MOLECULAR AND CELLULAR BIOLOGY, Apr. 1997, p. 2326–2335 Vol. 17, No. 4
0270-7306/97/$04.00+0

Identification and Characterization of Interactions between the Vertebrate Polycomb-Group Protein BMI1 and Human Homologs of Polyhomeotic

MARCO J. GUNSTER,[1] DAVID P. E. SATIJN,[1] KARIEN M. HAMER,[1] JAN L. DEN BLAAUWEN,[1] DIEDERIK DE BRUIJN,[2] MARK J. ALKEMA,[3] MAARTEN VAN LOHUIZEN,[3] ROEL VAN DRIEL,[1] AND ARIE P. OTTE[1]*

E. C. Slater Institute, University of Amsterdam, 1018 TV Amsterdam,[1] Department of Human Genetics, University Hospital, 6500 HB Nijmegen,[2] and Division of Molecular Genetics, Netherlands Cancer Institute, 1066 CX Amsterdam,[3] The Netherlands

Received 31 October 1996/Returned for modification 2 December 1996/Accepted 21 January 1997

In *Drosophila melanogaster*, the *Polycomb*-group (*PcG*) genes have been identified as repressors of gene expression. They are part of a cellular memory system that is responsible for the stable transmission of gene activity to progeny cells. PcG proteins form a large multimeric, chromatin-associated protein complex, but the identity of its components is largely unknown. Here, we identify two human proteins, HPH1 and HPH2, that are associated with the vertebrate PcG protein BMI1. HPH1 and HPH2 coimmunoprecipitate and cofractionate with each other and with BMI1. They also colocalize with BMI1 in interphase nuclei of U-2 OS human osteosarcoma and SW480 human colorectal adenocarcinoma cells. HPH1 and HPH2 have little sequence homology with each other, except in two highly conserved domains, designated homology domains I and II. They share these homology domains I and II with the *Drosophila* PcG protein Polyhomeotic (Ph), and we, therefore, have named the novel proteins HPH1 and HPH2. *HPH1*, *HPH2*, and *BMI1* show distinct, although overlapping expression patterns in different tissues and cell lines. Two-hybrid analysis shows that homology domain II of HPH1 interacts with both homology domains I and II of HPH2. In contrast, homology domain I of HPH1 interacts only with homology domain II of HPH2, but not with homology domain I of HPH2. Furthermore, BMI1 does not interact with the individual homology domains. Instead, both intact homology domains I and II need to be present for interactions with BMI1. These data demonstrate the involvement of homology domains I and II in protein-protein interactions and indicate that HPH1 and HPH2 are able to heterodimerize.

In *Drosophila melanogaster*, the genes of the *Polycomb* group (*PcG*) and *trithorax* group (*trxG*) are part of a cellular memory system that is responsible for the stable inheritance of gene activity. The *PcG* and *trxG* genes have been identified in *Drosophila* as repressors (*PcG*) (17, 22, 30) and activators (*trxG*) (16) of homeotic gene activity. In *PcG* mutants, the expression patterns of homeotic genes are initially normal, but in later embryonic phases, homeotic genes become expressed in regions of the embryo where they normally are repressed (22, 29). Besides *Polycomb* (*Pc*), an estimated 30 to 40 mutants exhibit similar, characteristic posterior homeotic transformations. These are collectively referred to as *Polycomb* group (*PcG*) mutants (15). The Pc protein binds to about 100 loci on polytene chromosomes in the *Drosophila* salivary gland (37). Also the PcG proteins Polyhomeotic (Ph), Polycomblike (Pcl) and Posterior sex combs (Psc) share many, but not all of these binding sites with Pc (10, 13, 18, 25). This is consistent with the idea that PcG proteins work together in a multimeric protein complex. Direct evidence that Pc is part of a multimeric complex comes from experiments which showed that the PcG protein Ph coimmunoprecipitates with Pc (13).

An important clue about the molecular mechanism underlying PcG action is the observation that the Pc protein shares a homologous domain with the *Drosophila* heterochromatin-binding protein HP1 (23). This discovery provides an important, direct link between regulation of gene activity and chromatin structure. It suggests that Pc and HP1 operate through common mechanisms, which may involve the formation of heterochromatin-like structures. This is further suggested by the finding that Pc is associated with chromatin over stretches of many kilobases. Pc specifically covers the silent part of the homeotic *bithorax* complex, which indicates that Pc represses gene activity by heterochromatinization (21).

PcG genes have originally been identified in *Drosophila*, but mammalian homologs of *PcG* genes have been identified recently. A mouse *Pc* homolog, M33 (24), and a *Xenopus Pc* homolog, *XPc* (26), have been characterized. Another *PcG* gene, *Posterior sex combs (Psc)*, possesses significant sequence homology to the murine oncogene *Bmi1* (6, 34, 35). Targeted deletion of *Bmi1* results in a posterior transformation of the embryo and an anterior shift in the expression of homeotic genes (32, 33). Vice versa, overexpression of *Bmi1* results in an anterior transformation and a posterior shift in the expression pattern of the homeotic gene *Hoxc-5* (3). Also, targeted deletion of *mel-18*, a closely related *Bmi1* homolog, results in a posterior transformation of the embryo and an anterior shift in the expression of homeotic genes (1). Collectively, these data show that, also in vertebrates, PcG proteins are involved in the regulation of homeotic genes.

Considering the many novel features of regulation of gene activity by changes in chromatin structure, amazingly little is known about the molecular nature of chromatin structure. In

* Corresponding author. Mailing address: E. C. Slater Institute, University of Amsterdam, Plantage Muidergracht 12, 1018 TV Amsterdam, The Netherlands. Phone: 31-20-5255115. Fax: 31-20-5255124. E-mail: arie.otte@chem.uva.nl.

particular, the composition of the chromatin-associated PcG complex is largely unknown. In order to characterize the molecular nature of the verbrate PcG protein complex, we employed the genetic two-hybrid system (7, 11, 12, 14). We screened a human leukocyte cDNA library for proteins that interact with the vertebrate PcG protein Bmi1. We report the identification of two novel human proteins that are associated with BMI1. The proteins have little sequence homology with each other, except in two conserved homology domains, I and II, which they share with the *Drosophila* PcG protein Ph. Based on the sequence homologies with Ph, we named these proteins HPH1 and HPH2. Two-hybrid analysis indicates that the homology domains are involved in protein-protein interactions and that HPH1 and HPH2 are able to heterodimerize through the homology domains.

MATERIALS AND METHODS

Isolation of the *HPH1* and *HPH2* genes. The full-length coding region of *XBmi1* (26) was cloned into the pAS2 vector (11, 14) (Clontech, Palo Alto, Calif.) and used as "bait" to screen for interacting proteins in a two-hybrid screen (12). The pAS2-*XBmi1* plasmid was cotransformed with a human leukocyte Matchmaker two-hybrid library (Clontech) into the yeast Y190 strain. The transformants were plated on selective medium lacking the leucine, tryptophan, and histidine amino acids, but containing 30 mM 3-aminotriazole (11, 14). Approximately 1.6×10^6 independent clones were obtained. Three hundred growing colonies were obtained, of which 16 were β-galactosidase positive. After DNA isolation and rescreening, six colonies remained histidine and β-galactosidase positive. These clones were further characterized. Two of these clones were 2.1 kb in length and were named *HPH2* based on its homologous domains with *Drosophila Ph* and *Rae28* (20). A total of 1.2 kb of this clone was the 3′ noncoding region. The coding region of *HPH2* as well as the first 5′ 1,000 bp of the coding region of *Rae28* were used to screen a λgt10 human fetal brain cDNA library (Clontech). Filters were hybridized overnight at 50°C in 0.5× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate), 10× Denhardt's solution, 10% dextran sulfate, 0.1% sodium dodecyl sulfate, (SDS), 100 μg of denatured herring sperm DNA per ml, and an [α-^{32}P]ATP-labeled probe (5×10^5 cpm/ml). After being washed two times for 60 min at 55°C in 0.5× SSC and 0.1% SDS, the filters were autoradiographed with intensifying screens for 2 days at −70°C. This led to the isolation of the *HPH1* cDNA.

RNA analysis. Multitissue Northern blots containing approximately 2 μg of poly(A)$^+$ RNA per lane from different human tissues or human cell lines were obtained commercially (Clontech). The U-2 OS cell line was not present on the commercial Northern blot. Poly(A)$^+$ RNA of U-2 OS was isolated and blotted, and the expression patterns of *HPH1*, *HPH2*, and *BMI1* were analyzed. To allow a comparison with the commercial Northern blot, we blotted poly(A)$^+$ RNA of SW480 cells, which is represented on the commercial blot, and all three genes are strongly expressed in this cell line. The blots were hybridized with [α-^{32}P]dATP-labelled DNA probes and the blots were autoradiographed with intensifying screens at −70°C with X-ray films.

Production of the Rae28 (Mph1) and HPH2 polyclonal rabbit antibodies. A fragment encompassing amino acids (aa) 413 to 760 of Rae28 (Mph1) was cloned into the expression vector pGEX2T (Pharmacia, Uppsala, Sweden) to produce a glutathione *S*-transferase–Mph1 fusion protein (4). This fragment encompasses a portion of the protein that is virtually identical to the corresponding region of HPH1, but which is absent from the HPH2 protein. The purified fusion protein was used to immunize a rabbit, and the serum was affinity purified with Affigel (Bio-Rad). A fragment encompassing aa 137 to 370 of HPH2 was cloned into the expression vector pET-23c (Novagen, Madison, Wis.). This fragment encompasses homology domain I, but lacks homology domain II. In homology domain I, this region shares homology with homology domain I of HPH1, but the remaining portion of the protein has no homology with HPH1. Purified fusion proteins were injected into a rabbit. The serum was affinity purified with CNBr Sepharose (Pharmacia, Uppsala, Sweden) to which the antigen was coupled. The production of the monoclonal antibody specific for BMI1 will be described elsewhere (4).

