Think globally, act locally

Ommatidia appear to use the morphogenetic furrow for anterior–posterior information (the furrow is assumed to be initiated in the posterior of the disc) and for dorsal–ventral information (the furrow is assumed to be initiated in the dorsal–ventral midline). Both are potentially sensed through short-range cell interactions[9,11,12], providing an excellent example of how local interactions can create a long-range pattern. The important thing seems to be that a firing centre provides an initial asymmetry, from which a pattern would then propagate outwards, by morphogenetic furrow movement (anterior–posterior) and by short-range signalling between ommatidia (dorsal–ventral). It makes sense to use a feature whose polarity is already defined, but the question now becomes: what decides that the morphogenetic furrow is initiated at the posterior midline? It could be that the initiation mechanism (which is largely unknown) responds to a global co-ordinate system in the disc. Alternatively, it could be induced by yet another prior feature in the disc, such as the nearby optic stalk (Fig. 1c).

References

1 Dietrich, W. (1909) *Z. Wiss. Zool.* 92, 465–539
2 Lawrence, P.A. (1992) *The Making of a Fly*, Blackwell Scientific
3 Heberlein, U., Wolff, T. and Rubin, G.M. (1993) *Cell* 75, 913–926
4 Ma, C., Zhou, Y., Beachy, P.A. and Moses, K. (1993) *Cell* 75, 927–938
5 Heberlein, U., Singh, C.M., Luk, A.Y. and Donohue, T.J. (1995) *Nature* 373, 709–711
6 Ma, C. and Moses, K. (1995) *Development* 121, 2279–2289
7 Pan, D. and Rubin, G.M. (1995) *Cell* 80, 543–552
8 Strutt, D.I., Wiersdorff, V. and Mlodzik, M. (1995) *Nature* 373, 705–709
9 Wehrli, M. and Tomlinson, A. (1995) *Development* 121, 2451–2459
10 Chanut, F. and Heberlein, U. (1995) *Development* 121, 4085–4094
11 Strutt, D.I. and Mlodzik, M. (1995) *Development* 121, 4247–4256
12 Zheng, L., Zhang, J. and Carthew, R.W. (1995) *Development* 121, 3045–3055
13 Gubb, D. (1993) *Development* (Suppl.), 269–277
14 Wolff, T. and Ready, D.F. (1991) *Development* 113, 841–850

Position effects and genetic disease

ERIC MILOT, PETER FRASER AND FRANK GROSVELD*
*vanderkamp@ch1.fgg.eur.nl

THE ERASMUS UNIVERSITY, MGC-DEPARTMENT OF CELL BIOLOGY AND GENETICS, PO BOX 1738,
3000 DR ROTTERDAM, THE NETHERLANDS.

The regulation of gene transcription is achieved by different levels of control. Most of the work in recent years has concentrated on the characterization of factors that act on proximal and distal DNA elements, which are responsible for the specific expression of a gene. Much less attention has been paid to the chromatin aspects of gene control. Recently, these two fields are rapidly starting to merge.

Position effects

The eukaryotic genome can be roughly divided into two cytologically distinguishable states: euchromatin and heterochromatin[1,2]. The euchromatin regions are decondensed in interphase. They replicate early and contain mostly single-copy sequences and genes. In contrast, the heterochromatin regions remain condensed throughout the cell cycle, usually replicate late and contain a high proportion of middle-repetitive and highly repetitive sequences[3]. Exactly how the state of chromatin modulates gene expression is still largely unknown, although it is very clear that the chromatin is not simply a passive structural scaffold[4].

Position effects (PEs), that is, a change in the expected level of gene expression, have been associated with integration or translocation of a gene with other regions of the genome. In some cases, heterochromatic and euchromatic regions can be juxtaposed, and this can cause a change in expression of the genes located in the vicinity of the breakpoint[5].

One particular type of PE, first observed in *Drosophila* and yeast, occurs through relocalization of a gene into a heterochromatic environment, which leads to a shut-down of expression of the gene in some of the cells. This cell-to-cell mosaic expression could be created by a differential spreading of the heterochromatin. This phenomena, which is clonal and heritable in daughter cells, is known as position-effect variegation (PEV; Ref. 6) and has yet to be explained fully at the molecular level.

