nature genetics

Long-range enhancers regulating *Myc* expression are required for normal facial morphogenesis

Veli Vural Uslu[1], Massimo Petretich[1], Sandra Ruf[1], Katja Langenfeld[1], Nuno A Fonseca[2], John C Marioni[2] & François Spitz[1]

Cleft lip with or without cleft palate (CL/P) is one of the most common congenital malformations observed in humans, with 1 occurrence in every 500–1,000 births[1,2]. A 640-kb noncoding interval at 8q24 has been associated with increased risk of non-syndromic CL/P in humans[3–5], but the genes and pathways involved in this genetic susceptibility have remained elusive. Using a large series of rearrangements engineered over the syntenic mouse region, we show that this interval contains very remote *cis*-acting enhancers that control *Myc* expression in the developing face. Deletion of this interval leads to mild alteration of facial morphology in mice and, sporadically, to CL/P. At the molecular level, we identify misexpression of several downstream genes, highlighting combined impact on the craniofacial developmental network and the general metabolic capacity of cells contributing to the future upper lip. This dual molecular etiology may account for the prominent influence of variants in the 8q24 region on human facial dysmorphologies.

CL/P is present in a large number of malformation syndromes with mendelian inheritance, but most cases correspond to isolated non-syndromic forms[1]. The influence of environmental risk factors[6] further complicates definition of the genetic components of CL/P. Genome-wide association studies (GWAS) have identified common variants contributing to increased risk of CL/P[3–5], defining several susceptibility loci. Of these, a 640-kb locus at 8q24 showed the strongest effect size in different populations, particularly in those with European ancestry[1,3–5]. This large interval corresponds to a gene desert (**Fig. 1**). Such situations are a common outcome of GWAS and constitute a challenge in moving from statistical associations to functional mechanisms[7], particularly for phenotypes that cannot be easily reproduced in cell lines.

To obtain insights into the role of the 8q24 noncoding interval during craniofacial development and to identify the gene(s) and pathway(s) it can influence, we sought to dissect its function *in vivo*. Similarity of craniofacial development in mice and humans[8] and conserved synteny of the 8q24 region (**Supplementary Fig. 1**) suggested that the mouse constitutes a suitable model. To identify regulatory activities present in this interval, we exploited a mouse line carrying, at the distal end of the locus, an insertion of a *LacZ* reporter gene whose activity reflects endogenous surrounding regulatory activities[9]. As this regulatory sensor lay in a *Sleeping Beauty* transposon, we produced, by repeated *in vivo* transposition, 38 new insertions in the locus (**Fig. 1b** and **Supplementary Table 1**). Strikingly, insertions within or adjacent to the region orthologous to the CL/P-associated interval showed specific LacZ expression in the developing fronto- and medionasal process of embryonic day (E) 11.5 mouse embryos, when this process fuses with the maxillary process to form the future upper lip[10] (**Fig. 1** and **Supplementary Table 1**).

To more precisely localize the region(s) responsible for this regulatory activity, we used the *lox*P site present in the transposon to engineer a series of rearrangements spanning the entire interval[9] (**Fig. 1b** and **Supplementary Table 2**). We first monitored the effects of the different deletions on expression of the reconstituted *LacZ* sensor (**Fig. 1f** and **Supplementary Fig. 2**). Confirming our earlier observations[9], del(8–17) showed no LacZ expression in the developing face. All deletions encompassing the 8–14 interval also led to loss of LacZ expression in the medionasal process, whereas deletions retaining the 10–13 interval had no or minor effects on this expression domain. We also analyzed mice carrying different duplications that 'moved' genomic intervals away from the CL/P-associated interval, linking them to a different regulatory domain (**Supplementary Fig. 3**). We observed LacZ expression in the medionasal process in dup(10–20) and dup(8–21) but not in dup(13–20) embryos (**Supplementary Fig. 3**), confirming the key role of the 10–13 interval. Altogether, the 14 overlapping rearrangements showed that the 280-kb region between 10a and 13a contains specific regulatory information relevant to CL/P, defining a medionasal enhancer (MNE) region. To identify potential regulatory modules constituting the MNE region, we carried out chromatin immunoprecipitation (ChIP) on face samples from E11.5 embryos for enhancer-associated histone modifications[11,12] (**Fig. 1g** and **Supplementary Table 3**). The 10–13 interval contained several regions enriched for acetylation of histone H3 at lysine 27 (H3K27ac) and monomethylation of histone H3 at lysine 4 (H3K4me1) (**Fig. 1g**, bottom). These regions are often evolutionarily conserved and were not detected in other tissues or cell types[13], suggesting that they have a specific function in the face. Supporting this hypothesis, it was recently shown that the human sequence

[1]Developmental Biology Unit, European Molecular Biology Laboratory (EMBL), Heidelberg, Germany. [2]European Bioinformatics Institute–European Molecular Biology Laboratory (EMBL-EBI), Wellcome Trust Genome Campus, Hinxton, Cambridge, UK. Correspondence should be addressed to F.S. (spitz@embl.de).

Received 13 December 2013; accepted 8 April 2014; published online 25 May 2014; doi:10.1038/ng.2971