Immunoprecipitation and Western blotting. One 175-cm^2 flask with human U-2 OS osteosarcoma cells which were grown to confluence was lysed in a ELB lysis buffer (250 mM NaCl, 0.1% Nonidet P-40, 50 mM HEPES [*N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid; pH 7.0], 5 mM EDTA, containing 0.5 mM dithiothreitol, 1 mM phenylmethylsulfonyl fluoride, and protease inhibitors, including leupeptin and aprotinin). The cell lysate was sonicated three times with bursts of 15 s. The lysate was centrifuged at 14,000 × *g* at 4°C for 10 min, and the supernatant (500 μl) was aliquoted and stored at −70°C. Fifty microliters of the supernatant was subsequently incubated with the indicated antibodies, for 4 h to overnight at 4°C. Goat anti-rabbit or goat anti-mouse immunoglobulin G (IgG) antibodies coupled to agarose beads (Sigma) and ELB buffer were added to enlarge the volume of the mixture to 300 μl. The mixture was incubated for 2 h at 4°C, under continuous mixing. The mixture was centrifuged at 1,500 × *g* at 4°C for 1 min, washed with 1 ml of ice-cold ELB buffer without protease inhibitors, and centrifuged at 1,500 × *g* at 4°C for 1 min. This washing procedure was repeated five times. After heating and centrifugation to remove the agarose beads, the proteins were separated by SDS-polyacrylamide gel electrophoresis (PAGE) and transferred to nitrocellulose. Typically, 10 μl of the SW480 cell extract was loaded as input, and 50 μl of the SW480 cell extract was used for the immunoprecipitation. The blots were probed with a 1:5,000 dilution of affinity-purified anti-Rae28/Mph1 or HPH2 antibodies or 1:3 dilution of tissue culture serum of the anti-Bmi1 monoclonal antibody. The secondary alkaline phosphatase-conjugated goat anti-rabbit or mouse IgG (H+L) antibodies (Jackson Immuno Research Laboratories) were diluted 1:10,000 and nitroblue tetrazolium–5-bromo-4-chloro-3-indolylphosphate toluidinium (Boehringer) was used as substrate.

TABLE 1. β-Galactosidase activities of Xbmi1, HPH1, and HPH2 interactions in the two-hybrid system

DNA-binding domain fusion (aa)	Activation domain fusion (aa)	Colony color[a]	% Relative β-galactosidase activity
HPH1 (722–1013)	HPH2 (137–432)	Blue	80
HPH1 (722–780)	HPH2 (137–432)	Blue	20
HPH1 (722–780)	HPH2 (104–190)	White	<1
HPH1 (722–780)	HPH2 (328–432)	Blue	20
HPH1 (925–1013)	HPH2 (137–432)	Blue	55
HPH1 (925–1013)	HPH2 (104–190)	Blue	30
HPH1 (925–1013)	HPH2 (328–432)	Blue	30
Xbmi1 (1–328)	HPH2 (137–432)	Blue	100[b]
Xbmi1 (1–328)	HPH2 (104–190)	White	<1
Xbmi1 (1–328)	HPH2 (328–432)	White	<1
Xbmi1 (1–328)	HPH2 (137–375)	White	<1

[a] White colonies were obtained on medium lacking both histidine and 3-aminotriazole. Blue colonies were obtained on medium lacking histidine, but in the presence of 3-aminotriazole.

[b] The average β-galactosidase activity in a triplicate experiment was 49 U. This activity was set at 100%.

Sucrose gradient ultracentrifugation. One hundred microliters of the U-2 OS cell extract in ELB buffer was loaded on top of a 5 to 30% sucrose gradient (5 ml) in ELB buffer. The gradient was centrifuged in a TST 54 (Kontron) rotor at 50,000 rpm for 8 h at 4°C. The gradient was prepared, and fractions were collected with a Buchler Auto Densi-Flow IIC. Fourteen 350-μl Fractions were collected, and proteins were precipitated by addition of 700 μl of ice-cold ethanol. The proteins were dissolved in SDS-PAGE loading buffer. The samples were split, and the proteins were separated on two SDS-PAGE gels and transferred to nitrocellulose. The upper halves of both blots were probed with anti-Rae28/Mph1 antibody to detect the approximately 120-kDa HPH1, and the lower halves were probed with, respectively, anti-HPH2 (54 kDa) or anti-BMI1 (44 to 46 kDa) antibody.

Analysis of interacting protein domains with the two-hybrid system. Indicated fragments of the cDNAs encoding HPH1 and HPH2 were derived by PCR (Expand, Boehringer). The fragments were subcloned into the pAS2 (GAL4-DNA binding domain [GAL4-BD]) or GAD10 (GAL4 activation domain [GAL4-AD]) vector. The fragments were sequenced over their entire length. The resulting plasmids were cotransformed into the yeast Y190 strain. The transformants were plated on medium lacking the leucine, tryptophan, and histidine amino acids, with or without 30 mM 3-aminotriazole. Interactions that were scored negative failed to grow in the presence of 30 mM 3-aminotriazole. Because of residual HIS3 promoter activity, they are able, however, to grow on medium that does not contain 3-aminotriazole (11, 14). Under these nonselective conditions, negative interactions were β-galactosidase negative, and the colony color was indicated as white (Table 1). Positive interactions meet the two criteria of growth in the presence of 30 mM 3-aminotriazole, and they are β-galactosidase positive. To quantitate the β-galactosidase activity, cultures of 2.5 ml were grown to an optical density at 600 nm of 1.0 to 1.2 in medium lacking the leucine, tryptophan, and histidine amino acids. No 3-aminotriazole was added in the case of negative interactors to allow them to grow. Cells were permeabilized, and β-galactosidase activity was measured as described previously (11, 14; technical instructions of Clontech). Measurements were performed in triplicate, starting with three independent colonies. To exclude the possibility that the negative interactors did not produce either one of the fusion proteins, we Western blotted equal amounts of protein and incubated the blots with monoclonal antibodies that specifically recognize the GAL4-BD or the GAL4-AD protein (Clontech). All positive and negative interactors expressed both GAL4-BD fusions and the GAL4-AD fusions at approximately the same levels (data not shown).

Immunofluorescence staining of tissue culture cells. U-2 OS and SW480 cells were cultured and labelled as described recently (28). Labelling has been ana-

```
MLRGGCSPRAPAAAPQPRPPPALPPRPRAPVPASRPGRPLLTPARPCGRM      50

RRGSPGPRLGGSRGERRRPAGRDPARVGPGQGLRRPARPGPAAWTETGQG     100
                                   I
IVHALTDLSIPGMTSGNGNSASSIAGTAPQNGENKPPQAIVKPQILTHVI     150
                                            ↑
EGFVIQEGADVSRWDARLLVGNLKKKYAQGFLPEKLPQQDHTTTTDSEME     200

EPYLQESKEEGAPLKLKCELCGRVDFAYKFKRSKRFCSMACAKRYNVGCT     250

KRVGLFHSDRSKLQKAGAATHNRRRPAKPVCHHLPRIPRSSQQALCPFRL     300

LLLCVTHSQEDSSRCSDNSSYEEPLSPISASSSTSAGDKASGTWSSPTCI     350
                    II
CGTWWAWDTTSCQVSHQVNVEDVYEFIRSLPGCQEIAEEFRAQEIDGQAL     400

LLLKEDHLMSVMNIKLGPALKIYARISMLKDS                        432
```

FIG. 1. Predicted amino acid sequence of HPH2. The homology domains I and II are indicated in boldface and underlined. The isolated two-hybrid clone encompasses homology domains I and II, starting from aa 137, which is indicated by an arrow.

lyzed by confocal laser scanning microscopy, of which single optical sections are shown. The first two pictures of each row represent the two different detection channels of the dual image, whereas the third picture in each row represent the false color overlay. For labelling, donkey anti-rabbit IgG coupled to Cy3 (Jackson ImmunoResearch Laboratories) and donkey anti-mouse IgG coupled to DTAF {5-([4,6-dichlorotriazin-2-yl]amino)fluorescein; Jackson} were used.

Nucleotide sequence accession number. The sequences for HPH1 and HPH2 have been deposited with GenBank under accession numbers U89277 and U89278, respectively.

RESULTS

Isolation and characterization of the human *HPH1* and *HPH2* genes. To identify genes encoding proteins that interact with or are part of the vertebrate PcG multimeric protein complex, we performed a genetic two-hybrid screen (7, 11, 12, 14). As target protein, we chose the *Xenopus* homolog of the vertebrate PcG protein Bmi1 (2, 3, 26, 35). At the protein level, XBmi1 is 90% identical and 95% similar to mouse and human BMI1 (26). We screened a human two-hybrid cDNA library. The full-length coding region of *XBmi1* (26) was cloned into the pAS2 vector (11, 14). The pAS2-*XBmi1* plasmid was co-transformed with a human leukocyte Matchmaker two-hybrid library (Clontech) into the yeast Y190 strain. The transformants were plated on selective medium lacking the histidine, leucine, and tryptophan amino acids and in the presence of 30 mM 3-aminotriazol (11, 14). Of approximately 1.6×10^6 independent clones, 300 colonies were His^+, of which 16 were also β-galactosidase positive. After DNA isolation and retransformation, six colonies remained His^+ β-Gal^+. These clones did not grow on His^- plates with 3-aminotriazol when cotransformed with either the empty pAS2 vector or with fusion proteins unrelated to the target protein.