This effect can be altered by the products of a number of genes known as enhancers or suppressors of PEV, which are proteins that are thought to change chromatin packaging. The spreading of such factors into neighbouring areas that contain the relocalized gene is thought to modulate the ability of transcription factors to bind their target DNA sequences. This is strongly supported by studies on telomeric silencing in yeast. Gene suppression is reduced by mutations occurring in *SIR2*, *SIR3*, *SIR4*, *NAT1*, *ARD1* and *RAP1*, which encode factors that bind to telomeric regions[7]. Other spatial effects on gene expression have been observed and correlated to the presence of these proteins. For example, *RAP1* is known to be present at high concentration in spots near the nuclear envelope[8] and is proposed to be involved in the repression of genes in these areas of the nucleus. Suppression by reducing accessibility of the *trans*-acting factor(s) can be monitored by the inability to methylate telomeric sequences[9]
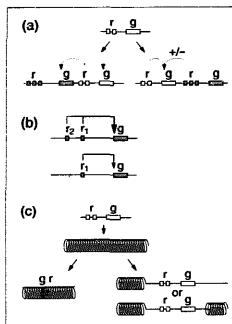
FIGURE 1. Different types of position effects. (a) An incoming (open boxes) gene (g) and distal regulatory elements (r) integrate next to resident gene or regulatory elements (grey boxes). Dotted line represents decreased frequency of interactions due to competition from the resident gene. '+/–' indicates that interactions between resident regulatory elements and the incoming gene could have positive or negative effects on transcription. (b) A resident gene with multiple, distal regulatory elements undergoes a deletion or translocation resulting in the juxtaposition of two originally non-adjacent chromatin regions. Differences in expression could be explained by the loss of a regulatory element. (c) An incoming gene with distal regulatory elements integrates into a region of heterochromatin. This could result in shutdown of the gene (black) through heterochromatinization or maintenance of an open chromatin structure of the gene through interaction with dominant, positive regulatory elements.

and can be counteracted by an increased concentration of a transcription factor whose target sequence is in the telomeric region[10]. Hence, suppression versus activation appears to be regulated by an equilibrium between (specific) positively acting factors and general negatively acting chromatin-binding proteins.

A completely different type of position effect can be envisaged when a gene and its regulatory regions end up next to another gene as a result of a genomic rearrangement (Fig. 1a). If the regulatory regions can interact with the other gene, then this will result in a competition between the genes, so that the original gene would

be expressed at a lower level. Alternatively, the gene might end up next to a novel regulatory region that could up- or downregulate its expression (Fig. 1a). Either way, such an effect could be described as a regulator-position effect (RPE), and would be seen as a change in the level of expression in a population of cells and in each single cell. However, when measured at the gene level in single cells[11], it might show a change in the frequency of transcription (see below).

A different type of position effect can be envisaged when a deletion or rearrangement (and these could be a large distance from the gene) removes part or all of the regulatory regions of a gene (Fig. 1b). This could be called a deletion-position effect (DPE), and would result in a change in expression in each cell.

Chromatin-mediated silencing in competition with activation as observed in PEV (Fig. 1c and see above) can often be distinguished from RPE and DPE, because this effect appears at the level of a single cell in an all-or-none phenomenon, that is, the relevant gene is on or off. When viewed from a population of cells it is measured as an overall decrease, or even complete silencing, of expression. These three types of PE could be fundamentally different or be variations on a theme.

Natural PE in mammals

It is known that RPE exists in mammals and that this can cause disease. One of the best examples is the coupling of the regulatory regions of the immunoglobulin genes to the *MYC* oncogene in Burkitt's lymphoma. This novel combination of gene and regulator results in inappropriate *MYC* expression[12].

As would be expected, DPE also exists in mammals. An example of DPE is that the deletion of the locus control region of the β-globin locus leads to inactivation of the β-globin genes up to 50 kb downstream. It is not obvious whether PEV plays a role in disease. However, a number of recent reports have changed this picture.

The first example is the study of a chromosomal disorder associated with campomelic dysplasia (CD), and autosomal sex reversal. The alteration of the chromosomal locus *SRA1* (17q25.1–q24.1) can lead to failure in testis development and sex reversal,

and to CD, a disease characterized by bone malformations[13]. This locus contains the *SOX9* gene, which encodes a member of the SOX family of DNA-binding proteins. The *SOX* family encodes proteins that contain an AMG box and includes the sex determining gene *SRY*. Foster *et al.*[14] and Wagner *et al.*[15] have recently described a chromosomal rearrangement in the *SOX9* locus of a patient having CD and sex reversal. The breakpoint of this rearrangement is 88 kb from *SOX9*. Examining this gene in other CD patients, they found that internal mutations in the *SOX9* coding region also lead to CD. There is no other gene known to be important for bone formation and for testis development in this locus. How, then, can a rearrangement 88 kb from *SOX9* lead to CD? As the authors pointed out, it could be that some unidentified parts of *SOX9* are disrupted by the rearrangement. Another possibility is that the rearrangement led to silencing of *SOX9* because of a PE. This could be created by the new environment of *SOX9* in this rearrangement (PEV or RPE), but also by the deletion of important regulatory regions (DPE).

