent interactions between them. A full color version of this figure is available at the *European Journal of Human Genetics* journal online.

how *MSX1* mutations alter molecular interactions and cause specific phenotypes. This not only increases our mechanistic insight and understanding of the pathogenicity of *MSX1* variants in craniofacial disorders, but could also expand the options to be considered for their precision treatment in the future.

CONFLICT OF INTEREST
The authors declare no conflict of interest.

1 Francis-West PH, Robson L, Evans DJ: Craniofacial development: the tissue and molecular interactions that control development of the head. *Adv Anat Embryol Cell Biol* 2003; **169**: III-VI 1–138.
2 Hill RE, Jones PF, Rees AR *et al*: A new family of mouse homeo box-containing genes: molecular structure, chromosomal location, and developmental expression of Hox-7.1. *Genes Dev* 1989; **3**: 26–37.
3 Jowett AK, Vainio S, Ferguson MW, Sharpe PT, Thesleff I: Epithelial-mesenchymal interactions are required for msx 1 and msx 2 gene expression in the developing murine molar tooth. *Development* 1993; **117**: 461–470.
4 Alappat S, Zhang ZY, Chen YP: Msx homeobox gene family and craniofacial development. *Cell Res* 2003; **13**: 429–442.
5 Catron KM, Wang H, Hu G, Shen MM, Abate-Shen C: Comparison of MSX-1 and MSX-2 suggests a molecular basis for functional redundancy. *Mech Dev* 1996; **55**: 185–199.
6 Bendall AJ, Abate-Shen C: Roles for Msx and Dlx homeoproteins in vertebrate development. *Gene* 2000; **247**: 17–31.
7 Jezewski PA, Vieira AR, Nishimura C *et al*: Complete sequencing shows a role for MSX1 in non-syndromic cleft lip and palate. *J Med Genet* 2003; **40**: 399–407.
8 Vastardis H, Karimbux N, Guthua SW, Seidman JG, Seidman CE: A human MSX1 homeodomain missense mutation causes selective tooth agenesis. *Nat Genet* 1996; **13**: 417–421.
9 Jumlongras D, Bei M, Stimson JM *et al*: A nonsense mutation in MSX1 causes Witkop syndrome. *Am J Hum Genet* 2001; **69**: 67–74.
10 Nieminen P, Kotilainen J, Aalto Y, Knuutila S, Pirinen S, Thesleff I: MSX1 gene is deleted in Wolf-Hirschhorn syndrome patients with oligodontia. *J Dent Res* 2003; **82**: 1013–1017.
11 Satokata I, Maas R: Msx1 deficient mice exhibit cleft palate and abnormalities of craniofacial and tooth development. *Nat Genet* 1994; **6**: 348–356.
12 Nadeau JH: Modifier genes in mice and humans. *Nat Rev Genet* 2001; **2**: 165–174.
13 Kim JW, Simmer JP, Lin BP, Hu JC: Novel MSX1 frameshift causes autosomal-dominant oligodontia. *J Dent Res* 2006; **85**: 267–271.
14 Nieminen P: Genetic basis of tooth agenesis. *J Exp Zool B Mol Dev Evol* 2009; **312B**: 320–342.
15 Yin W, Bian Z: The Gene Network Underlying Hypodontia. *J Dent Res* 2015; **94**: 878–885.
16 Liang J, Zhu L, Meng L, Chen D, Bian Z: Novel nonsense mutation in MSX1 causes tooth agenesis with cleft lip in a Chinese family. *Eur J Oral Sci* 2012; **120**: 278–282.
17 Schutte BC, Murray JC: The many faces and factors of orofacial clefts. *Hum Mol Genet* 1999; **8**: 1853–1859.
18 Vieira AR, Castillo Taucher S, Aravena T *et al*: [Mutational analysis of the muscle segment homeobox gene 1 (MSX1) in Chilean patients with cleft lip/palate]. *Revista Medica de Chile* 2004; **132**: 816–822.
19 Salahshourifar I, Halim AS, Wan Sulaiman WA, Zilfalil BA: Contribution of MSX1 variants to the risk of non-syndromic cleft lip and palate in a Malay population. *J Hum Genet* 2011; **56**: 755–758.
20 Suzuki Y, Jezewski PA, Machida J *et al*: In a Vietnamese population, MSX1 variants contribute to cleft lip and palate. *Genet Med* 2004; **6**: 117–125.
21 Tongkobpetch S, Siriwan P, Shotelersuk V: MSX1 mutations contribute to nonsyndromic cleft lip in a Thai population. *J Hum Genet* 2006; **51**: 671–676.
22 Zhang Z, Song Y, Zhao X, Zhang X, Fermin C, Chen Y: Rescue of cleft palate in Msx1-deficient mice by transgenic Bmp4 reveals a network of BMP and Shh signaling in the regulation of mammalian palatogenesis. *Development* 2002; **129**: 4135–4146.
23 Lange J, Wyrwicz LS, Vriend G: KMAD: knowledge-based multiple sequence alignment for intrinsically disordered proteins. *Bioinformatics* 2015; **32**: 932–936.
24 Dinkel H, Michael S, Weatheritt RJ *et al*: ELM–the database of eukaryotic linear motifs. *Nucleic Acids Res* 2012; **40**: D242–D251.
25 Chen G, Courey AJ: Groucho/TLE family proteins and transcriptional repression. *Gene* 2000; **249**: 1–16.
26 Lidral AC, Reising BC: The role of MSX1 in human tooth agenesis. *J Dent Res* 2002; **81**: 274–278.
27 Yamaguchi S, Machida J, Kamamoto M *et al*: Characterization of novel MSX1 mutations identified in Japanese patients with nonsyndromic tooth agenesis. *PloS One* 2014; **9**: e102944.
28 Chishti MS, Muhammad D, Haider M, Ahmad W: A novel missense mutation in MSX1 underlies autosomal recessive oligodontia with associated dental anomalies in Pakistani families. *J Hum Genet* 2006; **51**: 872–878.
29 Xuan K, Jin F, Liu YL *et al*: Identification of a novel missense mutation of MSX1 gene in Chinese family with autosomal-dominant oligodontia. *Arch Oral Biol* 2008; **53**: 773–779.
30 Mostowska A, Biedziak B, Jagodzinski PP: Novel MSX1 mutation in a family with autosomal-dominant hypodontia of second premolars and third molars. *Arch Oral Biol* 2012; **57**: 790–795.
31 Wong SW, Liu HC, Han D *et al*: A novel non-stop mutation in MSX1 causing autosomal dominant non-syndromic oligodontia. *Mutagenesis* 2014; **29**: 319–323.
32 Kimura M, Machida J, Yamaguchi S *et al*: Novel nonsense mutation in MSX1 in familial nonsyndromic oligodontia: subcellular localization and role of homeodomain/MH4. *Eur J Oral Sci* 2014; **122**: 15–20.
33 De Muynck S, Schollen E, Matthijs G, Verdonck A, Devriendt K, Carels C: A novel MSX1 mutation in hypodontia. *Am J Med Genet A* 2004; **128A**: 401–403.
34 Arte S, Parmanen S, Pirinen S, Alaluusua S, Nieminen P: Candidate gene analysis of tooth agenesis identifies novel mutations in six genes and suggests significant role for WNT and EDA signaling and allele combinations. *PloS One* 2013; **8**: e73705.
35 Bergendal B, Klar J, Stecksen-Blicks C, Norderyd J, Dahl N: Isolated oligodontia associated with mutations in EDARADD, AXIN2, MSX1, and PAX9 genes. *Am J Med Genet A* 2011; **155A**: 1616–1622.
36 AlFawaz S, Plagnol V, Wong FS, Kelsell DP: A novel frameshift MSX1 mutation in a Saudi family with autosomal dominant premolar and third molar agenesis. *Arch Oral Biol* 2015; **60**: 982–988.
37 Tatematsu T, Kimura M, Nakashima M *et al*: An aberrant splice acceptor site due to a novel intronic nucleotide substitution in MSX1 gene is the cause of congenital tooth agenesis in a Japanese family. *PloS One* 2015; **10**: e0128227.
38 Hu G, Vastardis H, Bendall AJ *et al*: Haploinsufficiency of MSX1: a mechanism for selective tooth agenesis. *Mol Cell Biol* 1998; **18**: 6044–6051.
39 Ogawa T, Kapadia H, Feng JQ, Raghow R, Peters H, D'Souza RN: Functional consequences of interactions between Pax9 and Msx1 genes in normal and abnormal tooth development. *J Biol Chem* 2006; **281**: 18363–18369.
40 Chen Y, Bei M, Woo I, Satokata I, Maas R: Msx1 controls inductive signaling in mammalian tooth morphogenesis. *Development* 1996; **122**: 3035–3044.
41 van den Boogaard MJ, Dorland M, Beemer FA, van Amstel HK: MSX1 mutation is associated with orofacial clefting and tooth agenesis in humans. *Nat Genet* 2000; **24**: 342–343.
42 Stenson PD, Mort M, Ball EV, Shaw K, Phillips A, Cooper DN: The Human Gene Mutation Database: building a comprehensive mutation repository for clinical and molecular genetics, diagnostic testing and personalized genomic medicine. *Hum Genet* 2014; **133**: 1–9.