Two of these clones were identical. They were 2.1 kb in length and contained a 900-bp coding region and a 1,200-bp 3′ noncoding region. The predicted protein coincides with two highly homologous domains of the Rae28 protein (Fig. 1 and 2). Rae28 (retinoic acid-activated Early-28) has been identified in a differential screen for genes which respond to retinoic acid (20). Employing a two-hybrid screen, with Bmi1 as the target protein, Rae28 has been identified as interacting with Bmi1 (4). Rae28 shares considerable homology with the *Drosophila* PcG protein Ph, particularly in two domains, designated homology domains I and II, that are, respectively, 29 and 67 amino acids long (9, 10, 20 [compare Fig. 2]). Sequence analysis indicated that the clone we isolated encompasses these two homology domains (Fig. 1 and 2). Homology domain I is 90 and 52% identical with homology domain I of Rae28 and Ph, respectively. Homology domain II is 79 and 56% identical with homology domain II of Rae28 and Ph, respectively (Fig. 2). Also, a Cys2-Cys2 zinc finger motif is conserved between the protein we isolated and the Rae28 and Ph proteins. Not only the structure of this zinc finger motif, but also its location between homology domains I and II is conserved (Fig. 2). Based on these homologies, we named the clone we isolated *HPH2*. Following the same reasoning, Alkema et al. recently renamed *Rae28 Mph1* (4).

The *HPH2* clone starts at homology domain I, suggesting that we did not isolate the full-length clone, but only the 3′ region of the gene. To isolate full-length cDNAs of *HPH2*, we screened a human fetal brain library. Initially we used the 1,300-bp 3′ noncoding region of *HPH2* as probe, but isolated

```
HPH1    METESEQNSNSTNGSSSSGGSSRPQIAQMSLYERQAVQALQALQRQPNAAQYFHQFMLQQQLSNAQLHSLAAVQQATIAASRQASSPNTSTTQQQTTTTQASINLATTSAAQLIS  115
Rae28   METESEQNS-STNGSSSSG-SSRPQIAQMSLYERQAVQALQALQRQPNAAQYFHQFMLQQQLSNAQLHSLAAVQQATIAASRQASSPN-ST-QQQT-TTQAS-NLATTSAAQLIS

HPH1    RSQSVSSPSATTLTQSVLLGNTTSPPLNQSQAQMYLRPQLGNLLQVNRTPGSNVPLASQLILMPNGAVAAVQQEVPSAQSPGVHADADQVQNLAVRNQQASAQGPQMQGSTQKAI  230
Rae28   RSQSVSSPSATTLTQSVLLGNTTSPPLNQSQAQMYLRPQLGNLLQVNRT-G-MVPLASQLILMPNGAVAAVQQEVP-AQSPGVHADADQVQNLAVRNQQASAQGPQM-GSTQKAI

HPH1    PPGASPVSSLSQASSQALAVAQASSGATNQSLNLSQAGGGSGNSIPGSMGPGGGGQAHGGLGQLPSSGMGGGSCPRKGTGVVQPLPAAQTVTVSQGSQTEAESAAAKKAEADGSG  345
Rae28   PPGASPVS-LSQ-SSQALAVAQASSGA--QSLNLSQAGGGSGNS-PGSMGPGGGGQA-GGLGQLPSSG--GGSCPRKGTGVVQPLPAAQTVTVSQGSQTEAESAAAKKAEADGSG

HPH1    QQNVGMNLTRTATPAPSQTLISSATYTQIQPHSLIQQQQQIHLQQKQVVIQQQIAIHHQQQFQHRQSQLLHTATHLQLAQQQQQQQQQQQQQQ=======PQATTLTAPQPPQV  460
Rae28   QQ-VGMNLTRTATPAPSQTLISSATYTQIQPHSLIQQQQQIHLQQKQVVIQQQIA-HHQQQFQHRQSQLLHTATHLQLAQQQQQQQQQQQQQQQQQQQQQQQQ-Q-TTLTAPQPPQV
Ph          G-------T-----TL---T----------QQQQ--L-QKQ---QQQ-----Q------Q-Q--------Q--QQ--QQQQ---QQQQ+++++++A----A----Q-

HPH1    PPTQQVPPSQSQQQAQTLVVQPMLQSSPLSLPPDAAPKPPIPIQSKPPVAPIKPPQLGAAKMSAAQQPPPHIPVQVVGTRQPGTAQAQALGLAQLAAAVPTSRGMPGTVQSGQAL  575
Rae28   PPTQQVPPSQSQQQAQTLVVQPMLQSSPL-LPP----KPPIPIQSKPPVAPIKPPQLGAAKMSA-QQPPPHIPVQVVGTRQPG-AQAQALGLAQLAAAVPT-RG--G-VQ-GQA-
Ph      ----Q-----------------------------------------------------------------------------------------------------------

HPH1    KA-SSPPSSQAPGALQECPPTLAPGMTLAPVQGTAHVVKGGATTSTPVVAQVPAAFYMQSVHLPGKPQTLAVKRKADSEEERDDVSTLGSMLPAKASPVAESPKVMDEKSSLGEK  690
Rae28   -ASS-P-S--APGALQECPP-LA-GMTLAPVQGTAHVVKGG-T-S-PVVAQVPAAFYMQSVHLPGK-QTLAVKRKA-SEEERDD-S-L-S-LP-KASP-AESPKV--EK-SLGEK
Ph      ----------------------------------------------------------------------------------------------------------------
HPH2                        ML--G---------A--------------A-VPA---------P--P--------------R----------PA---P-----------------

                                                  HOMOLOGY DOMAIN I
HPH1    AESVANVNANTPSSELVALTPAPSVPPPTLAMVSRQMGDSKPPQAIVKPQILTHIIEGFVIQEGAEPFPGGCSQLLKESEKPLQTGLPTGLTENQSGGPFGVDSPSAELDKKANL  805
Rae28   AE-VA--NAN-P-SDLVAL-P-PS-PPPTLA-VSRQMGDSKPPQAIVKPQILTHIIEGFVIQEGAEPFP-GCSQ-LKE-EKPLQ-GLPTGL---Q-G---G-DSPS-EL-KKANL
Ph      ----------------------------------G---P-A--KP--LTHVI-GF-IQE--EPFP-----------------------------------------------
HPH2    -E-----------------------------A----Q-G--KPPQAIVKPQILTHVIEGFVIQEGA----------K--Q--LP--L-----------------------

HPH1    LKCEYCGKYAPAEQFRGSKRFCSMTCAKRYNVSCSHQFRLKRKKMKEFQEANYARVRRRGPRRSSSDIARAKIQGKCHRGQEDSSRGSDNSSYDEALSPTSPGPLSVRAGHGER  920
Rae28   LKCEYCGKYAPAEQFRGSKRFCSMTCAKRYNVSCSHQFRLKRKKMKEFQEA-YARVRRRGPRRSSSDIARAKIQGK-HRGQEDSSRGSDNSSYDEALSPT-PGPLSVRAGHGER
Ph      --CE-CG----A------KR-CS--C-------------------------------------------------------------------------------------
HPH2    LKCE-CG----A--F--SKRFCSM-CAKRYNV-C-----L--------Q-A--A---RR-P------I-R---Q------QEDSSR-SDNSSY-E-LSP-S-------AG----

                                                  HOMOLOGY DOMAIN II
HPH1    DLGNPNTAPPTPELHGINPVFLSSNPSRWSVEEVYEFIASLQGCQEIAEEFRSQEIDGQALLLLKEEHLMSAMNIKLGPALKICAKINVLKET                        1013
Rae28   DLGN--T-P-TPEL-GINPVFLSSNPS-WSVEEVYEFIASLQGCQEIAEEFRSQEIDGQALLLLKEEHLMSAMNIKLGPALKICAKINVLKET
Ph      -------------------------S-WSV-DV--FI--LPGCQ-----F--QEIDGQALLLLKE-HL--AM--KLGPALKI-A-----KE-
HPH2    ------T----------------S-+----VEDVYEFI-SLPGCQEIAEEFR-QEIDGQALLLLKE-HLMS-MNIKLGPALKI-A-I--LK--
```

FIG. 2. Predicted amino acid sequence of HPH1. The sequence is aligned to the Rae28/Mph1, Ph, and HPH2 sequences. Identical amino acids are shown. The conserved amino acids of the putative zinc finger domain are shaded. Homology domains I and II are boxed.

no cDNAs which contained additional 5′ sequences. We then used the coding sequences of *HPH2* as probe, as well as the most upstream 5′ 1,000 bp of the coding region of *Rae28/Mph1*. We isolated a cDNA clone that overlapped with the *HPH2* two-hybrid clone, extending HPH2 with 140 aa (Fig. 1) and the *HPH2* cDNA to 2,554 bp. The length of the *HPH2* transcript and the molecular mass of HPH2 (see below) suggest that we isolated the full-length *HPH2* cDNA.

Besides isolation of the extended *HPH2* clone, we also isolated two overlapping cDNA clones which are, in contrast with *HPH2*, highly homologous with *Rae28/Mph1* (Fig. 2). The predicted 1,012-aa long protein is 95% identical with Rae28/Mph1 and is only homologous with HPH2 and Ph in homology domains I and II and the Cys2-Cys2 zinc finger (Fig. 2). Besides these sequence homologies, the protein shares a glutamine-rich stretch (aa 400 to 450) with Ph (Fig. 2 [9, 10, 20]). Given the high degree of homology with *Ph*, but also with the earlier characterized *Rae28/Mph1*, we named this clone *HPH1.*

Homology domain II is also shared with other proteins. The products of the recently identified PcG gene, *Sex comb on midleg* (*Scm*) and of the *Drosophila* gene *lethal(3)malignant brain tumor* (*l(3)mbt*) share homology domain II (referred to as the SPM domain) with Ph, HPH1, and HPH2 (5, 36). The SPM domain in the Scm protein is essential for its function (5). The occurrence of this domain in these diverse proteins suggests the existence of a family of proteins containing these homology domains.