A similar observation was made for inactivation of the *PAX6* gene, which, in the heterozygous state, leads to aniridia (absence of the iris)[16]. The analysis of two aniridia families revealed genomic rearrangements in the locus 11p13 containing the *PAX6* gene[17]. Precise mapping of these genomic alterations showed that the chromosomal breakpoints are located at least 85 kb distal from the 3' end of *PAX6*. Again, these results could be explained as an extreme PEV, in which the new chromatin environment leads to complete silencing of *PAX6* in the chromosomal rearrangements. Alternatively, the results could be explained simply as a loss of an unidentified, distant regulatory region. Another recent example of a disease with a possible PE concerns the *POU3F4* gene, which is involved in X-linked deafness[18,19].

The last example was obtained in a study of the *Steel* (*Sl*) locus in mice. The *Sl* locus is essential for the development of germ cells, haematopoietic cells and melanocytes[20]. The gene encodes the mast-cell growth factor, MGF, which is the ligand of the tyrosine kinase receptor encoded by *kit*. The molecular defects in two *Sl* mutant

alleles that lead to sterility in females were shown to be chromosomal rearrangements[21]. As in the cases for *SOX9* and *PAX6* described above, the rearrangements were mapped precisely with breakpoints at 115 and 195 kb from the *Mgf* coding region. In both rearrangements, *Mgf* transcripts seem to be normal but the level of expression is altered by the rearrangements. Again, this suggests that the new environment of the *Mgf* coding region could affect the level of expression, but it could equally well be due to a loss of regulatory regions.

These three reports show that PE in mammals can play a role in genetic diseases. Although it is attractive to suggest that PEV is the cause, it is actually not clear because RPE or DPE could also be the cause of the phenotypes described above. In fact, all PEs could be interrelated and could be the different outcomes of affecting the balance between activation and suppression. The data supporting this idea have been obtained from transgenic mouse experiments.

Transgenic PE in mammals

PEs have been frequently observed in transformation systems and transgenic animals[5]. These pose significant problems in the study of gene regulation. However, studies on the human β-globin gene locus led to the discovery of the locus control region (LCR) which overcomes PE (Ref. 22). Since then, a number of other LCRs have been identified[23]. Similar to most regulatory elements, LCRs are tissue specific. However, they differ from common regulatory elements because they enable expression independent of the position of integration in the host genome. In addition, they provide a level of expression dependent on the number of copies of the LCR-transgene construct. The β-globin LCR is one contiguous piece of DNA of 21 kb, which contains five tissue-specific hypersensitive sites[23,24]. In other cases, different numbers of hypersensitive sites have been found and they need not be immediately adjacent to each other. The hypersensitive regions bind to transcription factors and the suggestion is that this binding is responsible for the open chromatin configuration. This would partly explain why euchromatic, but not heterochromatic, regions of the genome are sensitive to DNaseI.

Another explanation for the prevention of PE would be provided by domain boundaries. This comes from the notion, originally based on cytological data[25,26], that genes or sets of genes are present in distinct domains separated by boundaries[27]. Two such candidate domain boundaries, scs and scs', were identified in *Drosophila*[28]. These elements are characterized by a nuclease-resistant sequence (250–350 bp) flanked by a pair of nuclease-hypersensitive sites. The nuclease-hypersensitive sites can insulate a reporter gene from PE in *Drosophila*[29]. The hypersensitive site (5'HS4) from the chicken β-globin locus was also reported to act as an insulator when tested in *Drosophila*[30]. In contrast, we failed to find such activity in transgenic mice using the *Drosophila* scs elements or the human counterpart of chicken 5'HS4 (i.e. human HS5; Ref. 24).