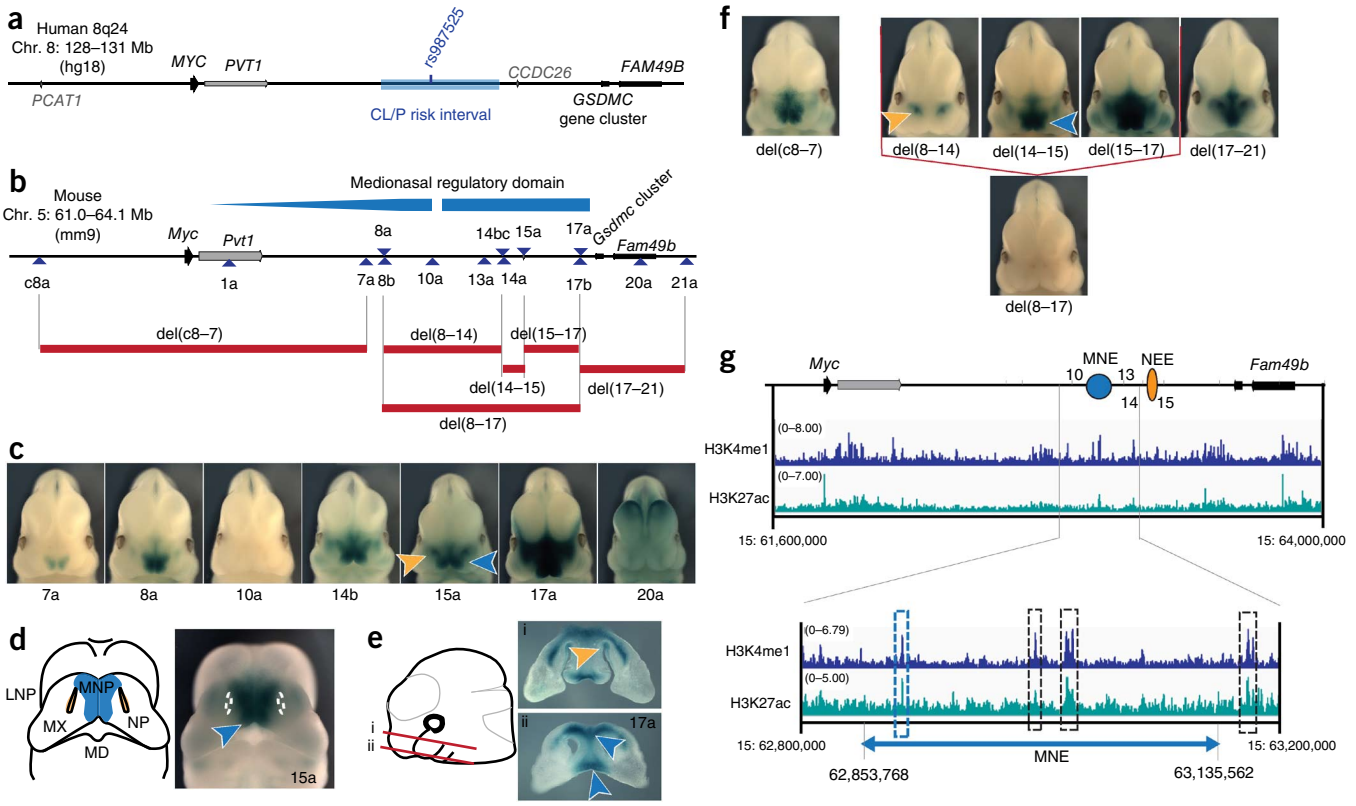

Figure 1 Functional characterization of the 8q24 CL/P regulatory landscape. (**a**) The human 8q24 interval associated with CL/P risk. Genes are shown as plain arrows (black, protein-coding genes; gray, annotated noncoding transcripts). The CL/P interval[3] and the most significantly associated SNP (rs987525) are indicated in blue. (**b**) Syntenic organization of the mouse locus, depicting transposon insertions (blue triangles) and deletions (red bars) used in this study. An expanded list of insertions and alleles is given in **Supplementary Table 1**. The expression patterns of adjacent insertions (shown in **c**) define a broad 'medionasal regulatory domain' indicated by a blue bar whose width represents relative LacZ expression levels. (**c**) LacZ staining of E11.5 embryos with various insertions, with insertion number indicated below. Arrowheads indicate expression in the medionasal process (blue) and in the nasal epithelium (orange). Insertions located in the *Gsdmc* gene cluster did not result in any expression, whereas the more telomeric ones around *Fam49b* resulted in a different expression pattern. (**d**,**e**) Magnified view (**d**) and sections (**e**) of E11.5 embryos with strong LacZ expression in the medionasal process (MNP; blue) and nasal epithelium (NP; orange). MX, maxillary process; MD, mandibulary process; LNP, lateral nasal process. Most insertions in the 7a–17a interval also showed LacZ expression in somites and in the limb mesenchyme of E11.5 embryos but not in heart or liver[9]. (**f**) LacZ staining in the faces of E11.5 embryos carrying the different deletions. The medionasal process and nasal epithelium expression domains can be separated by different deletions, highlighting two distinct regulatory regions, the MNE and NEE regions, respectively. See also the description of additional lines and sections in **Supplementary Figures 2** and **3**. Precise genomic positions are given in **Supplementary Table 2**. (**g**) Enrichment for enhancer-associated marks (blue, H3K4me1; turquoise, H3K27ac) profiled by ChIP sequencing (ChIP-seq) of E11.5 facial tissues highlighted several candidate face-specific regulatory regions (dashed boxes; positions and evolutionary conservation are given in **Supplementary Table 3**). One representative profile of two biological replicates is shown for each track. The blue boxed peak corresponds to Vista element hs1877, which showed enhancer activity in a number of regions, including the facial mesenchyme[14].

corresponding to one of these face-specific H3K27ac and H3K4me1 peaks showed, in E11.5 transgenic mouse embryos, enhancer activity in the face, among other domains[14].

Most of the genes surrounding this large region are expressed in the developing face (**Supplementary Fig. 4**), and their promoters show active chromatin marks (**Supplementary Table 3**). To identify MNE region–regulated gene(s), we measured by quantitative PCR (qPCR) the expression of these genes in dissected tissues from wild-type E11.5 embryos and those with deletion of the MNE region (**Fig. 2a**). In the face of del(8–17) embryos, *Myc* expression was reduced to 15% of its normal expression level. None of the other surrounding genes tested showed significant ($P < 0.05$) changes in expression. Notably, *Myc* levels in embryonic heart and liver were unaffected (**Fig. 2b**), showing that the 8–17 interval controls *Myc* expression only in tissues where it has enhancer activity. Whole-mount *in situ* hybridization confirmed strong reduction in *Myc* transcript levels in the face but not in the liver of del(8–17) homozygous embryos (**Fig. 2c**). The smaller deletion of the 8–14 interval, which retains the MNE region, led to a similar reduction in *Myc* expression in the developing face, whereas deletions that did not include the MNE region (**Fig. 2d**) or simple transposon insertions (**Supplementary Fig. 5**) had no or minor effects. From these experiments, we conclude that the MNE region controls *Myc* expression specifically in the future upper lip. If the MNE region regulates *Myc* expression in *trans*, both *Myc* alleles should be equally affected. Instead, in compound embryos carrying del(8–17) on one chromosome and a *GFP*-tagged *Myc* gene[15] on the other, *GFP* expression was not diminished (but was instead slightly higher), whereas the *Myc* allele on the same chromosome as del(8–17) was downregulated to 30% of wild-type levels, after normalization to *GFP* levels (**Fig. 2e**). The very different response of the two alleles demonstrates that the MNE region acts on *Myc* in *cis* across more than 1 Mb of DNA. The intermediate effects of heterozygous deletions of the MNE region on *Myc* expression also supported direct regulation in *cis*.