In conclusion, a two-hybrid screen with the *Xenopus* homolog of the vertebrate PcG protein Bmi1 as the target resulted in the isolation of a human cDNA which encodes a protein that shares two homology domains with Rae28/Mph1 and the PcG protein Ph. Subsequent screening of a cDNA library to isolate full-length clones of this cDNA led to the isolation of a second cDNA which is the human homolog of *Rae28/Mph1*. We named the proteins HPH1 and HPH2. The longer HPH1 protein is the virtually identical human homolog of Rae28/Mph1. The shorter HPH2 protein which was isolated in the two-hybrid screen shares the homology domains, but further it has only limited homology with HPH1.

Distribution of *HPH1* and *HPH2* transcripts in human tissues and cancer cell lines. Besides limited sequence homology outside the homology domains, the *HPH2* cDNA clone is considerably shorter (2.55 kb) than the *HPH1* and *Rae28/Mph1* cDNA clones (3.2 and 3.5 kb, respectively [20]). To verify this difference in length at the level of mRNA transcripts, we characterized the expression levels of the *HPH1* and *HPH2* genes. We analyzed multiple tissue Northern blots containing poly(A)$^+$ mRNA from different human tissues or human cell lines (Clontech). As probes, we selected the 5′ coding region from the *HPH1* cDNA, since this region is not present in the *HPH2* cDNA (Fig. 2) and the 3′ noncoding region from the *HPH2* cDNA.

We detected approximately 4.4- and 6-kb transcripts for the *HPH1* gene and a single approximately 2.5-kb transcript for the *HPH2* gene (Fig. 3 and 4). The length of the transcripts confirms that *HPH1* and *HPH2* are different genes. The highest level of expression of *HPH1* is found in thymus (lane 2), testis (lane 4), and ovary (lane 5) (Fig. 3). Low expression levels were detected in spleen (lane 1), prostate (lane 3), small intestine (lane 6), colon (lane 7), and peripheral blood leukocytes (lane 8). In contrast, *HPH2*, like *BMI1*, was expressed ubiquitously (Fig. 3).

The differences in the abundance of *HPH1*, *HPH2*, and *BMI1* transcripts are more pronounced in human cell lines than in normal human tissues (Fig. 4). The *HPH1* gene was strongly expressed in the Burkitt's lymphoma Raji (lane 5),

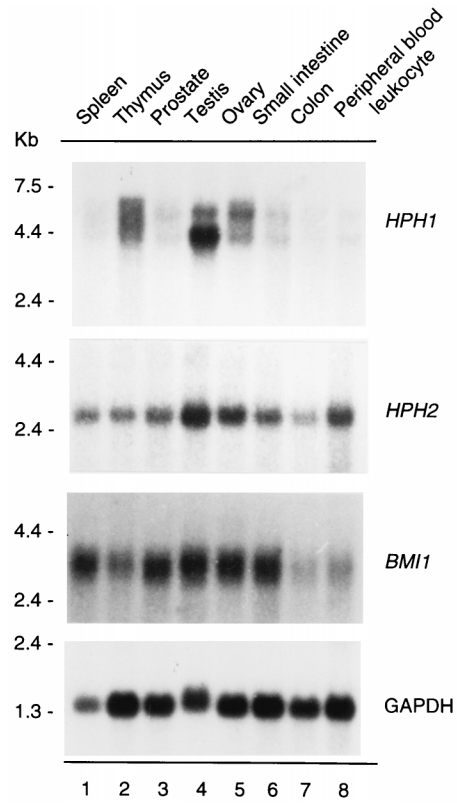

FIG. 3. Expression patterns of *HPH1*, *HPH2*, and *BMI1* in human tissues. Shown are expression levels in spleen (lane 1), thymus (lane 2), prostate (lane 3), testis (lane 4), ovary (lane 5), small intestine (lane 6), colon (lane 7), and peripheral blood leukocytes (lane 8). As probes, we selected the full-length *BMI1* cDNA, the 5′ coding region from the *HPH1* cDNA, since this region is not present in the *HPH2* cDNA, and the 3′ noncoding region from the *HPH2* cDNA. The filter was rehybridized with a probe for GAPDH to verify the loading of RNA in each lane.

colorectal adenocarcinoma SW480 (lane 6), melanoma G381 (lane 8), and U-2 OS osteosarcoma (lane 10) cell lines and at a much lower level in the other cell lines (Fig. 4). In contrast, both *HPH2* and *BMI1* were expressed at a low level in Burkitt's lymphoma Raji cells. As in normal human tissues, no major differences in expression levels of *HPH2* were observed, except in Burkitt's lymphoma Raji cells (lane 5). An aberrant *HPH2* transcript was detected, specifically in the colorectal adenocarcinoma SW480 (lane 6). This approximately 5.5-kb transcript is expressed at a much lower level than the 2.5-kb transcript. No aberrant transcripts were detected in the mRNA from the other cell lines, not even after longer exposures. *BMI1* expression is more varied than *HPH2* expression, *BMI1* being most strongly expressed in the K-562 (lane 3), SW480 (lanes 6 and 9), and U-2 OS (lane 10) cell lines.

In summary, the *HPH1*, *HPH2* and *BMI1* genes are differentially expressed in normal human tissues and cancer cell lines. The length of the *HPH1* and *HPH2* transcripts confirms the existence of two different genes and also verifies that the *HPH2* cDNA is considerably shorter than the *HPH1* cDNA.

HPH1 and HPH2 coimmunoprecipitate with BMI1. In order to test whether HPH1, HPH2, and BMI1 exist in vivo as part of a common protein complex, we performed immunoprecipitation experiments with antibodies raised against the Rae28/Mph1, HPH2, and Bmi1 proteins. The predicted molecular mass of Rae28/MPh1, which is virtually identical to that of

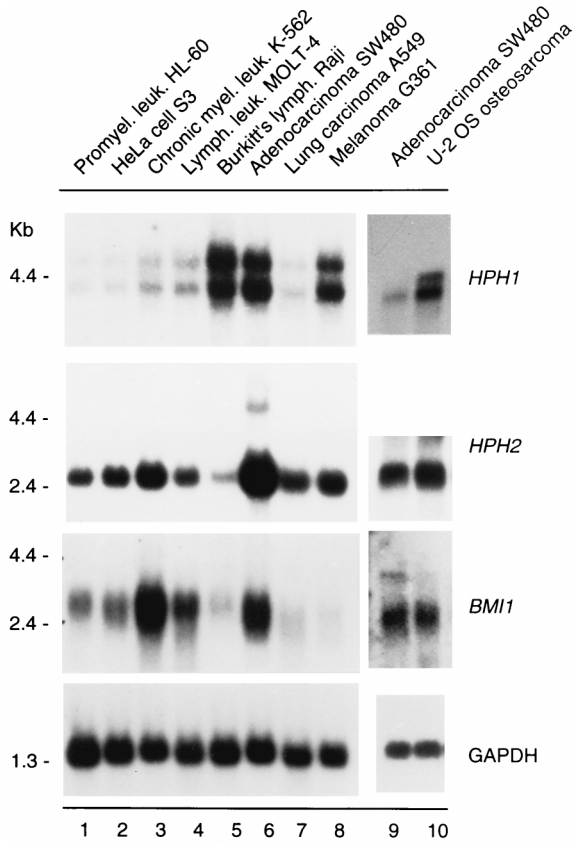

FIG. 4. Expression patterns of *HPH1*, *HPH2*, and *BMI1* in human cell lines. Shown are expression levels in the promyelocytic leukemia HL-60 (lane 1), HeLa S3 (lane 2), chronic myelogenous leukemia K-562 (lane 3), lymphoblastic leukemia MOLT-4 (lane 4), Burkitt's lymphoma Raji (lane 5), colorectal adenocarcinoma SW480 (lane 6), lung carcinoma A549 (lane 7), and melanoma G361 (lane 8) cell lines. The same probes as in Fig. 3 were used. Lanes 1 to 8 represent a commercially obtained Northern blot. We also isolated and blotted poly(A)$^+$ RNA from U-2 OS (lane 10). To allow comparison with the commercial multiple tissue Northern blot, we isolated and blotted poly(A)$^+$ RNA from SW480 cells. Note that although the levels of expression of *HPH2* and *BMI1* in SW480 are comparable, the level of expression of *HPH1* in SW480 cells is substantially lower in our hands (compare lanes 6 and 9).

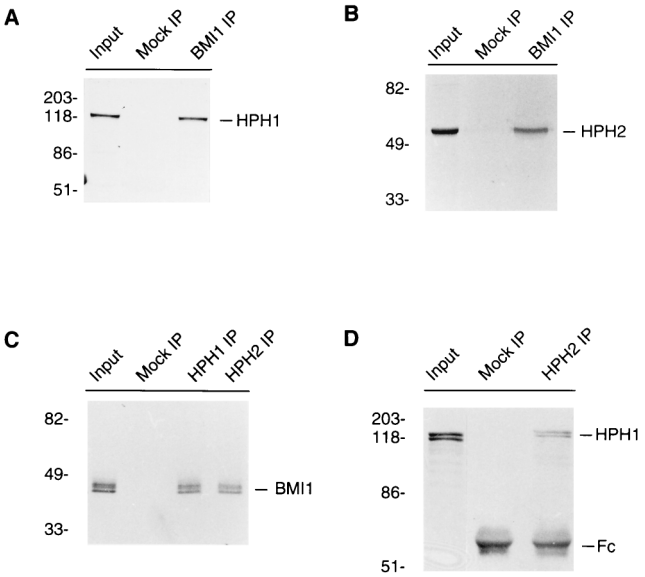

FIG. 5. Coimmunoprecipitation of HPH1 and HPH2 with BMI1 from human U-2 OS osteosarcoma cells. (A) Extracts of U-2 OS human osteosarcoma cells were incubated with a mouse anti-Bmi1 antibody and subsequently with agarose-coupled goat anti-mouse IgG antibodies. The resulting IP was Western blotted and incubated with rabbit anti-Rae28/Mph1 antibody. The approximately 120-kDa HPH1 protein (Input) was detected in the BMI1 BMI1 IP. In this particular experiment, HPH1 runs as a single protein, but when a lower percentage of polyacrylamide is used, the HPH1 protein runs as a doublet (compare with panel D). When the Bmi1 antibody was omitted, no HPH1 protein was detected (Mock IP). (B) The mouse anti-Bmi1 antibody immunoprecipitated the approximately 54-kDa HPH2 protein (Input) from extracts of U-2 OS cells (BMI1 IP). No HPH2 protein was detected when the Bmi1 antibody was omitted from the immunoprecipitation (Mock IP). (C) The HPH1 (HPH1 IP) and HPH2 (HPH2 IP) antibodies immunoprecipitated the approximately 44- and 46-kDa BMI1 proteins (Input) from extracts of U-2 OS cells. No BMI1 protein was detected when the HPH antibodies were omitted from the immunoprecipitation (Mock IP). (D) The HPH2 antibody immunoprecipitated the approximately 120-kDa HPH1 protein (Input) from extracts of U-2 OS cells (HPH2 IP). No HPH1 protein was detected when another, unrelated antibody was used in the immunoprecipitation (Mock IP). Note the rabbit heavy chain (Fc) in the HPH2 and mock IP.