There have been some reports of PE in transgenic mice when LCR sequences are used in the construct[31–34]. In one transgenic line reported by Strouboulis et al.[32], all of the human transgenes in the locus were expressed at a lower level. This was probably caused by integration next to an endogenous gene that competes with the globin genes for the interaction with the LCR (i.e. a case of RPE). When predictable levels of expression are required (e.g. in gene therapy), it would, therefore, be important to keep the distance between the promoter and the regulators as small as possible and, thus, lower the risk of RPE.

One study using LCRs shows PEV (Ref. 33), and two other studies discuss PE in general[31,34]. In these three cases, the LCRs used were incomplete or modified. The PEV observed by Elliott et al.[33] occurred when the CD2 LCR is used in combination with the immunoglobulin enhancer, but not when the CD2 LCR is used alone[35]. Bonifer et al.[31] and Robertson et al.[34] used an LCR containing one or more hypersensitive sites[36,37]. Although no firm conclusions can be made from these results, they suggest that interference with an LCR, in the form of deletions or additions, make it sensitive to PEs.

Hence, deletions of the CD2 and the β-globin LCR were made and tested in transgenic mice[38] (E. Milot et al., unpublished). In the case of the CD2 LCR, the deletion of the most 3'

hypersensitive site (HS3 downstream of the gene) yielded a number of mice with a classical, clonally heritable PEV. In the case of the β-globin LCR, the deletion of different hypersensitive sites was tested, and PE and PEV were both found (E. Milot et al., unpublished).

The most interesting observation of both sets of experiments was the perfect correlation between the observed effects and integration of the deletion constructs into heterochromatic regions of the host genome. Integration into euchromatin had no such effect. A possible explanation for these results, and those described above, comes from work describing the dynamics of globin LCR–gene interactions *in vivo*[1]. The major implication from that work is that LCR–gene interactions are not static, but that complexes form and dissociate continually. These stochastic interactions are dependent on three parameters: (1) the frequency of interaction, which is, itself, dependent on the distance of the gene to the LCR (Refs 39–41); (2) the affinity of the LCR for the gene; and (3) the stability of the LCR–gene complex. The latter two are dependent on the balance of DNA-binding proteins in the nucleus. The net conclusion is that the LCR determines, at least partly, the level of gene expression by determining the frequency of expression. In other words, the LCR would not increase the amount of polymerases that can be loaded on a gene at a given time but would, rather, increase the frequency of periods during which polymerases can be loaded on a gene[26]. It is easy to explain DPE and RPE in terms of the creation of novel combinations of regulatory elements and genes. It would appear to be more difficult to explain PEV and clonal inheritance, were it not for the fact that the cells in an apparently homogenous population are actually not transcriptionally identical at any given time[11]. Hence, a processive shut-down mechanism spreading along the chromosome, as in heterochromatinization, might be in competition with a stochastic activation event. The net result of these two processes might not be the same in every cell and could result in a stable shut-down in some of the cells, which is inherited by an unknown mechanism through replication and division. The same percentage of expressing versus silent

cells will be found in animals in the next generation because all the parameters underlying the stochastic events will be similar in the offspring to those in the parents.

In conclusion, studies of genetic diseases and transgenic models indicate that PE phenomena are important for mammalian gene regulation. Thus, genomic rearrangements should not be regarded as a 'local' reorganization of the genome only, but should also be regarded as events that could alter long-range chromatin interactions. At the moment, it is not clear whether DPE, RPE or PEV is the cause of the human diseases described above and, hence, whether PEV plays a role in mammalian disease at all. However, transgenic studies suggest that these different types of PE can be observed in mammals and that chromatin modulators, such as LCRs and domain boundaries, can be used to clarify the chromatin organization. Further molecular characterization of the diseases mentioned above, and others, should also help to understand the complexity of the organization of the mammalian genome. Some diseases are particularly good candidates because of the chromosomal position of the gene(s) involved, such as the gene responsible for faciocapulohumeral muscular dystrophy, which is located close to the telomere of chromosome 4 (Ref. 42).