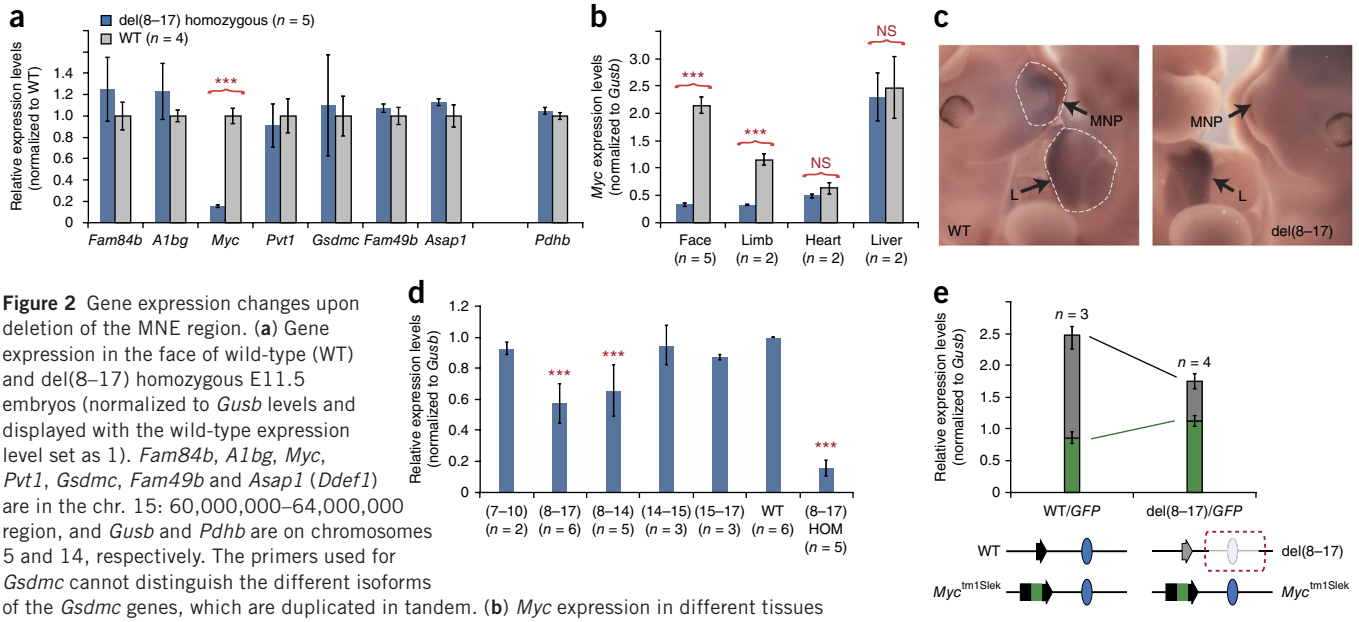

Figure 2 Gene expression changes upon deletion of the MNE region. (**a**) Gene expression in the face of wild-type (WT) and del(8–17) homozygous E11.5 embryos (normalized to *Gusb* levels and displayed with the wild-type expression level set as 1). *Fam84b*, *A1bg*, *Myc*, *Pvt1*, *Gsdmc*, *Fam49b* and *Asap1* (*Ddef1*) are in the chr. 15: 60,000,000–64,000,000 region, and *Gusb* and *Pdhb* are on chromosomes 5 and 14, respectively. The primers used for *Gsdmc* cannot distinguish the different isoforms of the *Gsdmc* genes, which are duplicated in tandem. (**b**) *Myc* expression in different tissues from E11.5 embryos. Values are expressed after normalization to *Gusb* levels (set as 1 in each tissue). (**c**) Loss of *Myc* expression detected by whole-mount *in situ* hybridization in the medionasal process (MNP) but not in the liver (L) of del(8–17) homozygous embryos relative to wild-type E11.5 embryos. Staining in control embryos is highlighted by dashed circles. (**d**) *Myc* expression in the face of E11.5 embryos heterozygous for different deletions. Values are expressed after normalization to *Gusb* levels (set as 1 in each tissue). (**e**) Allele-specific analysis of *Myc* expression in the faces of E11.5 embryos showed that the MNE region acts in *cis*. The reference allele $Myc^{tm1Slek}$ carries an insertion of *GFP* in the second exon of *Myc*[15], and its expression is detected with *GFP*-specific primers. Expression levels of the other alleles (wild type or del(8–17)) were determined by subtracting *GFP* expression levels from the overall *Myc* expression levels (the *Myc*-specific primers used amplified both alleles). *Myc* expression levels are significantly lower in del(8–17) than in wild-type mice (Student's *t* test, $P < 5.6 \times 10^{-5}$), whereas *GFP* expression is instead higher. In all charts, error bars represent s.d. ***$P < 0.005$, Student's *t* test; NS, not statistically significant ($P > 0.05$).

We then investigated the physiological consequences of the deletion of the MNE region on facial development. Mice homozygous for del(8–17) were born in normal mendelian proportions, without gross abnormalities or the several global skull malformations observed upon *Myc* deletion in the neural crest lineage[16]. Homozygous del(8–17) mice usually showed impaired postnatal growth, and some died before weaning, suggesting that the 8–17 interval contains additional elements[9] that may regulate *Myc* expression in other organs and tissues. Interestingly, compared to their littermates, del(8–17) homozygous mice had a significantly smaller snout ($P < 0.005$) (**Fig. 3** and **Supplementary Fig. 6a**), with an abnormal suture of the nasal and frontal bones (**Fig. 3b**). Other craniofacial regions were unaffected (**Fig. 3c**), and we did not observe palatal clefts in any of the pups. Changes in face morphology could already be detected in embryos, with a reduction in the width of the medionasal process in del(8–17) embryos compared to control E11.5 embryos (**Fig. 3d**). We did not detect any change in apoptosis in the developing medionasal prominence at E11.5 (data not shown). However, we found a small but significant reduction in the number of mitotic cells ($P < 0.05$) (**Fig. 3e** and **Supplementary Fig. 6b**), suggesting abnormal cell proliferation. Notably, we recovered a small number of E14–E15 embryos heterozygous and homozygous for del(8–17) with cleft lip and/or abnormal facial development (**Fig. 3f**): out of 121 embryos produced from heterozygous × heterozygous crosses, 4 del(8–17) embryos showed cleft lip and 4 other del(8–17) embryos showed cleft palate. The normal rate of spontaneous CL/P on the C57BL/6J background, on which our alleles are maintained, is extremely low[17], and none of the 35 wild-type embryos from these crosses had abnormal craniofacial morphology (Fisher's exact probability test $P = 0.0589$). Thus, these sporadic cases further support the involvement of this region in susceptibility to CL/P.

To identify genetic pathways that might contribute to the increased incidence of CL/P, we compared the mRNA transcriptomes of the medionasal regions in del(8–17) and control embryos at E11.5. We found 100 genes that showed significant (false discovery rate (FDR) < 0.05), albeit mild, differences in expression (**Fig. 4a**, **Supplementary Figs. 7** and **8**, and **Supplementary Table 4**). We tested and confirmed differences in expression for a subset of these genes using qPCR in del(8–17) and del(8–14) homozygous embryos. We also showed that the expression of these genes was not altered in del(14–17) embryos, in which the deleted interval does not include the MNE region (**Fig. 4b**). Consistent with our qPCR analysis, *Myc* expression was reduced to 17% of wild-type levels, and we found no significant changes for the genes located in the surrounding 5 Mb of sequence (**Supplementary Table 5**). Gene Ontology (GO) analysis showed that the majority of the genes misexpressed in del(8–17) embryos were involved in ribosome assembly and translational control (**Supplementary Tables 6** and **7**). For these genes, the fold change in expression was small (20–25%), but all showed lower expression levels in del(8–17) embryos and enrichment of the associated terms was maintained even when considering only highly expressed genes. Interestingly, MYC has already been shown to be a direct regulator of ribosome biogenesis in cancer and cellular models[18]. Our results stress the notion that physiological, non-oncogenic expression levels of *Myc* modulate ribosomal and translation-associated protein expression. We also found significant yet mild misexpression of a number of transcription factors and signaling molecules known to be important for facial development (**Supplementary Table 8**). Intriguingly, the altered frontonasal sutures in del(8–17) mice resembled the phenotype induced by facial-specific knockout of *Tfap2a*[19]. Whereas *Tfap2a* expression seemed unaltered in del(8–17) embryos, *Nr2f1*, whose product, together with TFAP2A, binds to a large subset of neural crest cell (NCC) enhancers[20], showed