HPH1, is 108 kDa. On Western blots, the affinity-purified antibody recognizes a protein with a size of approximately 120 kDa (Fig. 5A). However, when a lower-percentage gel is used, it becomes apparent that HPH1 runs as a doublet protein (Fig. 5D). It is possible that the HPH1 protein doublet is a reflection of the two, 4.4- and 6-kb *HPH1* transcripts (Fig. 3 and 4). It is, however, also possible that the HPH1 protein doublet arises from posttranslational modifications.

The highest expression levels of the *HPH1*, *HPH2*, and *BMI1* genes are found in the SW480 and U-2 OS cell lines (Fig. 4). We used extracts of U-2 OS human osteosarcoma cells to perform coimmunoprecipitation experiments. Extracts of U-2 OS cells were incubated with a mouse anti-Bmi1 antibody and subsequently with agarose-coupled goat anti-mouse IgG antibodies. This monoclonal antibody is raised against murine Bmi1 protein and recognizes the human BMI1 protein (2). The resulting immunoprecipitate was Western blotted and incubated with rabbit anti-Rae28/Mph1 antibody. The 120-kDa HPH1 protein was detected in the BMI1 immunoprecipitate (Fig. 5A, BMI1 IP). The predicted molecular mass of HPH2, based on amino acid composition, is 47 kDa. On Western blots, the affinity-purified antibody recognizes a single 54-kDa protein (Fig. 5B, Input). The mouse anti-Bmi1 antibody was able to immunoprecipitate the HPH2 protein from extracts of U-2 OS cells (Fig. 5B, BMI1 IP). The BMI1 immunoprecipitations were specific since no HPH1 or HPH2 proteins were detected when the BMI1 antibody was omitted from the immunoprecipitations (Fig. 5A and B, Mock IP). Also when the unrelated mouse monoclonal antibody 5E10, which recognizes the PML protein (31), was used in the immunoprecipitation, no HPH1 or HPH2 protein was detected (data not shown), underlining the specificity of the immunoprecipitation of HPH1 and HPH2 by the BMI1 antibody.

The predicted molecular mass of human BMI1, based on amino acid composition, is 38 kDa (2). On Western blots, the antibody recognizes a doublet of 44- and 46-kDa proteins (Fig. 5C, Input). Both the Rae28/Mph1 and HPH2 antibodies were able to immunoprecipitate the BMI1 protein from extracts of U-2 OS cells (Fig. 5C, HPH1 IP and HPH2 IP, respectively). No BMI1 was detected when the antibodies were omitted (Fig. 5C, Mock IP) or when an unrelated antibody was added to the immunoprecipitations.

These results suggest that in vivo, HPH1 and HPH2 are part of a protein complex which contains BMI1. We further strengthened this notion by showing that the HPH2 antibody was able to immunoprecipitate the HPH1 protein from ex-

tracts of U-2 OS cells (Fig. 5D, HPH2 IP). The HPH2 immunoprecipitations were specific, since no HPH1 proteins were detected when an unrelated antibody was added to the immunoprecipitations (Fig. 5D, Mock IP). We also tried to immunoprecipitate HPH2 by using the Rae28/Mph1 antibodies. Since both Rae28/Mph1 and HPH2 antibodies are rabbit derived, the heavy chains (Fc) of the rabbit anti-HPH2 antibodies (Fig. 5D, Mock IP and HPH2 IP) or the rabbit anti-Rae28/Mph1 are detected by the goat-anti rabbit alkaline-phosphatase-conjugated antibody. Unfortunately, HPH2 migrates at the same position of these Fc heavy chains (compare Fig. 5B and D), thus preventing detection of HPH2 in IPs by the Rae28/Mph1 antibodies.

The major part of the immunoprecipitated or coimmunoprecipitated HPH1, HPH2, and BMI1 proteins was present in the pellet, indicating that the proteins indeed coimmunoprecipitated. We could, however, still detect residual amounts of BMI1, HPH1, or HPH2 proteins in the supernatants after immunoprecipitation with the BMI1, Rae-28/Mph1, or HPH2 antibodies (data not shown). Therefore, neither the immunoprecipitation nor the coimmunoprecipitation can be considered quantitative. However, since the majority of the proteins were removed from the supernatants and present in the pellets, this indicates that HPH1, HPH2, and BMI1 proteins coimmunoprecipitate from extracts of U-2 OS cells. This suggests that in vivo, the proteins are part of a common protein complex.

HPH1, HPH2, and BMI1 cofractionate in a sucrose gradient. The coimmunoprecipitation experiments suggest that HPH1, HPH2, and BMI1 may be part of a high-molecular-mass protein complex. To test this idea, we performed velocity gradient ultracentrifugation experiments. We loaded extracts of U-2 OS cells, which had been prepared as in the immunoprecipitation experiments, on a 5 to 30% sucrose gradient (5 ml). After centrifugation 14 350-μl fractions were collected. We calibrated the lower-molecular-mass range of the gradient with catalase (232 kDa) and thyroglobulin (669 kDa), which were detected in fractions 1 and 3, respectively (Fig. 6). Lack of suitable markers prevented calibration of the higher-molecular-mass range of the gradient. The proteins in the 14 fractions of the gradient on which the U-2 OS cell extract was loaded were separated by SDS-PAGE, blotted on nitrocellulose, and probed with antibodies against HPH1, HPH2, and BMI1. We found that the largest portion of the HPH1, HPH2, and BMI1 proteins cofractionate in fractions 7 through 11 (Fig. 6). This result indicates that the HPH1, HPH2, and BMI1 proteins are present in a complex with a molecular mass higher than 669 kDa.

These results are in agreement with earlier results that demonstrated that the *Drosophila* Polycomb and Polyhomeotic proteins cofractionate by gel filtration and velocity gradient ultracentrifugation (13). We used a gradient ranging from 5 to 30% sucrose, which we found to result in a better separation than when the reported 5 to 20% gradient was used (13). Similar to what has been reported previously (13), we found a portion of the HPH1 and BMI1 proteins to be present in the pellet of the gradient (Fig. 6, fraction 14). We also found a substantial portion of the HPH1 and BMI1 proteins to be present in fraction 1 (Fig. 6). In contrast, no HPH2 protein was detected in either fraction 1 or fraction 14. The meaning of these observations is unclear. It has been speculated that the presence of the proteins in the lower-molecular-mass range of the gradient can be explained by partial dissociation of the complex during the preparation (13). Alternatively, it could indicate a heterogeneous composition of the multimeric complex (13). At present, it is not clear which is the right interpretation. The presence of the HPH1, HPH2, and BMI1 proteins in fractions 7 through 11, however, supports the idea that the proteins are part of a common, multimeric protein complex.

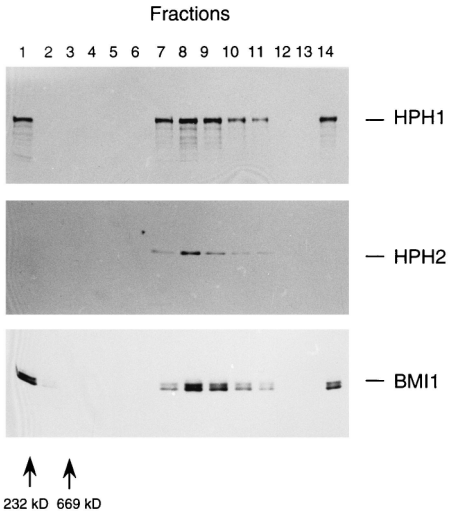

FIG. 6. Cofractionation of HPH1, HPH2, and BMI1 from human U-2 OS osteosarcoma cells. Cell extract was loaded on a 5 to 30% sucrose gradient, and the gradient was centrifuged in a TST50 (Kontron) rotor at 50,000 rpm for 8 h. Fourteen 350-μl fractions (indicated on top) were collected, and proteins were precipitated and separated on two SDS-PAGE gels and transferred to nitrocellulose. The upper halves of both blots were probed with anti-Rae28/Mph1 antibody to detect the approximately 120-kDa HPH1, and the lower halves were probed with, respectively, anti-HPH2 (54-kDa) (middle panel) or anti-BMI1 (44- to 46-kDa) (lowest panel) antibody. This procedure was followed since the similar molecular mass ranges of HPH2 and BMI1 excluded the possibility of detecting all three proteins on a single blot. By using a separate 5 to 30% gradient, we calibrated the lower-molecular-mass range of the gradient by using catalase (232 kDa) and thyroglobulin (669 kDa) (Pharmacia, Uppsala, Sweden) as markers. The fractions in which the markers were detected are indicated by arrows. The fractions containing a higher percentage sucrose were not calibrated because of a lack of suitable markers.

