MCB Accepts, published online ahead of print on 11 October 2010
Mol. Cell. Biol. doi:10.1128/MCB.00465-10

Pygo2 associates with MLL2 histone methyltransferase (HMT) and GCN5 histone acetyltransferase (HAT) complexes to augment Wnt target gene expression and breast cancer stem-like cell expansion

Running title: Pygo2-HMT/HAT interaction in breast cancer cells

Jiakun Chen[1,†], Qicong Luo[1,†] , Yuanyang Yuan[1,2,†], Xiaoli Huang[1], Wangyu Cai[1], Chao Li[1], Tongzhen Wei[1], Ludi Zhang[1], Meng Yang[1], Qingfeng Liu[1], Guodong Ye[1], Xing Dai[2,1,*], and Boan Li[1,*]

[1]Department of Biomedical Sciences and the Key Laboratory of the Ministry of Education for Cell Biology and Tumor Cell Engineering, School of Life Sciences, Xiamen University, Xiamen, Fujian 361005, China
[2]Department of Biological Chemistry, School of Medicine, University of California, Irvine, CA 92697, USA

Word count for the Materials and Methods section: 1290

Word count for the Instruction, Results, and Discussion sections: 4425

* Corresponding author.
Mailing address for Boan Li: Department of Biomedical Sciences, Xiamen University School of Life Sciences, 422 South Siming Rd., Xiamen, 361005, Fujian province, China. Phone: 86-592-2181987. Fax: 86-592-2181984. E-mail: bali@xmu.edu.cn.
Mailing address for Xing Dai: Department of Biological Chemistry, College of Medicine, D250 Med Sci I, University of California, Irvine, CA 92697-1700, USA. Phone: 949-824-3101. Fax: 949-824-2688. E-mail: xdai@uci.edu

† These authors contributed equally.

ABSTRACT

Resent studies have identified Pygopus as a core component of the β-catenin/T-cell factor (TCF)/ lymphoid-enhancing factor 1(LEF) transcriptional activation complex required for the expression of canonical Wg/Wnt target genes in *Drosophila*. However, the biochemical involvement of mammalian Pygopus proteins in β-catenin/TCF/LEF gene activation remains controversial. In this study, we perform a series of molecular/biochemical experiments to demonstrate that Pygo2 associates with histone modifying enzymatic complexes, specifically the MLL2 histone methytransferase (HMT) and STAGA histone acetyltransferase (HAT) complexes, to facilitate their interaction with β-catenin and to augment Wnt1-induced, TCF/LEF-dependent transcriptional activation in breast cancer cells. We identify a critical domain in Pygo2 encompassing the first 47 amino acids that mediates its HMT/HAT interaction. We further demonstrate the importance of this domain in Pygo2's ability to transcriptionally activate both artificial and endogenous Wnt target genes, and to expand breast cancer stem-like cells in culture. This work now links mechanistically Pygo2's role in histone modification to its enhancement of the Wnt-dependent transcriptional program and cancer stem-like cell expansion.

INTRODUCTION

Epigenetic regulation underlies tissue development, homeostasis and tumorigenesis, and includes the modification of the chromatin in transcriptional activation or repression. The basic repeating unit of the chromatin is the nucleosome consisting of 146 bp of DNA wrapped around a histone octamer containing two copies of each histones H2A, H2B, H3 and H4. Methylation and acetylation of lysine (K) residues on histone H3 and H4 tails confer either activating or silencing effects on transcription. Dimethylation (me2) and trimethylation (me3) of H3K4 and acetylation (Ac) of H3K9/K14 are associated with transcriptional activation, while H3K9 and H3K27 methylation is associated with transcriptional repression(1). Histone methylation is catalyzed by histone methyltransferases (HMTs) and reversed by histone demethylases, whereas the steady-state acetylation levels of histone proteins are achieved by the actions of histone acetyltransferases (HATs) and histone deacetylases (HDACs)(2, 3).