1670

43 Li WL, Cui JJ, Fang QY, Mei LX: [A novel mutation of MSX1 gene in a Chinese pedigree with oligodontia]. *Zhonghua Kou Qiang Yi Xue Za Zhi* 2008; **43**: 157–159.

44 Zhong Q, Simonis N, Li QR *et al*: Edgetic perturbation models of human inherited disorders. *Mol Syst Biol* 2009; **5**: 321.

45 Krieger E, Vriend G: YASARA View - molecular graphics for all devices - from smartphones to workstations. *Bioinformatics* 2014; **30**: 2981–2982.

46 Wang X, Wei X, Thijssen B, Das J, Lipkin SM, Yu H: Three-dimensional reconstruction of protein networks provides insight into human genetic disease. *Nat Biotechnol* 2012; **30**: 159–164.

47 van Bokhoven H, Brunner HG: Splitting p63. *Am J Hum Genet* 2002; **71**: 1–13.

48 Faivre L, Collod-Beroud G, Loeys BL *et al*: Effect of mutation type and location on clinical outcome in 1,013 probands with Marfan syndrome or related phenotypes and FBN1 mutations: an international study. *Am J Hum Genet* 2007; **81**: 454–466.

49 Sahni N, Yi S, Zhong Q *et al*: Edgotype: a fundamental link between genotype and phenotype. *Curr Opin Genet Dev* 2013; **23**: 649–657.

50 Sahni N, Yi S, Taipale M *et al*: Widespread macromolecular interaction perturbations in human genetic disorders. *Cell* 2015; **161**: 647–660.

51 Carter H, Hofree M, Ideker T: Genotype to phenotype via network analysis. *Curr Opin Genet Dev* 2013; **23**: 611–621.