HPH1 and HPH2 colocalize with BMI1 in nuclei of U-2 OS and SW480 cells. We next analyzed the subcellular localization of the HPH1 and HPH2 proteins in relation to the PcG protein BMI1 by performing immunofluorescence labelling experiments. The use of mouse anti-Bmi1 and rabbit anti-Rae28/Mph1 and HPH2 allows double-labelling experiments. We used U-2 OS human osteosarcoma cells and SW480 human colorectal adenocarcinoma cells, in which we found that *HPH1*, *HPH2*, and *BMI1* are all expressed at a high level (Fig. 4). In extracts of U-2 OS cells, we found that HPH1 and HPH2 coimmunoprecipitate and cofractionate with each other and with BMI1 (Fig. 5 and 6). In U-2 OS cells, both HPH1 and BMI1 are found in the nucleus and throughout the nucleoplasm, and they colocalize in large, brightly labelled domains (Fig. 7A to C). Also HPH2 and BMI1 colocalize in the same brightly labelled domains (Fig. 7D to F). HPH1 and HPH2 colocalize with BMI1 in the large domains, but as yet, this is not clear for the more homogeneously distributed pattern. The fine granular pattern is too complex to allow analysis of any systematic colocalization.

The observed labelling patterns of HPH1, HPH2, and BMI1 are in contrast with recently published labelling patterns of the Bmi1 protein (8). It should be pointed out, however, that these workers detected overexpressed Bmi1 protein in Rat1 fibroblasts or COS cells in which no endogenous Bmi1 could be detected (8). We, on the other hand, detected endogenous levels of BMI1 in cell lines which express the *BMI1* gene at a high level (Fig. 4). This is a likely explanation for the observed differences in labelling patterns. The significance of the do-

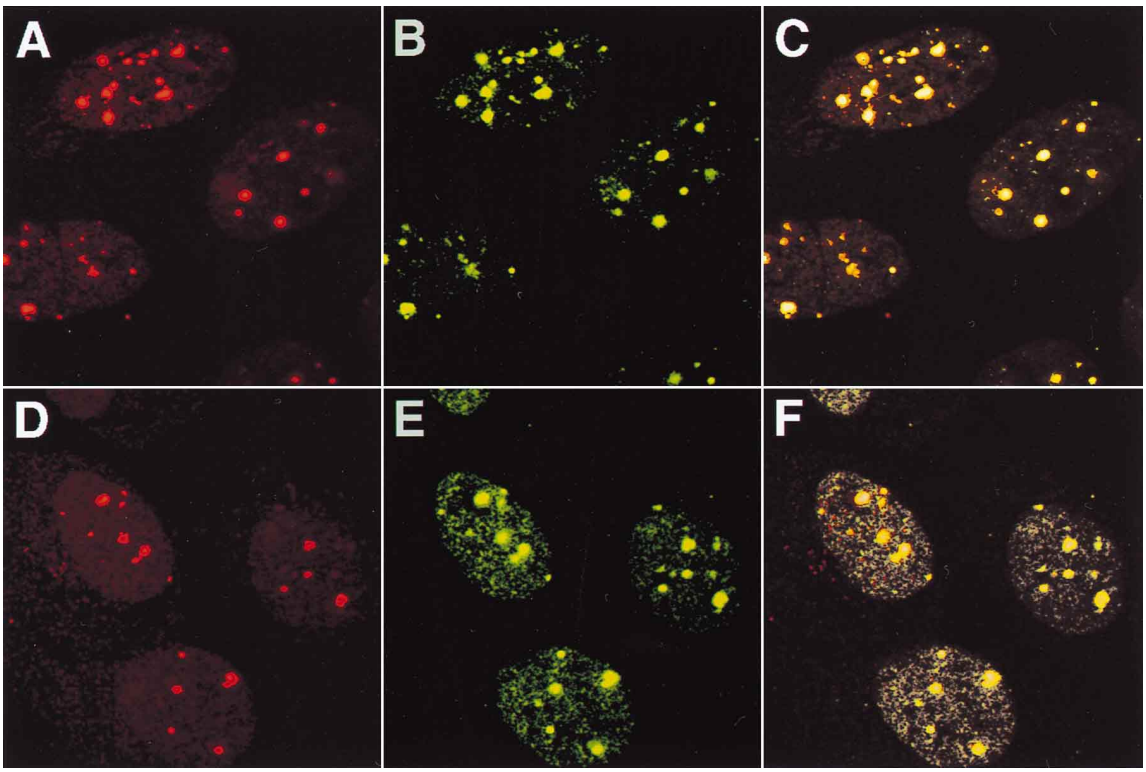

FIG. 7. Colocalization of HPH1 and HPH2 with BMI1 in nuclei of U-2 OS cells. (A to C) Rabbit anti-Rae28/Mph1 (A) and mouse anti-BMI1 (B) labelling is shown in single confocal optical sections. Panels A and B represent the two different detection channels of the dual image, whereas panel C represents the false color overlay. HPH1 and BMI1 are homogeneously distributed in the nucleus, but are also concentrated in large, brightly labelled domains. HPH1 and BMI1 colocalize in these large domains (C) (indicated by yellow). With rabbit anti-HPH2 (D) and mouse anti-BMI1 (E) antibodies, similar distribution patterns and colocalizations (F) were observed.

mains in which HPH1, HPH2, and BMI1 colocalize is further underlined by our recent finding that a human homolog of the Polycomb protein also colocalizes with BMI1 in these domains (27). This human Pc homolog also coimmunoprecipitates with BMI1 (27).

We further analyzed the subcellular localization of HPH1 and HPH2 in relation to BMI1 in SW480 cells. As in U-2 OS cells, we found complete colocalization between HPH1 (Fig. 8A to C) and HPH2 (Fig. 8D to F) and BMI1 in large, brightly labelled domains. These results demonstrate that the colocalization of HPH1 and HPH2 with the PcG protein BMI1 is not restricted to one cell type.

We conclude that HPH1 and HPH2 colocalize with BMI1 in large domains in the nuclei of U-2 OS and SW480 cells. These results further underline the validity of the identified two-hybrid interaction between HPH2 and BMI1. They strengthen the notion that HPH1, HPH2, and BMI1 are part of a multimeric protein complex.

Interactions between homology domains I and II of HPH1 and HPH2. We found that HPH2 is able to immunoprecipitate HPH1 (Fig. 5D), that HPH1 and HPH2 cofractionate in a sucrose gradient (Fig. 6), and that HPH1 and HPH2 colocalize with BMI1 in the nuclei of interphase cells (Fig. 7 and 8). These results suggest that HPH1 and HPH2 are part of a common protein complex, but they do not address the question of whether HPH1 and HPH2 interact with each other. Homology domains I and II are potentially involved in protein-protein interactions. Significant homology has been found between the Rae28/Mph1 and Ph homology domain II and the recently characterized protein encoded by the PcG gene *Sex comb on midleg* (*Scm*) (5). This domain has been named SPM domain and it has been suggested that this domain is involved in protein-protein interactions (5). By using the two-hybrid assay, we therefore tested whether the homology domains of HPH1 and HPH2 can mediate protein-protein interactions. The region encompassing homology domains I and II from HPH1 (aa 722 to 1013 [Fig. 9A]) was fused to the GAL4 DNA binding domain. The plasmid was cotransformed with the HPH2 two-hybrid clone, encompassing homology domains I and II (aa 137 to 432 [Fig. 9A]) (see Fig. 1). The transformants were able to grow on selective medium lacking histidine, in the presence of 3-aminotriazole, and were also β-galactosidase positive. This result indicates an interaction between homology domains I and II of HPH1 and HPH2, and we subsequently tested potential interactions between the individual homology domains. We found that homology domain I of HPH1 interacts with homology domain II, but not with homology domain I of HPH2. Also, homology domain II of HPH1 interacts with both homology domains I and II of HPH2 (Fig. 9A). In contrast, BMI1 does not interact with individual homology domains I and II. For an interaction with BMI1, both intact homology domains I and II need to be present in the fusion protein (Fig. 9B).

To quantify the strength of the interactions between the different portions of the HPH1 and HPH2 proteins, we prepared lysates of the transformants and measured the β-galactosidase activity. The interaction between Xbmi1 and HPH2 (aa 137 to 375), which was identified in our original two-hybrid screen (Fig. 1), was found to be the strongest. The relative strengths of the other interactions are given as percentage of the Xbmi1-HPH2 (aa 137 to 375) interaction (Table 1). The second strongest interaction was found between HPH1 (aa 722