In yeast, a multi-subunit complex containing *Drosophila* Trithorax-related protein Set1 has been shown to be responsible for mono-, di-,and trimethylation of histone H3K4(4). In humans, multiple Set1-like HMT complexes with H3K4 HMT activities have been identified (5). Each of these complexes contains the SET domain-containing homologs of yeast Set1, including hSet1 (human Set1), MLL1 (mixed-lineage leukemia 1, also known as MLL, HRX, ALL1, or KMT2A), MLL2 (mixed-lineage leukemia 2, also known as HRX2, KMT2B), MLL3 (mixed-lineage leukemia 3, also known as HALR, KMT2C), and MLL4 (mixed-lineage leukemia 4, also known as ALR, KMT2D) (6-11) , which carry the enzymatic activity for the associated complexes. RbBP5, WDR5, and ASH2L,which are homologs of yeast Set1 complex subunits Swd1, Swd3, and Bre2, respectively, are the core components and shared by various human Set1-like HMT

93 complexes(7, 12).

94 The first histone-specific HAT, histone acetyltransferase A, was isolated from Tetrahymena
95 and has been shown to be the homologue of the yeast GCN5 putative transcriptional adaptator.
96 Thus, it is the first enzyme to link histone acetylation and transcriptional activation(13). Since
97 then, many HATs have been identified. In mammalian cells, two GCN5-containing complexes
98 have been found, the STAGA [SPT3-TAF(II)31-GCN5L acetylase] complex(14) and the TFTC
99 [TATA-binding-protein-free TAF(II)-containing complex] complex(15). Each of the two
100 complexes contains distinct but overlapping subunits(16). GCN5 has been found to have a
101 preference for K9 and K14 on histone H3, while it also acetylates K8 and K16 on histone H4,
102 albeit to a lesser extent(17, 18).

103 The canonical Wnt signaling pathway plays a central role in normal development and
104 tumorigenesis(19-21). A key output of this pathway is the stabilization and accumulation of
105 β-catenin(22). Without stimulation by Wnt ligands, β-catenin is assembled into the destruction
106 complex composed of Axin, GSK3-β, APC and CK1-α, where β-catenin is sequentially
107 phosphorylated and earmarked for ubiquitin-mediated degradation. Stimulation by Wnts leads to
108 the inhibition of phosphorylation and degradation of β-catenin, which then enters the nucleus and
109 binds to a member of the LEF/TCF family of transcription factors to regulate the expression of
110 target genes involved in diverse cellular processes(23, 24). The identification of many of its
111 nuclear interacting partners has significantly added to our understanding of β-catenin function as
112 a transcription regulator. Many of these factors, including SET1-like HMT complexes and
113 GCN5-containing HAT complexes, are involved in chromatin structure and RNA polymerase II
114 (RNA Pol II) regulation (25, 26). Mutations or abnormal expression of Wnt signaling
115 components have been linked to a number of human cancers derived from multiple tissues(21,

27-29). For instance, Wnts can promote tumorigenesis in mammary epithelium and can enhance the self-renewal of mammary stem cells(30, 31). In addition, several studies have implicated the relevance of Wnt signaling in transformation of breast stem/progenitor cells (see review(32)). However, the molecular mechanisms by which Wnt signaling components promote breast cancer tumorigenesis still remain poorly understood.

Drosophila genetics has identified two additional nuclear components of Wnt signaling, Pygopus (Pygo) and Lgs/Bcl9(33-36). Pygo proteins are thought to promote β-catenin/LEF/TCF transcriptional activation through regulation of β-catenin nuclear retention, and/or by binding to β-catenin via the adapter protein BCL9 and recruiting transcriptional activation complexes(reviewed in (37)), (38-41). More recent studies have suggested an H3K4me3-decoding function for mammalian Pygo proteins(42, 43). Furthermore, studies from others and us have shown that these proteins are involved in promoting H3K4me3 and H3K9/K14Ac(41, 42, 44). However, the molecular mechanisms by which Pygo proteins regulate Wnt target gene expression is still an issue of debate that requires further investigation. Additionally, Pygo2 is an excellent entry point to probe into the epigenetic control mechanisms downstream of Wnt signaling that operate in breast cancer cells. In this work, we report the identification of specific Pygo2-interacting HMT and HAT complexes in breast cancer cells, and map the critical domain in Pygo2 that mediates such interactions. We also provide evidence for the functional involvement of these interactions in β-catenin-dependent transcription and breast cancer stem-like cell expansion.