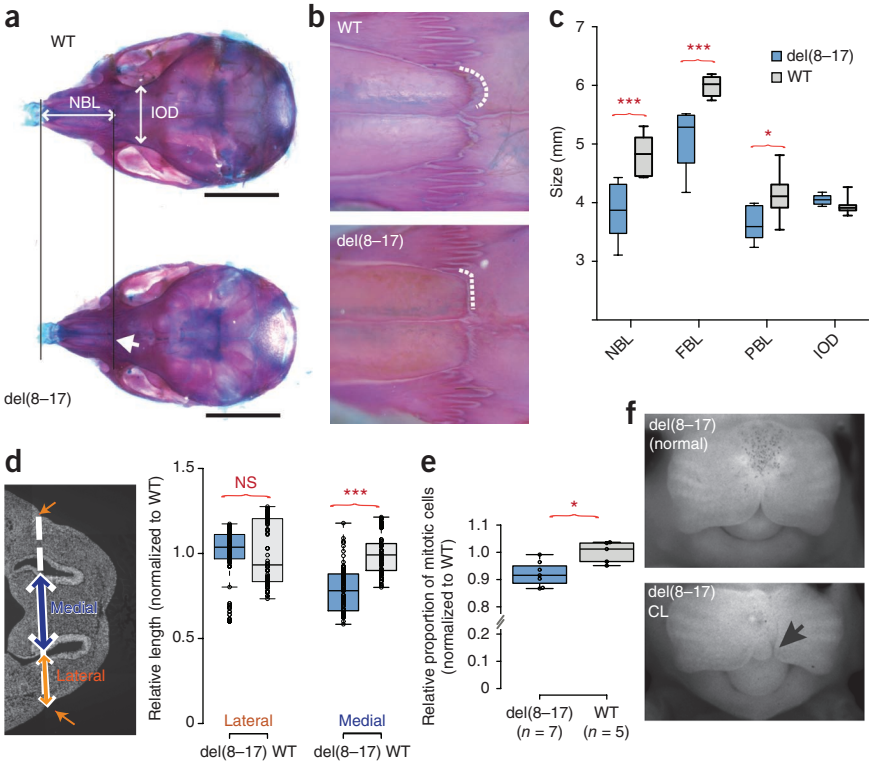

Figure 3 Facial dysmorphologies upon deletion of the CL/P-associated 8q24 region. (**a**) Comparison of the skulls of del(8–17) homozygous mice to those of their wild-type littermates. Dorsal views of representative skulls from 3-week-old wild-type and mutant (del(8–17)) mice, stained with Alcian blue (cartilage) and Alizarin red (bone). IOD, interorbital distance; NBL, nasal bone length. Scale bars, 5 mm. (**b**) Enlarged views of the frontonasal regions showing the abnormal suture in del(8–17) mice (highlighted by dashed lines). (**c**) Comparison of different bone lengths and skull measures (FBL, frontal bone length; PBL, parietal bone length) in 3-week-old (del(8–17), $n = 5$; wild type, $n = 7$) mice. Del(8–17) mice showed reduced nasal and frontal bone lengths (Student's t test, $P = 0.0033$ and 0.0028, respectively). Box plots show medians and first and third quartiles. Whiskers indicate minimum and maximum values. (**d**) Comparison of the widths of the lateral and medial parts of the developing face in wild-type ($n = 3$) and del(8–17) ($n = 4$) E11.5 embryos. The landmarks (arrows) used to compare the lateral and medial widths of the developing faces of E11.5 embryos are shown on the left. NS = not significant. (**e**) Quantification of proliferation in the faces of wild-type ($n = 5$) and (del(8–17) homozygous ($n = 7$) E11.5 embryos. Values represent the proportion of mitotic cells, determined by staining for phosphorylated histone H3, averaged from 10–30 serial sections per embryo. In **d** and **e**, box plots show medians and first and third quartiles. Whiskers indicate 1.5 times the interquartile ranges of the first and third quartiles. $^*P = 0.011$ (Student's t test). (**f**) Most E14.5 del(8–17) embryos have normal face morphology (top), but a minority have CL/P (arrow) with other craniofacial malformations (bottom). In all panels, statistical significance was determined with Student's t tests. $^{***}P < 0.005$, $^*P < 0.05$.

lower expression levels. TFAP2A also interacts both genetically and directly with MYC[21,22]. These interactions may contribute to propagating the consequences of *Myc* downregulation in the facial NCC gene regulatory network (GRN).

The distinct facial morphology of mice carrying deletion of the MNE region and the sporadic CL/P cases observed demonstrate the key functional role of this region in craniofacial morphogenesis. These findings support the implication of genetic variation at 8q24 in CL/P[3–5] and in normal variation in facial shape[23,24] in humans. They refine the proposed critical interval to a smaller region, distal to the marker most significantly associated with CL/P, rs987525 (ref. 3; **Supplementary Fig. 9**).

At the molecular level, we show that this region is a key tissue-specific *cis*-regulatory enhancer for *Myc*. Its extremely remote location further emphasizes the role of distant regulatory elements in controlling gene activity[14,25], not only for tissue-specific developmental regulators but also for genes with more general functions[26]. The MNE