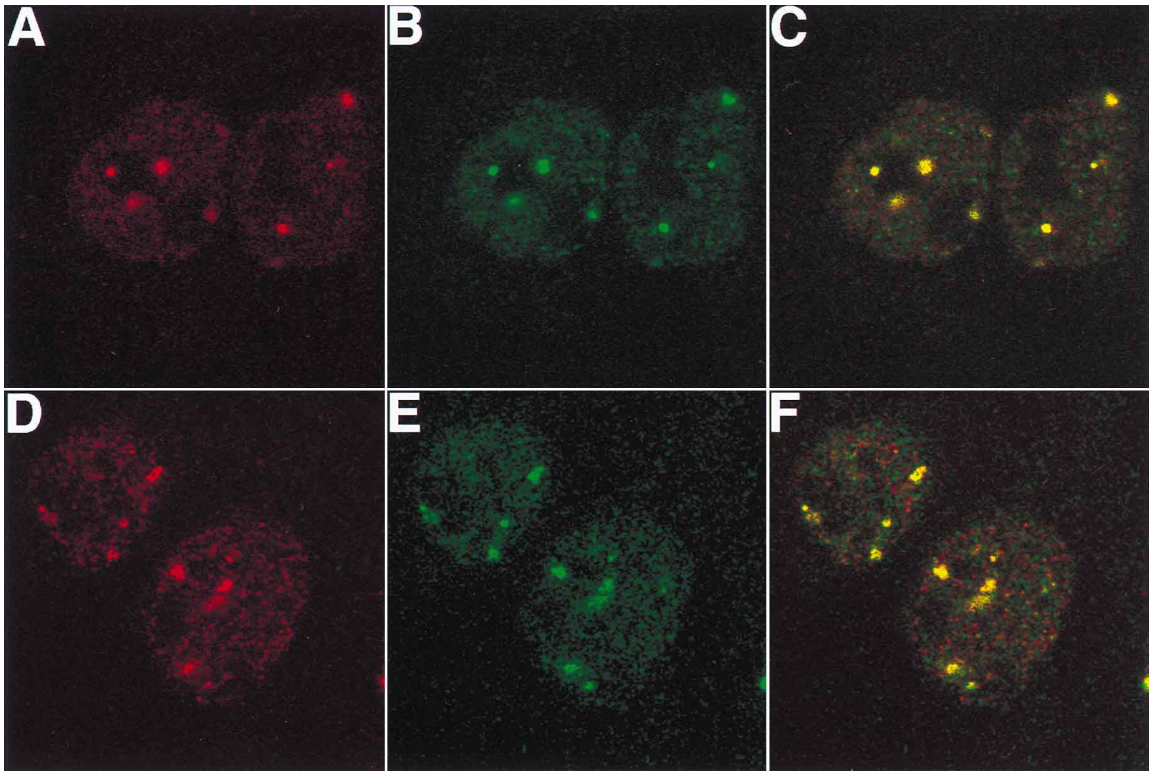

FIG. 8. Colocalization of HPH1 and HPH2 with BMI1 in nuclei of SW480 cells. (A to C) Rabbit anti-Rae28/Mph1 (A) and mouse anti-BMI1 (B) labelling and the false color overlay (C) are shown. Similar to in U-2 OS cells (Fig. 7), HPH1 and BMI1 colocalize in large, brightly labelled domains (C) (indicated by yellow). With rabbit anti-HPH2 (D) and mouse anti-BMI1 (E) antibodies, similar distribution patterns and colocalizations (F) were observed.

to 1013) and HPH2 (aa 137 to 432), in which both intact homology domains of HPH1 and HPH2 were present. When single homology domains were tested against each other, a decrease in the strength of the interactions was observed (Table 1).

Taken together, these results demonstrate the importance of the homology domains in protein-protein interactions. They further suggest that HPH1 and HPH2 can heterodimerize in vivo through these homology domains.

DISCUSSION

Are HPH1, HPH2, and the PcG protein BMI1 part of a multimeric protein complex? It has been proposed that PcG proteins repress homeotic gene expression via the formation of multimeric complexes. This model is based on the observation that different PcG proteins bind in overlapping patterns to about 100 loci on polytene chromosomes in the *Drosophila* salivary gland (10, 13, 18, 25, 37). However, so far, only the PcG protein Polyhomeotic, but no other known proteins, has been shown to coimmunoprecipitate or cofractionate with Pc (13).

Employing the genetic two-hybrid system, we identified a protein that specifically interacts with the vertebrate PcG protein BMI1. Based on the occurrence of two conserved homology domains that are shared with Ph and the vertebrate Rae28/Mph1 protein, we name this protein HPH2 (see below). We also identified the human homolog of the protein expressed by the *Rae28/Mph1* gene, which we name HPH1. Our results indicate that mammals have at least two proteins that bear similarity to the *Drosophila* Ph protein. The homology domains appear to be involved in mediating interactions between HPH1 and HPH2 as well as between Xbmi1 and HPH2. We further show that HPH1 and HPH2 coimmunoprecipitate with BMI1 and with each other, indicating an in vivo association between these proteins. This notion is further supported by our finding that HPH1, HPH2, and BMI1 cofractionate in a sucrose gradient. Furthermore, HPH1, HPH2, and BMI1 colocalize in interphase nuclei of two different human lines. Together, these data provide evidence for an in vivo association and indicate that HPH1 and HPH2 are part of a human multimeric protein complex which also contains the vertebrate PcG protein BMI1.

In summary, four different approaches, namely (i) the two-hybrid assay, (ii) coimmunoprecipitation experiments, (iii) cofractionation experiments, and (iv) immunocolocalization experiments, all indicate that HPH1, HPH2, and BMI1 are part of a common, multimeric protein complex.

Are HPH1 and HPH2 homologs of the PcG protein Polyhomeotic? Two lines of evidence suggest that HPH1 and HPH2 are functional homologs of the *Drosophila* PcG protein Polyhomeotic (Ph). In the first place, based on four different approaches (see above) we find evidence for the notion that HPH1, HPH2, and BMI1 are part of a multimeric protein complex. BMI1 is a vertebrate PcG protein (3, 32, 33) and possibly a functional vertebrate homolog of the *Drosophila* PcG protein Posterior sex combs (6, 35). It is feasible that proteins that form a complex with PcG proteins are also PcG proteins. In the second place, the HPH1 and HPH2 proteins share extensive sequence and overall structural homology with the *Drosophila* PcG protein Ph. Both HPH1 and HPH2 share the conserved homology domains I and II with Ph. In addition, a particular Cys2-Cys2 zinc finger motif is conserved, both in terms of sequence homology and in terms of its position between the two homology domains. Also, a glutamine-rich re-

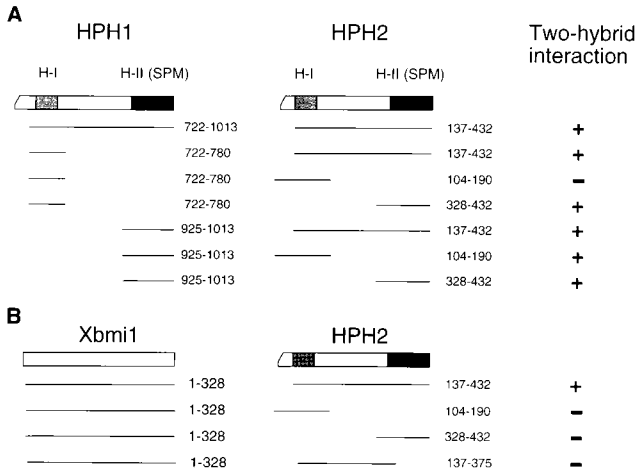

FIG. 9. Interactions between homology domains I and II (H-I and H-II, respectively) of HPH1 and HPH2. (A) Indicated portions of HPH1 were fused to the GAL4 DNA binding domain. These HPH1 regions encompass homology domains I and II (aa 722 to 1013), homology domain I (aa 722 to 780), and homology domain II (aa 925 to 1013) alone. The plasmids were cotransformed with the originally isolated HPH2 two-hybrid clone (aa 137 to 432), which is fused to the GAL4 transactivation domain. Also, homology domain I (aa 104 to 190) and homology domain II (aa 328 to 432) of HPH2 were tested against the indicated portions of HPH1. All interactions were positive (+), as judged by the ability of the transformants to grow on selective medium lacking histidine, and they were also β-galactosidase positive. (B) Full-length Xbmi1 (aa 1 to 328), fused to the GAL4 DNA binding domain, was tested for interaction against indicated portions of HPH2, which are fused to the GAL4 transactivation domain. The quantitative analysis of the interactions is represented in Table 1. Homology domain II has also been termed SPM (5).

gion is located in HPH1 in the same region of the protein as in Ph. In this context, a similar situation with the murine *Pc* homolog M33 is very relevant. M33 shares only the conserved chromodomain (52 aa long) and C terminus (17 aa long) with Pc (24), but it is still able to partly rescue the *Drosophila Pc* phenotype, when overexpressed in the *Pc* mutant (19). Although the homology is restricted to small regions and the overall composition of the protein, M33 can be considered to be a functional homolog of Pc.

Both lines of evidence are supportive for the idea that HPH1 and HPH2 are human homologs of Ph. Although the sequence and structural homologies between Ph and Rae28 have been noted before (20), at that time it was unclear whether this implied that the Rae28 protein is a functional, mammalian homolog of the PcG protein Ph. We now show that the human Rae28 homolog, HPH1, as well as HPH2, interacts with and colocalizes with a noted member of the vertebrate PcG family, BMI1. We are, therefore, tempted to conclude that we have identified components of a human PcG protein complex. It will be of considerable interest to test whether loss-of-function mutations of the mouse *Rae28/Mph1* gene and the murine homolog of *HPH2* result in homeotic transformations.

Possible functional significance of the existence of multiple HPH proteins. We have presented evidence that HPH1, HPH2, and BMI1 are part of a multimeric protein complex. Our data also indicate that homology domains I and II of HPH1 and HPH2 are involved in mediating protein-protein interactions. These data suggest that HPH1 and HPH2 are able to form heterodimers through their respective homology domains I and II. This may imply that their function is exerted via heterodimerization and that the human multimeric PcG complex requires the presence of both HPH1 and HPH2.

The occurrence of two proteins with homology to Ph may also point towards the interesting possibility of the existence of different mammalian multimeric PcG complexes. This idea is supported by the differential expression patterns of *HPH1*, *HPH2*, and *BMI1* (Fig. 3 and 4). *HPH2*, like *BMI1*, is expressed ubiquitously in all tissues and cell lines tested, whereas *HPH1* has a more limited expression pattern. Since we screened a peripheral blood leukocyte two-hybrid library, it is not amazing that we isolated HPH2 as a BMI1 binding protein. *HPH2* is abundantly expressed in this tissue, and *HPH1* is expressed only at a very low level. This might indicate that in leukocytes, HPH2 and BMI1 are components of the PcG complex and HPH1 is not.

