

candidate genes (*AGER*, *TNSI*, *HTRA*, *THSD4* and *GSTCD*) were shown to be expressed in lung tissue². These loci are all biologically plausible candidates that are involved in pathways known to be important in lung biology, including immune function, muscle function or inflammation. These studies did not replicate the *SERPINE2* association and also did not find evidence for association to the 15q *CHRNA5-CHRNA3-CHRNA4* locus previously associated with nicotine dependence, lung cancer and COPD^{11–13}.

Clinical relevance

Lung function levels depend strongly on age, height and gender, with distinct changes from childhood to adulthood (Fig. 1). Thus, because of study-design considerations, some gene variants identified in the two studies reported here may simply influence lung function in nondiseased subjects and will have no independent effect on patients with airways disease. However, the high within-

individual consistency of spirometric lung function measures over time, and the links of the identified candidate genes to lung development, suggest that genetic variations associated with lung function development might be important genetic determinants of lung function in both healthy individuals and those with airways disease (asthma and COPD). To test this hypothesis, the newly associated variants should be tested in both individuals with asthma and those with COPD. These studies should also be extended to include bronchodilator and bronchoconstrictor responses.

These loci account for a small proportion of the variation in FEV₁ and the FEV₁/FVC ratio. Repapi *et al.*² estimates that the five new loci identified in their study account for ~0.14% of the variation in FEV₁/FVC ratio. Translating this to clinical prediction remains further away, and the first steps toward this goal include additional validation, refinement of the genetic loci and determination

of the specific genes and their functional variants, as well as more accurate estimation of effect sizes from additional studies of unaffected and diseased subjects.

1. O'Connor, G., Sparrow, D. & Weiss, S.T. *Am. Rev. Respir. Dis.* **140**, 225–252 (1989).
2. Repapi, E. *et al. Nat. Genet.* **42**, 36–44 (2010).
3. Hancock, D.B. *et al. Nat. Genet.* **42**, 45–52 (2010).
4. Redline, S. *et al. Am. J. Epidemiol.* **129**, 827–836 (1989).
5. Xu, X. *et al. Am. J. Respir. Crit. Care Med.* **160**, 1928–1933 (1999).
6. Wilk, J.B. *et al. BMC Med. Genet.* **8** (Suppl. 1), S8 (2007).
7. Wilk, J.B. *et al. PLoS Genet.* **5**, e1000429 (2009).
8. Pillai, S.G. *et al. PLoS Genet.* **5**, e1000421 (2009).
9. DeMeo, D.L. *et al. Am. J. Hum. Genet.* **79**, 253–264 (2006).
10. Zhu, G. *et al. Am. J. Respir. Crit. Care Med.* **176**, 167–173 (2007).
11. Amos, C.I. *et al. Nat. Genet.* **40**, 616–622 (2008).
12. Hung, R.J. *et al. Nature.* **452**, 633–637 (2008).
13. Thorgeirsson, T.E. *et al. Nature.* **452**, 638–642 (2008).
14. Speizer, F.E. & Tager, I.B. *Epidemiol. Rev.* **1**, 124–142 (1979).

Getting connected in the globin interactome

Tobias Ragozy & Mark Groudine

A new study provides compelling evidence that transcriptional regulation and three-dimensional genomic architecture are linked. The alpha- and beta-globin loci associate with hundreds of active genes across the genome at transcription factories in erythroid cells, and specialized Klf1-containing transcription factories mediate the association of Klf1-regulated genes.

The organization of a eukaryotic nucleus reflects its specific expression profile^{1,2}. On the genomic scale, this translates to preferred tissue-specific chromatin folding and chromosome positioning within interphase nuclei, including intermingling of looped segments within and between chromosomes^{3,4}. Consequently, co-regulated genes tend to localize near each other at activity hubs. However, the description of colocalized and coexpressed genes has been anecdotal and limited to a few genes. What had been missing until recently^{5,6} was information on how extensive and consistent these associations are at the genomic level. On page 53 of this issue, Peter Fraser and colleagues⁷ use an array of biochemical and cytogenetic approaches to elegantly describe

such a genomic transcription interactome by exploring the gene associations of the mouse α - (*Hba*) and β -globin (*Hbb*) loci in erythroid cells.