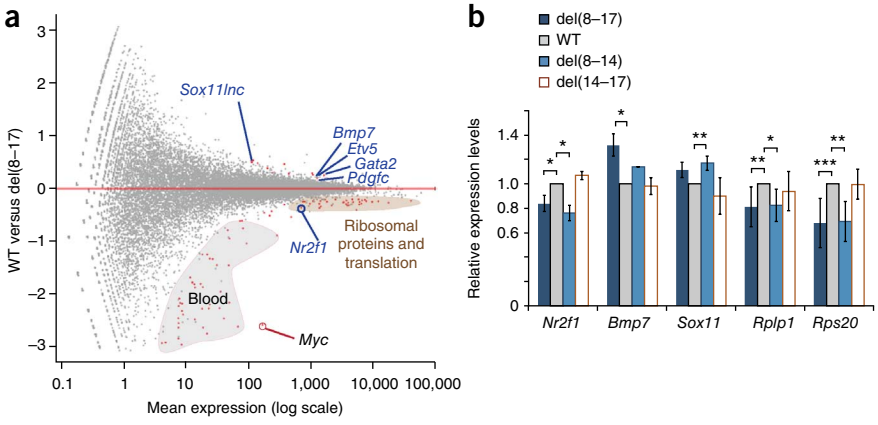

Figure 4 Alterations in gene expression upon deletion of the CL/P-associated 8q24 region. (**a**) Changes in expression measured by RNA sequencing in the medial faces of del(8–17) homozygous embryos (four replicate libraries from four different embryos) compared to wild-type controls (four replicate libraries from littermates). Genes with significant (FDR < 0.05) changes in expression are shown in red. See **Supplementary Table 4** for a list of the misregulated genes and **Supplementary Figure 7** for heat-map representations of the data. RNA sequencing analysis also showed downregulation of several blood-specific genes; their presence arose from the small blood vessels in the dissected tissues and suggested an additional role of the 8–17 interval in hematopoiesis, in which *Myc* has an important role[49]. Notably, the genomic region involved appears to be distinct from the MNE region, as the downregulation of *Apoe* and *Csf1* was observed in del(14–17) but not in del(8–14) embryos (**Supplementary Fig. 8**). (**b**) Validation by qPCR of expression changes for some of the genes identified by RNA sequencing. The levels of *Rplp1* and *Rps20* are significantly ($^{***}P < 0.005$, $^{**}P < 0.01$, Student's t test) decreased to similar extents in the medial faces of del(8–17) ($n = 4$ to 7) and del(8–14) ($n = 3$) homozygote E11.5 embryos but not in del(14–17) ($n = 3$) homozygote E11.5 embryos. *Nr2f1* levels also appeared lower ($^*P < 0.05$, Student's t test). Error bars, s.d.

region comprises several adjacent blocks carrying enhancer-associated epigenetic marks, which may act collectively to control *Myc* expression, as described for other loci[27–29]. In such situations, deletions of individual modules may show only limited or subtle effects[14,27] and may lead to an underestimation of their collective role. Indeed, it was recently proposed that GWAS may identify haplotypes with multiple polymorphisms affecting adjacent enhancers instead of a single 'master' element[30]. For this reason, larger rearrangements, such as the ones used here for 8q24 or earlier to characterize the 9p21 coronary artery disease (CAD)-associated interval[31], may be a useful strategy in following up GWAS findings, particularly for intervals with multiple candidate enhancer regions[32].

The different types of genes found to be misexpressed upon deletion of the MNE region give insights into the nature of the risk associated with variants in 8q24 (**Supplementary Fig. 10**). MYC is known for its role in NCCs[16,33], and our analyses suggest that the altered expression levels of downstream genes involved in facial development observed with deletion of the MNE region may account for the slight but penetrant concurrent facial dysmorphism. In this context, orofacial clefts may constitute the extreme end of a spectrum of normal phenotypic variation[23,34]. In addition to this hypothesis, we suggest that the global misexpression of ribosomal subunits and translation factors observed in del(8–17) mice may contribute to CL/P incidence through a distinct process. Heterozygous mutations of ribosomal genes, including *Rps7* and *Rps20*, whose expression was downregulated in del(8–17) mice, have phenotypic consequences in mice and humans[35–37]. Of note, humans with Diamond-Blackfan anemia, which is caused by mutations affecting diverse ribosomal proteins, are also prone to CL/P malformations[38]. Abundance of ribosomal proteins therefore constitutes a limiting factor: even mild, reduced expression of ribosomal and translation regulatory genes in del(8–17) mice may affect cellular metabolic capacity, leading to a general but spatially restricted sensitivity. This phenomenon may account for the influence of diverse metabolic and environmental conditions (for example, tobacco or alcohol use, diabetes or folate deficiency) that have been proposed to contribute to CL/P etiology[6] by transforming broad-spectrum but minor growth perturbations into growth defects affecting a specific embryonic process. The combination of these different effects might explain why this region has such a high population attributable risk[3] and a prominent influence on facial (dys)morphologies.

Interestingly, GWAS for other traits—prominently, susceptibility to diverse types of cancer—have identified associations with other regions within 8q24 (refs. 39–42), which were proposed to coincide with variants in tissue-specific enhancers for *Myc* or other genes[43–47]. As illustrated recently[48], validating or determining the role of these variants *in vivo* will be essential. The series of overlapping deletions and rearrangements described here opens up avenues to further dissect this important regulatory landscape.

URLs. TRACER, http://tracerdatabase.embl.de/; HTSeq-count, http://www-huber.embl.de/users/anders/HTSeq/doc/index.html.

METHODS

Methods and any associated references are available in the online version of the paper.

Accession codes. ChIP-seq data have been deposited in the Gene Expression Omnibus (GEO) under accessions GSM1279691, GSM1279692 and GSM1279693. RNA sequencing data have been deposited in GEO under accessions GSM1279694–GSM1279701.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

We thank B. Sleckman (Washington University, St. Louis) and A. Trumpp (DKFZ, Heidelberg) for providing the *Myc*^{tm1Slek} strain. We thank members of the EMBL Laboratory Animal Resources Facility for animal welfare and husbandry; the EMBL Genomics Core Facility for advice and support in processing ChIP and RNA sequencing experiments; and Genome Biology Computational Support for help with the analyses. We thank members of the Spitz laboratory and colleagues at EMBL for sharing reagents and helpful comments. V.V.U. and M.P. were supported by PhD fellowships from the Jeff Schell Darwin Trust and the EMBL International PhD program, respectively. This work was supported by EMBL.

AUTHOR CONTRIBUTIONS

F.S. designed the experiments. V.V.U., M.P., S.R. and K.L. performed the experiments. N.A.F. and J.C.M. performed RNA-seq data, bioinformatics and statistical analyses. V.V.U., M.P. and F.S. analyzed the data. F.S. wrote the manuscript with V.V.U., M.P. and J.C.M.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