MATERIALS AND METHODS

Cell culture. HEK 293T and human breast cancer cell lines MDA-MB-231, MCF7 and T-47D were maintained in DMEM supplemented with 10% fetal calf serum (Gibco). All cell lines were grown at 37 °C with 5% carbon dioxide.

Antibodies. Rabbit polyclonal anti-Pygo2 antibody was raised against a GST fusion protein containing amino acids 1–114 of hPygo2 and affinity-purified. Mouse monoclonal anti-hPygo2 antibody was purchased from Santa Cruz Biotechnology. ADA3, WDR5, GCN5, SPT3, and TAF5 antibodies were from Abcam. ASH2L, RbBP5, SET1, MLL1, MLL2 and TRRAP antibodies were from Bethyl Lab. Antibodies against β-catenin, H3, H3K4me2, H3K4me3, H3K9me2 and H3K9/K14Ac were purchased from Upstate Biotechnology. Mouse monoclonal anti-HA, -Myc, -Flag and -β-actin antibodies were Sigma products.

Plasmids. Full-length GCN5 and TRRAP expression constructs were obtained by cloning cDNA-derived PCR products into pCMV-MYC and p3Xflag-CMV-10 vectors, respectively. pcDNA3-Flag-RbBP5 was a generous gift from Ge Kai (National Institutes of Health). PCR fragments containing full-length or Δ1-47 of Pygo2 were obtained using clone pENTR221 (invitrogen, Clone ID ISO22981) as a template, and subcloned into the NdeI/XbaI sites of pCMV5-HA. The SV40 nuclear location signal (NLS) was amplified by PCR and inserted in-frame to Pygo2Δ1-47. All constructs were verified by nucleotide sequencing.

RNAi. The siRNA duplexes of Pygo2 (sc-76303), MLL2 (sc75796), ASH2L (sc-43556),

RbBP5 (sc-76373), GCN5 (sc-37946), TRRAP (sc-36746), ADA3 (sc-78466) and a negative control scrambled sequence (sc-37007) were purchased from Santa Cruz Biotechnology. The shRNA(5'-AAAAGGGATTTGGTCCCATGATCTCTTGGATCCAAGAGATCATGGGACCAA ATCCC-3') of Pygo2 and shRNA(5'-AAAAGGATGTGGATACCTCCCAAGTTTGGATCCAA ACTTGGGAGGTATCCACATCC-3') of β-catenin were subcloned into pLV-H1-EF1α-puro RNAi vector (BIOSETTIA). Cells were transfected using Lipofectamine 2000 (Invitrogen) according to manufacturer's protocol.

Nuclear extract preparation. Nuclear extracts were produced as follows: cells were re-suspended in buffer A (10 mM Hepes pH7.9, 10 mM KCl, 0.15% NP-40, 0.1 mM EDTA pH 8.0; 0.1 mM EGTA pH 8.0, 1 mM DTT and protease inhibitors), incubated on ice for 15 min and nuclei collected by centrifugation. Nuclei were re-suspended in buffer B (20 mM Hepes pH 7.9, 400 mM NaCl, 0.5% NP-40, 1 mM EDTA pH 8.0, 1 mM EGTA pH 8.0, 1 mM DTT and protease inhibitors), rotated at 4 ℃ for 1 h and cleared by centrifugation. The supernatants were dialyzed against dialysis buffer (20 mM Hepes pH 7.9, 0.2 mM EDTA pH 8.0, 20% glycerol, 0.1 M KCl, 0.5 mM DTT and protease inhibitors) at 4 ℃ before use.