Globins cast a wide net

The globin genes are highly expressed in developing erythrocytes, and they consistently associate with transcription factories, which are foci of hyperphosphorylated RNA polymerase II (RNAPII) and are sites of active transcription. Using double-label RNA fluorescence *in situ* hybridization (FISH), Schoenfelder *et al.*⁷ discovered that several expressed erythroid-specific genes (of a set of 33 examined loci on 15 chromosomes) frequently localize to the same transcription factories as *Hba* or *Hbb*. To expand on this initial observation, Schoenfelder *et al.* used a modified chromosome conformation capture technique (enhanced ChIP-4C, or e4C) to establish the genomic interactome of the globin loci. They precede the assay with a chromatin immunoprecipitation (ChIP)

against hyperphosphorylated RNAPII and a bait-specific (globin) enrichment step involving a biotin pulldown. These clever adaptations ensure that the detected associations are restricted to actively transcribed genes and thus result in a highly improved signal-to-noise ratio. Consequently, the results provide the first truly robust and reproducible large-scale detection of long-range and interchromosomal associations between active genes. Microarray analysis of the e4C material confirms known interactions between *Hbb* and erythroid-specific genes (for example, *Eraf* and *Uros*) and hundreds more. Strikingly, around 90% of these detected interactions were in *trans* for both *Hbb* and *Hba*. The authors performed elaborate controls to demonstrate that virtually all ligation products obtained via the e4C protocol originated from individual cross-linked complexes and thus represent bona fide long-distance interactions. Moreover, paired-end tag sequencing (PET) of phosphorylated RNAPII ChIP material confirmed

Tobias Ragozy and Mark Groudine are in the Division of Basic Sciences, Fred Hutchinson Cancer Research Center, Seattle, Washington, USA.
e-mail: markg@fhcrc.org

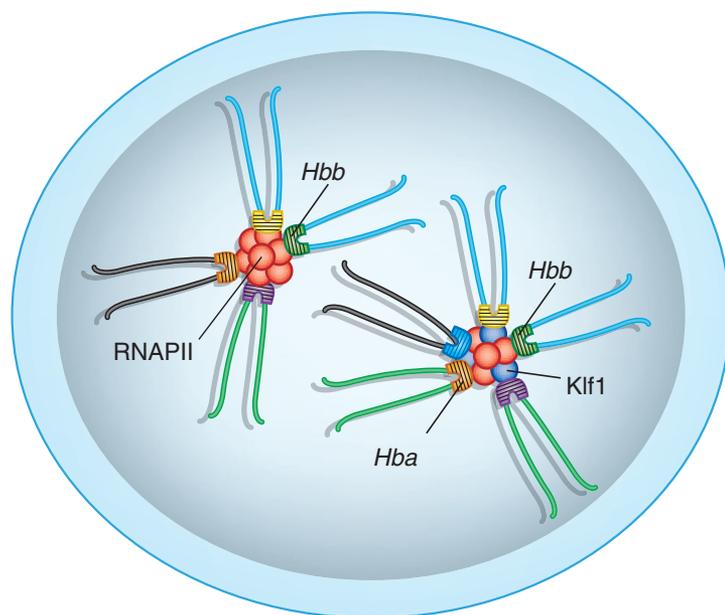


Figure 1 Gene associations at transcription factories in the erythroid nucleus. The *Hbb* locus (like *Hba*) associates with other active genes both *in cis* and *in trans* at transcription factories (left). Specialized transcription factories containing the transactivator Klf1 mediate the preferential interaction between Klf1-regulated genes to boost their expression (right). *Hba* and *Hbb* may interact with members of their own network and/or each other at the same factory. RNAPII, red; Klf1 protein, blue; genes, striped regions. Chromatin loops from the same chromosome associating *in cis* at a factory are depicted in the same color (for example, blue, green).

that between 80% and 90% of e4C hits overlapped with at least one annotated gene, though a correlation with actual expression data might have provided further confirmation that interacting genes were indeed transcriptionally active. Multicolor RNA FISH also revealed that up to three different RNA species (limited to three only by technical constraints) could colocalize at a common transcription factory, strongly suggesting that multiple genes can be transcribed at a single factory (Fig. 1). Importantly, the data reveal that the *Hba* and *Hbb* networks are gene specific, as close to 80% of associated genes were unique to either *Hba* or *Hbb* and only a minority was common to both.