The idea that PcG complexes exist with different compositions is reinforced by the observation that the PcG protein Psc shares many, but not all binding sites with the PcG proteins Pc, Ph, and Pcl on salivary gland polytene chromosomes (10, 18, 25, 37). Distinct compositions of the PcG complexes could lead to specificity in the repression of different loci in different tissues. This provides flexibility, and it increases the possibilities of regulating a wide range of target genes with only a limited number of components.

ACKNOWLEDGMENTS

M.J.G. and D.P.E.S. contributed equally to this work.

We thank Wouter Schul and Ineke van der Kraan for help with confocal laser microscopy and Thijs Hendrix for raising rabbit antibodies.

REFERENCES

1. **Akasaka, T., M. Kanno, R. Balling, M. A. Mieza, M. Taniguchi, and H. Koseki.** 1996. A role for *mel-18*, a *Polycomb* group-related vertebrate gene, during the anteroposterior specification of the axial skeleton. Development **122:**1513–1522.
2. **Alkema, M. J., J. Wiegant, A. K. Raap, A. Berns, and M. van Lohuizen.** 1993. Characterization and chromosomal localization of the human proto-oncogene BMI-1. Hum. Mol. Genet. **2:**1597–1603.
3. **Alkema, M. J., N. M. T. van der Lugt, R. C. Bobeldijk, A. Berns, and M. van Lohuizen.** 1995. Transformation of axial skeleton due to overexpression of bmi-1 in transgenic mice. Nature (London) **374:**724–727.
4. **Alkema, M. J., M. Bronk, E. Verhoeven, A. P. Otte, L. J. van't Veer, A. Berns, and M. van Lohuizen.** 1997. Genes Dev. **11:**226–240.
5. **Bornemann, D., E. Miller, and J. Simon.** 1996. The *Drosophila Polycomb* group gene *Sex comb on midleg (Scm)* encodes a zinc finger protein with similarity to Polyhomeotic protein. Development **122:**1621–1630.
6. **Brunk, B. P., E. C. Martin, and P. N. Adler.** 1991. *Drosophila* genes *Posterior sex combs* and *Suppressor (2) of zeste* encode proteins with homology to the murine *bmi-1* oncogene. Nature (London) **353:**351–353.
7. **Chien, C.-T., P. L. Bartel, R. Sternglanz, and S. Fields.** 1991. The two-hybrid system: a method to identify and clone genes for proteins that interact with a protein of interest. Proc. Natl. Acad. Sci. USA **88:**9578–9582.
8. **Cohen, K. J., J. S. Hanna, J. E. Prescott, and C. V. Dang.** 1996. Transformation by the Bmi-1 oncoprotein correlates with its subnuclear localization but not its transcriptional suppression activity. Mol. Cell. Biol. **16:**5527–5535.
9. **Deatrick, J., M. Daly, N. B. Randsholt, and H. W. Brock.** 1991. The complex genetic locus *polyhomeotic* in *Drosophila melanogaster* potentially encodes two homologous zinc-finger proteins. Gene **105:**185–195.
10. **DeCamillis, M., N. Cheng, D. Pierre, and H. W. Brock.** 1992. The *polyhomeotic* gene of *Drosophila* encodes a chromatin protein that shares polytene chromosome-binding sites with *Polycomb*. Genes Dev. **6:**223–232.
11. **Durfee, T., K. Becherer, P.-L. Chen, S.-H. Yeh, Y. Yang, A. E. Kilburn, W.-H. Lee, and S. J. Elledge.** 1993. The retinoblastoma protein associates with the protein phosphatase type I catalytic subunit. Genes Dev. **7:**555–569.
12. **Fields, S., and O. K. Song.** 1989. A novel genetic system to detect protein-protein interactions. Nature (London) **340:**245–246.
13. **Franke, A., M. DeCamillis, D. Zink, N. Cheng, H. W. Brock, and R. Paro.** 1992. *Polycomb* and *polyhomeotic* are constituents of a multimeric protein complex in chromatin of *Drosophila melanogaster*. EMBO J. **11:**2941–2950.
14. **Harper, J. W., G. R. Adami, N. Wei, K. Keyomars, and S. J. Elledge.** 1993. The p21 Cdk-interacting protein Cip1 is a potent inhibitor of G1 cyclin-dependent kinases. Cell **75:**805–816.
15. **Jürgens, G.** 1985. A group of genes controlling the spatial expression of the *bithorax* complex in *Drosophila*. Nature (London) **316:**153–155.
16. **Kennison, J. A.** 1995. The *Polycomb* and *trithorax* group proteins of *Drosophila*: trans-regulators of homeotic gene function. Annu. Rev. Genet. **29:** 289–303.

17. **Kingston, R. E., C. A. Bunker, and A. N. Imbalzano.** 1996. Repression and activation by multiprotein complexes that alter chromatin structure. Genes Dev. **10:**905–920.
18. **Lonie, A., R. D'Andrea, R. Paro, and R. Saint.** 1994. Molecular characterisation of the *Polycomblike* gene of *Drosophila melanogaster*, a *trans*-acting negative regulator of homeotic gene expression. Development **120:**2629–2636.
19. **Müller, J., S. J. Gaunt, and P. A. Lawrence.** 1995. Function of the Polycomb protein is conserved in mice and flies. Development **121:**2847–2852.
20. **Nomura, M., Y. Takihara, and K. Shimada.** 1994. Isolation and characterization of retinoic acid-inducible cDNA clones in F9 cells: one of the early inducible clones encodes a novel protein sharing several highly homologous regions with a *Drosophila Polyhomeotic* protein. Differentiation **57:**39–50.
21. **Orlando, V., and R. Paro.** 1993. Mapping *Polycomb*-repressed domains in the *Bithorax* complex using *in vivo* formaldehyde cross-linked chromatin. Cell **75:**1187–1198.
22. **Paro, R.** 1990. Imprinting a determined state into the chromatin of *Drosophila*. Trends Genet. **6:**416–421.
23. **Paro, R., and D. S. Hogness.** 1991. The *Polycomb* protein shares a homologous domain with a heterochromatin-associated protein of *Drosophila*. Proc. Natl. Acad. Sci. USA **88:**263–267.
24. **Pearce, J. J. H., P. B. Singh, and S. J. Gaunt.** 1992. The mouse has a *Polycomb*-like chromobox gene. Development **114:**921–929.
25. **Rastelli, L., C. S. Chan, and V. Pirotta.** 1993. Related chromosome binding sites for *zeste*, suppressors of *zeste* and *Polycomb* group proteins in *Drosophila* and their dependence on *Enhancer of zeste* function. EMBO J. **12:**1513–1522.
26. **Reijnen, M. J., K. M. Hamer, J. L. den Blaauwen, C. Lambrechts, I. Schoneveld, R. van Driel, and A. P. Otte.** 1995. *Polycomb* and *bmi-1* homologs are expressed in overlapping patterns in *Xenopus* embryos and are able to interact with each other. Mech. Dev. **53:**35–46.
27. **Satijn, D. P. E., M. J. Gunster, J. van der Vlag, K. M. Hamer, W. Schul, M. J. Alkema, A. J. Saurin, P. S. Freemont, R. van Driel, and A. P. Otte.** Submitted for publication.
28. **Schul, W., B. Groenhout, K. Koberna, Y. Takagaki, A. Jenny, E. M. M. Manders, I. Raska, R. van Driel, and L. de Jong.** 1996. The RNA 3′ cleavage factors CstF 64 kDa and CPSF 100 kDa are concentrated in nuclear domains closely associated with coiled bodies and newly synthesized RNA. EMBO J. **15:**2883–2892.
29. **Simon, J., A. Chiang, and W. Bender.** 1992. The different *Polycomb* group genes are required for spatial control of the *abdA* and *AbdB* homeotic gene products. Development **114:**493–505.
30. **Simon, J.** 1995. Locking in stable states of gene expression: transcriptional control during *Drosophila* development. Curr. Opin. Cell Biol. **7:**376–385.
31. **Stuurman, N., A. de Graaf, A. Floore, A. Josso, B. Humbel, L. de Jong, and R. van Driel.** 1992. A monoclonal antibody recognizing nuclear matrix-associated nuclear bodies. J. Cell Sci. **101:**773–784.
32. **van der Lugt, N. M. T., J. Domen, K. Linders, M. van Roon, E. Robanus-Maandag, H. te Riele, M. van der Valk, J. Deschamps, M. Sofroniew, M. van Lohuizen, and A. Berns.** 1994. Posterior transformation, neurological abnormalities, and severe hematopoietic defects in mice with targeted deletion of the *BMI1* proto-oncogene. Genes Dev. **8:**750–769.
33. **van der Lugt, N. M. T., M. Alkema, A. Berns, and J. Deschamps.** 1996. The *Polycomb*-group homolog Bmi-1 is a regulator of murine *Hox* gene expression. Mech. Dev. **58:**153–164.
34. **van Lohuizen, M., S. Verbeek, B. Scheijen, E. Wientjes, H. van der Gulden, and A. Berns.** 1991. Identification of cooperating oncogenes in E(mu)-*myc* transgenic mice by provirus tagging. Cell **65:**737–752.
35. **van Lohuizen, M., M. Frasch, E. Wientjes, and A. Berns.** 1991. Sequence similarity between the mammalian *BMI1* proto-oncogene and the *Drosophila* regulatory genes *Psc* and *Su(z)2*. Nature (London) **353:**353–355.
36. **Wismar, J., T. Loffler, N. Habtermichael, O. Vef, M. Geissen, R. Zirwes, W. Altmeyer, H. Sass, and E. Gateff.** 1995. The *Drosophila melanogaster* tumor suppressor gene *lethal(3)malignant brain tumor* encodes a proline-rich protein with a novel zinc finger. Mech. Dev. **53:**141–154.
37. **Zink, B., and R. Paro.** 1989. *In vivo* binding pattern of a transregulator of homeotic genes in *Drosophila melanogaster*. Nature **337:**468–471.