1. Dixon, M.J., Marazita, M.L., Beaty, T.H. & Murray, J.C. Cleft lip and palate: understanding genetic and environmental influences. *Nat. Rev. Genet.* **12**, 167–178 (2011).
2. Mangold, E., Ludwig, K.U. & Nöthen, M.M. Breakthroughs in the genetics of orofacial clefting. *Trends Mol. Med.* **17**, 725–733 (2011).
3. Birnbaum, S. *et al.* Key susceptibility locus for nonsyndromic cleft lip with or without cleft palate on chromosome 8q24. *Nat. Genet.* **41**, 473–477 (2009).
4. Beaty, T.H. *et al.* A genome-wide association study of cleft lip with and without cleft palate identifies risk variants near *MAFB* and *ABCA4*. *Nat. Genet.* **42**, 525–529 (2010).
5. Ludwig, K.U. *et al.* Genome-wide meta-analyses of nonsyndromic cleft lip with or without cleft palate identify six new risk loci. *Nat. Genet.* **44**, 968–971 (2012).
6. Murray, J.C. Gene/environment causes of cleft lip and/or palate. *Clin. Genet.* **61**, 248–256 (2002).
7. Freedman, M.L. *et al.* Principles for the post-GWAS functional characterization of cancer risk loci. *Nat. Genet.* **43**, 513–518 (2011).
8. Gritli-Linde, A. The etiopathogenesis of cleft lip and cleft palate: usefulness and caveats of mouse models. *Curr. Top. Dev. Biol.* **84**, 37–138 (2008).
9. Ruf, S. *et al.* Large-scale analysis of the regulatory architecture of the mouse genome with a transposon-associated sensor. *Nat. Genet.* **43**, 379–386 (2011).
10. Jiang, R., Bush, J.O. & Lidral, A.C. Development of the upper lip: morphogenetic and molecular mechanisms. *Dev. Dyn.* **235**, 1152–1166 (2006).
11. Rada-Iglesias, A. *et al.* A unique chromatin signature uncovers early developmental enhancers in humans. *Nature* **470**, 279–283 (2011).
12. Heintzman, N.D. *et al.* Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature* **459**, 108–112 (2009).
13. Shen, Y. *et al.* A map of the *cis*-regulatory sequences in the mouse genome. *Nature* **488**, 116–120 (2012).
14. Attanasio, C. *et al.* Fine tuning of craniofacial morphology by distant-acting enhancers. *Science* **342**, 1241006 (2013).
15. Huang, C.-Y., Bredemeyer, A.L., Walker, L.M., Bassing, C.H. & Sleckman, B.P. Dynamic regulation of *c-Myc* proto-oncogene expression during lymphocyte development revealed by a *GFP–c-Myc* knock-in mouse. *Eur. J. Immunol.* **38**, 342–349 (2008).
16. Wei, K. *et al.* Neural crest cell deficiency of *c-myc* causes skull and hearing defects. *Genesis* **45**, 382–390 (2007).
17. Karolyi, J., Erickson, R.P. & Liu, S. Genetics of susceptibility to 6-aminonicotinamide–induced cleft palate in the mouse: studies in congenic and recombinant inbred strains. *Teratology* **37**, 283–287 (1988).
18. van Riggelen, J., Yetil, A. & Felsher, D.W. MYC as a regulator of ribosome biogenesis and protein synthesis. *Nat. Rev. Cancer* **10**, 301–309 (2010).
19. Nelson, D.K. & Williams, T. Frontonasal process–specific disruption of AP-2α results in postnatal midfacial hypoplasia, vascular anomalies, and nasal cavity defects. *Dev. Biol.* **267**, 72–92 (2004).
20. Rada-Iglesias, A. *et al.* Epigenomic annotation of enhancers predicts transcriptional regulators of human neural crest. *Cell Stem Cell* **11**, 633–648 (2012).
21. de Crozé, N., Maczkowiak, F. & Monsoro-Burq, A.H. Reiterative AP2a activity controls sequential steps in the neural crest gene regulatory network. *Proc. Natl. Acad. Sci. USA* **108**, 155–160 (2011).
22. Gaubatz, S. *et al.* Transcriptional activation by Myc is under negative control by the transcription factor AP-2. *EMBO J.* **14**, 1508–1519 (1995).

23. Boehringer, S. *et al.* Genetic determination of human facial morphology: links between cleft-lips and normal variation. *Eur. J. Hum. Genet.* **19**, 1192–1197 (2011).
24. Liu, F. *et al.* A genome-wide association study identifies five loci influencing facial morphology in Europeans. *PLoS Genet.* **8**, e1002932 (2012).
25. Visel, A., Rubin, E.M. & Pennacchio, L.A. Genomic views of distant-acting enhancers. *Nature* **461**, 199–205 (2009).
26. Hallikas, O. *et al.* Genome-wide prediction of mammalian enhancers based on analysis of transcription-factor binding affinity. *Cell* **124**, 47–59 (2006).
27. Montavon, T. *et al.* A regulatory archipelago controls *Hox* genes transcription in digits. *Cell* **147**, 1132–1145 (2011).
28. Marinić, M., Aktas, T., Ruf, S. & Spitz, F. An integrated holo-enhancer unit defines tissue and gene specificity of the Fgf8 regulatory landscape. *Dev. Cell* **24**, 530–542 (2013).
29. Hnisz, D. *et al.* Super-enhancers in the control of cell identity and disease. *Cell* **155**, 934–947 (2013).
30. Corradin, O. *et al.* Combinatorial effects of multiple enhancer variants in linkage disequilibrium dictate levels of gene expression to confer susceptibility to common traits. *Genome Res.* **24**, 1–13 (2014).
31. Visel, A. *et al.* Targeted deletion of the 9p21 non-coding coronary artery disease risk interval in mice. *Nature* **464**, 409–412 (2010).
32. Harismendy, O. *et al.* 9p21 DNA variants associated with coronary artery disease impair interferon-γ signalling response. *Nature* **470**, 264–268 (2011).
33. Bellmeyer, A., Krase, J., Lindgren, J. & LaBonne, C. The protooncogene *c-Myc* is an essential regulator of neural crest formation in *Xenopus*. *Dev. Cell* **4**, 827–839 (2003).
34. Parsons, T.E. *et al.* Phenotypic variability and craniofacial dysmorphology: increased shape variance in a mouse model for cleft lip. *J. Anat.* **212**, 135–143 (2008).
35. Terzian, T. & Box, N. Genetics of ribosomal proteins: "curiouser and curiouser.". *PLoS Genet.* **9**, e1003300 (2013).
36. Watkins-Chow, D.E. *et al.* Mutation of the Diamond-Blackfan anemia gene *Rps7* in mouse results in morphological and neuroanatomical phenotypes. *PLoS Genet.* **9**, e1003094 (2013).
37. McGowan, K.A. *et al.* Ribosomal mutations cause p53-mediated dark skin and pleiotropic effects. *Nat. Genet.* **40**, 963–970 (2008).
38. Ito, E., Konno, Y., Toki, T. & Terui, K. Molecular pathogenesis in Diamond-Blackfan anemia. *Int. J. Hematol.* **92**, 413–418 (2010).
39. Gudmundsson, J. *et al.* Genome-wide association study identifies a second prostate cancer susceptibility variant at 8q24. *Nat. Genet.* **39**, 631–637 (2007).
40. Tomlinson, I.P.M. *et al.* A genome-wide association study identifies colorectal cancer susceptibility loci on chromosomes 10p14 and 8q23.3. *Nat. Genet.* **40**, 623–630 (2008).
41. Shete, S. *et al.* Genome-wide association study identifies five susceptibility loci for glioma. *Nat. Genet.* **41**, 899–904 (2009).
42. Goode, E.L. *et al.* A genome-wide association study identifies susceptibility loci for ovarian cancer at 2q31 and 8q24. *Nat. Genet.* **42**, 874–879 (2010).
43. Tuupanen, S. *et al.* The common colorectal cancer predisposition SNP rs6983267 at chromosome 8q24 confers potential to enhanced Wnt signaling. *Nat. Genet.* **41**, 885–890 (2009).
44. Pomerantz, M.M. *et al.* The 8q24 cancer risk variant rs6983267 shows long-range interaction with *MYC* in colorectal cancer. *Nat. Genet.* **41**, 882–884 (2009).
45. Ahmadiyeh, N. *et al.* 8q24 prostate, breast, and colon cancer risk loci show tissue-specific long-range interaction with *MYC*. *Proc. Natl. Acad. Sci. USA* **107**, 9742–9746 (2010).
46. Jia, L. *et al.* Functional enhancers at the gene-poor 8q24 cancer-linked locus. *PLoS Genet.* **5**, e1000597 (2009).
47. Sotelo, J. *et al.* Long-range enhancers on 8q24 regulate *c-Myc*. *Proc. Natl. Acad. Sci. USA* **107**, 3001–3005 (2010).
48. Sur, I.K. *et al.* Mice lacking a *Myc* enhancer that includes human SNP rs6983267 are resistant to intestinal tumors. *Science* **338**, 1360–1363 (2012).
49. Wilson, A. *et al.* c-Myc controls the balance between hematopoietic stem cell self-renewal and differentiation. *Genes Dev.* **18**, 2747–2763 (2004).