Searching for common motifs among *Hba*- and *Hbb*-interacting genes, the authors discovered that many share binding sites for the erythroid-specific transactivator Klf1. As it turns out, in erythroid nuclei, Klf1 is present in relatively few foci that overlap with a subset of transcription factories. It is these Klf1-containing RNAPII foci that associate

prominently with Klf1-regulated genes and are thus best described as specialized transcription factories (Fig. 1). Supporting evidence comes from the analysis of Klf1 knockout cells, in which associations between *Hbb* and other Klf1-regulated genes, as well as their colocalization with transcription factories, are disrupted. Non-Klf1 target associations, on the other hand, remained largely unaffected.

Clearly, the demonstrated interactions of *Hbb* and *Hba* with other transcribed genes do not occur simultaneously in each erythroid nucleus. Instead, given the dynamics of nuclear organization and the stochastic nature of gene transcription, at any given time a smaller and varied subset of interactions is to be expected. Still, the study reveals a genomic view of preferred active interaction partners of two network hubs (*Hba* and *Hbb*), as well as the existence of long-suspected, but until now not demonstrated, specialized transcription factories that boost the expression of erythroid-specific genes.

Anchors away

Since the description of the 3C method by Dekker *et al.*⁸, variations of this technique have been used to probe the local conformation of complex gene loci as well as genomic regions of ever-increasing size^{9–12}. Almost all of these studies, including the work by Schoenfelder *et al.*⁷, have been anchored to a particular locus with specific bait primers, which increases the specificity of the assay but also biases it. Interactions between other loci of the network (such as that between *Uros* and *Eraf* here) are not registered. Another recent study describes the interactome of estrogen receptor-bound chromatin in a human epithelial adenocarcinoma cell line by ChIP and PET without the use of an anchor⁶. The results revealed an unbiased map of estrogen receptor binding sites and their interactions. However, the interactome failed to be truly genomic, as interchromosomal interactions between binding sites were not reproducibly detected, possibly in part due to the variable and abnormal karyotype of these cells. A promising alternative for probing genomic interactions may be the recently published Hi-C method⁵. As this method is also sequencing based and unbiased, it is capable of reliably uncovering highly detailed chromosomal interactions both *in cis* and *in trans*. It could easily be adapted to include a ChIP step to restrict the analysis to tissue-, pathway- or activity-specific chromatin-gene networks. Regardless, an important remaining question is whether the preferential gene associations at common transcription factories (in particular the specialized type) or other nuclear bodies occur by self-assembly or active recruitment.

1. Rajapakse, I. *et al. Proc. Natl. Acad. Sci. USA* **106**, 6679–6684 (2009).
2. Kosak, S.T. & Groudine, M. *Genes Dev.* **18**, 1371–1384 (2004).
3. Misteli, T. *Bioessays* **27**, 477–487 (2005).
4. Branco, M.R. & Pombo, A. *PLoS Biol.* **4**, e138 (2006).
5. Lieberman-Aiden, E. *et al. Science* **326**, 289–293 (2009).
6. Fullwood, M.J. *et al. Nature* **462**, 58–64 (2009).
7. Schoenfelder, S. *et al. Nat. Genet.* **42**, 53–61 (2010).
8. Dekker, J., Rippe, K., Dekker, M. & Kleckner, N. *Science* **295**, 1306–1311 (2002).
9. Tolhuis, B., Palstra, R.J., Splinter, E., Grosveld, F. & de Laat, W. *Mol. Cell* **10**, 1453–1465 (2002).
10. Skok, J.A. *et al. Nat. Immunol.* **8**, 378–387 (2007).
11. Simonis, M. *et al. Nat. Genet.* **38**, 1348–1354 (2006).
12. Dostie, J. *et al. Genome Res.* **16**, 1299–1309 (2006).