References

1 Heintz, E. (1928) *Int. Jahrb. Wiss. Bot.* 69, 762–818
2 Eissenberg, J.C., Elgin, S.C.R. and Paro, R. (1995) in *Chromatin Structure and Gene Expression* (Elgin, S.C.R., ed.), pp. 147–171, Oxford University Press
3 John, B. (1988) in *Heterochromatin: Molecular and Structural Aspects* (Verma, R.S., ed.), pp. 141–147, Cambridge University Press
4 Pillus, L. and Grunstein, M. (1995) in *Chromatin Structure and Gene Expression* (Elgin, S.C.R., ed.), pp. 123–146, Oxford University Press
5 Wilson, C., Hugo, G., Bellen, H. and Gehring, W. (1990) *Annu. Rev. Cell. Biol.* 6, 679–714
6 Karpen, G.H. (1994) *Curr. Opin. Genet. Dev.* 4, 281–291
7 Palladino, F. and Gasser, S.M. (1994) *Curr. Opin. Cell Biol.* 6, 373–379
8 Palladino, F. et al. (1993) *Cell* 75, 531–542
9 Gottschling, D.E. (1992) *Proc. Natl Acad. Sci. USA* 89, 4062–4065
10 Aparicio, O.M. and Gottschling, D.E. (1994) *Genes Dev.* 8, 1133–1146
11 Wijgerde, M., Grosveld, F. and Fraser, P. (1995) *Nature* 377, 209–213
12 Joos, S. et al. (1992) *Hum. Mol. Genet.* 1, 625–632
13 Tommerup, N. et al. (1993) *Nat. Genet.* 4, 170–174
14 Foster, J.W. et al. (1994) *Nature* 372, 525–530
15 Wagner, T. et al. (1994) *Cell* 79, 1111–1120
16 Ton, C.C.T. et al. (1991) *Cell* 67, 1059–1074
17 Fantes, J. et al. (1995) *Hum. Mol. Genet.* 4, 415–422
18 de Kok, Y.J.M. et al. (1995) *Science* 267, 685–688
19 de Kok, Y.J.M. et al. (1995) *Hum. Mol. Genet.* 4, 2145–2150
20 Silvers, W.K. (1979) *The Coat Colors of Mice: A Model for Mammalian Gene Action and Interaction*, Springer-Verlag
21 Bedell, M.A. et al. (1995) *Genes Dev.* 9, 455–470
22 Grosveld, F., van Assendeift, G.B., Greaves, D.R. and Kollias, G. (1987) *Cell* 51, 975–985
23 Dillon, N. and Grosveld, F. (1993) *Trends Genet.* 9, 134–137
24 Zafarana, G. et al. (1995) in *Volume 1 of the Proceedings of the 9th Conference on Hemoglobin Switching* (Stamatoyannopoulos, G., ed.), pp. 39–44
25 Gall, J. (1956) in *Mutation*, Brook-Haven Symp. Biol. 8, 17–26
26 Alfert, M. (1956) *Int. Rev. Cytol.* 3, 131–140
27 Schedl, P. and Grosveld, F. (1995) in *Chromatin Structure and Gene Expression* (Elgin, S.C.R., ed.), pp. 174–196, Oxford University Press
28 Udvardy, A., Maine, E. and Schedl, P. (1985) *J. Mol. Biol.* 185, 341–358
29 Kellum, R. and Schedl, P. (1991) *Cell* 64, 941–950
30 Chung, J.H., Whiteley, M. and Felsenfeld, G. (1993) *Cell* 74, 505–514
31 Bonifer, C. et al. (1994) *Nucleic Acids Res.* 22, 4202–4210
32 Strouboulis, J., Dillon, N. and Grosveld, F. (1992) *Genes Dev.* 6, 1857–1864
33 Elliott, J.I., Festenstein, R., Tolaini, M. and Kioussis, D. (1995) *EMBO J.* 14, 575–584
34 Robertson, G. et al. (1995) *Proc. Natl Acad. Sci. USA* 92, 5371–5375
35 Lang, G. et al. (1988) *EMBO J.* 7, 1675–1682
36 Philipsen, S. et al. (1990) *EMBO J.* 9, 2159–2167
37 Pruzir a, S. et al. (1991) *Nucleic Acids Res.* 9, 1413–1419
38 Festenstein, R. et al. *Science* (in press)
39 Hanscombe, O. et al. (1991) *Genes Dev.* 5, 1387–1394
40 Pederson, K.R. and Stamatoyannopoulos, G. (1993) *Mol. Cell. Biol.* 13, 4836–4843
41 Dillon, N., Strouboulis, J. and Grosveld, F. in *Molecular Biology of Hemoglobin Switching: Proceedings of the 9th Conference on Hemoglobin Switching* (Stamatoyannopoulos, G., ed.), (in press)
42 Zatz, M. et al. (1995) *Am. J. Hum. Genet.* 56, 99–105