ONLINE METHODS

Mouse alleles. Transposon insertion sites were generated and mapped as described previously[9]. Additional information on each insertion can be found on the TRACER website[50]. To distinguish the different *Myc* alleles, we used *Myc*^{tm1Slek} (ref. 15). The reference allele *Myc*^{tm1Slek} carries an insertion of the *GFP* gene in the second exon of *Myc*, and the encoded fusion protein appeared functional[15]. Yet, we noticed that the *GFP*-tagged allele was expressed at lower levels than the wild-type allele, suggesting that it might be a hypomorph. *Myc* is known to negatively regulate its own transcription through its promoter region[51], and this feedback can account for overall *Myc* levels in compound mice, on top of the specific effect of del(8–17) in *cis*.

Mouse alleles were genotyped by PCR of DNA purified from tail biopsies (for mice after birth) or yolk sacs (for embryos). Sequences for the primers for each insertion are provided in **Supplementary Table 9**. Deletions and duplications were generated by TAMERE[52] using the *Hprt*$^{tm1(cre)Mnn}$ line[53,54]. Rearrangements were further verified by long-range PCR (Expand Long-Range dNTPack, Roche), using primers flanking the reconstituted transposon. Mouse lines were maintained by backcrossing on a C57BL/6J background. Both male and female embryos were used for the experiments. Sample sizes were not predetermined.

Mouse experiments were conducted in accordance with the principles and guidelines in place at EMBL, as defined and overseen by its Institutional Animal Care and Use Committee, in accordance with European Convention 18/3/1986 and Directives 86/609/EEC and 2010/63/EU.

LacZ staining and *in situ* hybridization. LacZ staining and mRNA whole-mount *in situ* hybridization were performed as previously described[9]. Except where indicated, LacZ staining was performed on embryos heterozygous for the corresponding insertion, obtained by mating a transgenic male with a wild-type female.

Quantitative RT-PCR. Total RNA was purified from the dissected faces, forelimbs, hearts and livers of E11.5 embryos (45–48 somites) (RNeasy Mini kit, Qiagen) and eluted in 30–50 µl of nuclease-free water. For each RNA sample, we first measured concentration and purity with a Nanodrop ND-1000 spectrophotometer, and we then checked quality with a Bioanalyzer (Agilent Technologies) (**Supplementary Fig. 11a**). Possible DNA contamination was checked for by PCR using the 179039L and 179039R primers. Samples that did not meet quality control standards were discarded. First-strand cDNA was synthesized from 250–750 ng of total RNA using the Protoscript M-MuLV kit (NEB) with random hexamers. Quantitative RT-PCR was carried out in 0.1-ml and 0.2-ml wells of 96-well plates. We included 1 µl of cDNA, 1 µM primer mix and 10 µl of SYBR Green PCR Master Mix (Applied Biosystems) in a 20-µl final volume. Primers are listed in **Supplementary Table 10**. PCR was carried out on StepOnePlus Real-Time PCR systems (Life Technologies) and AB7500 Real-Time PCR systems (Applied Biosystems). Each primer pair was tested for efficiency and specificity following recommended procedures[55] (**Supplementary Fig. 11b**). We included *Gusb* and *Pdhb* as references. As *Gusb* expression levels (C_t values) were more in the range of those for the genes of interest, we used this gene for normalization, although the same significant differences were observed using normalization to *Pdhb* levels. Gene expression data were processed using Microsoft Excel and StepOne Real-Time PCR Software v2.0. For each group, two to five biological replicates were used (specific numbers are indicated in the figures), and, for each sample, quantification was performed on two to three technical replicates.

Skeletal preparation and analysis. Skulls from postnatal day (P) 19–21 and P40 mice were prepared following a standard procedure[56] and stained with 0.3% Alcian blue or 0.1% Alizarin red. For morphological measurements, ten landmarks from the top view of the skull and six landmarks from the bottom view were used[57]. Measurements were compared between age-matched groups. No mice were excluded from the analyses. Comparisons between the del(8–17) homozygous mice and wild-type controls were carried out with Student's *t* tests, and the statistical variance was calculated by *f* test. Exclusion of any of the samples did not have a critical impact on the statistical significance of the difference between genotypes.

Cell proliferation assays. Antibody to phosphorylated histone H3 (Ser10; 06-570, Millipore, rabbit polyclonal) was used at a 1:200 dilution in 10% FCS, 0.2% PBS with Tween-20 (PBST) blocking solution on 5-µm paraffin sections prepared from E11.5 embryos. Antigen retrieval was carried out in citrate buffer, pH 6.0 (Dako), and an Alexa Fluor 594–conjugated anti-rabbit secondary antibody was used at a 1:500 dilution with 1× 4′,6-diamidino-2-phenylindole (DAPI) in the blocking solution. Images were acquired at 10× and 20× magnification with a Zeiss CellObserver HS Automated wide-field microscope. Four to 10 images were stitched together to obtain a full E11.5 embryonic face section. Image analysis was carried out with Fiji software[58], using an automated macro to minimize subjective errors in counting nuclei positive for phosphorylation of histone H3. In brief, for each section, phosphorylated histone H3 signal was captured with the DsRed filter, and the signal obtained with the GFP-filtered image was subtracted out to remove background autofluorescence. The resulting signal image (S) was subjected to background rolling, Gaussian blur and Renyi Entropy dark thresholding. To define parameters for the automated counting of particles positive for phosphorylated histone H3 in 10× and 20× wide-field images, mitosis-phase particles were counted manually on 63× magnification images of the two sections that covered the whole face, obtained by stitching images recorded with a Spinning-Disk Microscope (PE Ultraview VoX) and using Volocity software (PerkinElmer). Particle counting was then performed and overlaid on the DAPI image: particle signals that were too elongated or did not overlap with DAPI signal were eliminated manually to obtain the final number of particles positive for phosphorylated histone H3 (pSc). The average number of pSc particles for a wild-type sample was 291. Cell numbers were estimated by measuring DAPI area (aD) on each section, as current Fiji plug-ins for DAPI signal counting were computationally inefficient for large samples. We verified that DAPI area and cell count performed with the Cell Counter plug-in (I. Levenfus, TU Dresden) showed a linear correlation of >99.5% in the range of 20 to 8,000 cells. Data were collected for sections from seven del(8–17) homozygous embryos and five wild-type littermate controls. The proliferation rate was defined as pSc/aD for each section. On average, proliferation rate was higher in the 130 sections from the 5 wild-type embryos than in the 165 sections from the 7 del(8–17) homozygous embryos (Student's *t* test, $P < 1 \times 10^{-5}$) (**Supplementary Fig. 6b**). The average proliferation rate was also calculated for each embryo (as the average of the values measured for all the sections from one embryo). Again, the seven del(8–17) homozygous embryos showed lower values than the five controls (Student's *t* test, $P < 0.05$) (**Fig. 3e**).

ChIP-seq experiments. For a given ChIP experiment, 25 facial mesenchyme samples were collected from E11.5 embryos, pooled, minced and fixed for 10 min in 1% formaldehyde in PBS. Formaldehyde was quenched through incubation with 250 mM glycine for 1 min on ice, and cells were lysed in buffer A (10 mM HEPES, pH 8.0, 1 mM EDTA, 0.5 mM EGTA, 0.25% Triton X-100 and 1:100 freshly added PMSF) for 10 min at 4 °C on a rotating wheel. Cells were then incubated with buffer B (10 mM HEPES, pH 8.0, 200 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 0.01% Triton X-100 and 1:100 freshly added PMSF) for 10 min at 4 °C on a rotating wheel. Cells were resuspended in 300 µl of sonication buffer (10 mM Tris-HCl, pH 8.0, 1 mM EDTA, pH 8.0, 0.1% SDS and 1:100 freshly added PMSF) and sonicated with a Bioruptor for 24 min, with a duty cycle of 0.5 on the high-power setting. An aliquot of 2% of chromatin was taken as the input control. Chromatin was then equilibrated to 300 µl of RIPA buffer and immunoprecipitated with 1.2 µg of antibody to H3K27ac (Abcam, ab4729) or 0.75 µg of antibody to H3K4me1 (Abcam, ab8580) with rotation overnight at 4 °C. Protein A Dynabeads (10001D, Life Technologies) were added (20 µl per tube), incubated for 5 h in rotation and washed four times in RIPA buffer and 1 time in Tris-EDTA. Immunoprecipitated complexes were eluted in 100 µl of elution buffer, and cross-links were reversed at 65 °C overnight. Afterward, 20 µg of proteinase K was added to each tube, and samples were incubated for 2 h in 55 °C. Finally, DNA was purified with QIAquick columns.

Libraries were prepared with NEB Next ChIP-seq Sample Prep Master Mix Set I (E6240S, NEB) and NEB Next Multiplex Oligos for Illumina (E7335S, NEB) and sequenced with a depth of 100 million reads for the input samples and 30 million reads for the ChIP samples.

RNA sequencing. The facial tissues of four del(8–17) homozygous and four wild-type E11.5 littermate embryos were dissected and treated independently as biological replicates. *Drosophila melanogaster* S2 cells (72,000) were spiked into each tube to detect any global perturbation in mRNA levels. Total RNA was isolated with an RNeasy Mini kit (Qiagen). RNA quality and quantity were measured with a 2100 Bioanalyzer (Agilent Technologies). RNA samples were prepared according to the TruSeq RNA sample preparation guide (Illumina). We performed 50-bp, single-end sequencing on an Illumina HiSeq instrument. The eight samples were barcoded and run on the same flow cell. RNA sequencing data were processed with the iRAP (version 0.2.0; N.A.F., R. Petryszak, J.C.M. and A. Brazma, unpublished data) analysis pipeline (with default settings). More specifically, reads were mapped to the mouse reference genome (GRCm38.72) using TopHat[59] (version 1.4.1) with the default number of mismatches. Read counts for each gene were estimated by running HTSeq-count (version 0.5.3.p9) with intersection non-empty mode. Data normalization and identification of genes with significantly different expression levels were performed using DESeq (version 1.10.1)[60]. *P* values were adjusted for multiple testing using the Benjamini-Hochberg procedure, with a corrected *P* value of 0.05 being used to identify differentially expressed genes. The ratio of reads mapped to the mouse genome and to the *Drosophila* genome was similar for all samples, suggesting the absence, in this context and for this range of *Myc* expression levels, of a global amplifier effect that would change the expression levels of most genes[61,62].

50. Chen, C.-K. *et al.* TRACER: a resource to study the regulatory architecture of the mouse genome. *BMC Genomics* **14**, 215 (2013).
51. Facchini, L.M., Chen, S., Marhin, W.W., Lear, J.N. & Penn, L.Z. The Myc negative autoregulation mechanism requires Myc-Max association and involves the *c-myc* P2 minimal promoter. *Mol. Cell. Biol.* **17**, 100–114 (1997).
52. Hérault, Y., Rassoulzadegan, M., Cuzin, F. & Duboule, D. Engineering chromosomes in mice through targeted meiotic recombination (TAMERE). *Nat. Genet.* **20**, 381–384 (1998).
53. Tang, S.-H.E., Silva, F.J., Tsark, W.M.K. & Mann, J.R. A Cre/*lox*P-deleter transgenic line in mouse strain 129S1/SvImJ. *Genesis* **32**, 199–202 (2002).
54. Wu, S., Ying, G., Wu, Q. & Capecchi, M.R. Toward simpler and faster genome-wide mutagenesis in mice. *Nat. Genet.* **9**, 922–930 (2007).
55. Bustin, S.A. *et al.* The need for transparency and good practices in the qPCR literature. *Nat. Methods* **10**, 1063–1067 (2013).
56. Nagy, A., Gertsenstein, M., Vintersten, K. & Behringer, R.R. *Manipulating the Mouse Embryo* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 2003).
57. Kawakami, M. & Yamamura, K.-I. Cranial bone morphometric study among mouse strains. *BMC Evol. Biol.* **8**, 73 (2008).
58. Schindelin, J. *et al.* Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).
59. Trapnell, C., Pachter, L. & Salzberg, S.L. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* **25**, 1105–1111 (2009).
60. Anders, S. & Huber, W. Differential expression analysis for sequence count data. *Genome Biol.* **11**, R106 (2010).
61. Nie, Z. *et al.* c-Myc is a universal amplifier of expressed genes in lymphocytes and embryonic stem cells. *Cell* **151**, 68–79 (2012).
62. Lin, C.Y. *et al.* Transcriptional amplification in tumor cells with elevated c-Myc. *Cell* **151**, 56–67 (2012).

doi:10.1038/ng.2971