GST-pull down and HMT/HAT activity assays. Glutathione S-transferase (GST) fusion proteins were expressed in *Escherichia coli* strain BL21. To purify the GST fusion proteins, cells were lysed by sonication in lysis buffer (PBS, 1% Triton X100, 2% β-mercaptoethanol, 0.1mM PMSF), and the resulting lysates were incubated for 1h at 4℃ with glutathione-Sepharose beads. The beads were pelleted by centrifugation and washed with dialysis buffer for subsequent experiments. Nuclear extracts were then incubated with resin-bound proteins by rotating at 4 ℃

for 3 h, washed four times in washing buffer (20 mM Hepes pH 7.9, 0.2 mM EDTA pH 8.0, 20% glycerol, 0.15 M KCl, 0.2% NP-40) and analyzed by Western blotting using appropriate antibodies, or for HMT/HAT enzymatic activities.

In vitro HMT assay was carried out in a final volume of 50 µl. The sample beads were incubated at 30 ℃ for 2 h with 2 µg core histones (Upstate) , 1 µl ^{3}H-SAM (55 Ci/mmol, Perkinelmer) in methylase activity buffer (50 mM Tris pH 8.5, 20 mM KCl, 10 mM $MgCl_2$, 10 mM β -mercaptoethanol, 250 mM sucrose), spotted on to P81 phosphocellulose squares (Upstate), washed 4 X 15 min with 50 mM $NaHCO_3$, pH 9.0, completely dried and read in a scintillation counter. Three independent experiments were performed and each sample was read in triplicate. In complementary experiments, the reactions above were electrophoresed on 15% SDS-PAGE and subjected to fluorography or Western blot analysis. HAT activity assay was carried out according to manufacturer's instructions (Upstate, Cat # 17-289RF). Acetylated histone H3 and H4 peptides (Upstate) were used as positive controls.

Co-immunoprecipitation (IP). Nuclear extracts were prepared as above. Transfected cells were lysed with lysis buffer (20 mM Tris-HCl pH7.5, 150mM NaCl, 1mM EDTA pH8.0, 1mM EGTA pH8.0, 1%Triton) for subsequent Co-IP. Nuclear extracts or lysates were then precleared with protein A/G beads for 1 hour at 4℃ with agitation. Specific or control IgG antibodies were added to the precleared samples and incubated with rotation at 4℃ for 4h or overnight. The immune complexes were captured with 20µl of protein A/G beads at 4℃ for 1 h, washed three times with washing buffer and subjected to SDS-PAGE for subsequent Western blot analysis.

Chromatin immunoprecipitation (ChIP) and real-time PCR. ChIP assay was carried out following the Upstate Biotechnology protocol. Briefly, cells were fixed with 1% paraformaldehyde at room temperature for 10 min, washed, and lysed with SDS lysis buffer (50 mM Tris-HCl, 1% SDS, 10 mM EDTA and protease inhibitors). The lysates were sonicated to reduce DNA lengths to be between 500 and 1000 bp. The soluble fraction was diluted, precleared with salmon sperm DNA/protein A-agarose, then divided into two tubes and incubated with specific antibodies or control IgG. The immune complexes were then precipitated with protein A/G beads and eluted with elution buffer (0.1 M NaHCO3, 1% SDS). The eluted samples were reverse-crosslinked and treated with proteinase K. DNA was purified by phenol/chloroform extraction and dissolved in distilled water. Real-time PCR quantification of ChIP samples were performed in triplicate using THUNDERBIRD SYBR qPCR Mix (TOYOBO) and primers for the c-Myc enhancer (Forward: 5'-GTGAATACACGTTTGCGGGTTAC-3'; Reverse: 5'-CGGTTTTTTT-CACAAGGGTCTCT-3'), Lef1 enhancer(Forward: 5'-TCCTGGATTCCTTCACCAAC-3'; Reverse: 5'-TCAGGCTGCTGAACATTGAA-3').

Quantitative RT-PCR. Total RNAs were isolated from MDA-MB231, T-47D and MCF7 breast cancer cell lines using the Trizol Reagent (Invitrogen). cDNAs were prepared from these RNAs using ReverTra Ace qPCR RT kit from TOYOBO. Real-time PCR experiments were performed as above using the following primers: c-Myc Forward: 5'-CTTCTCTCCGTCCTCG-GATTCT-3'; c-Myc Reverse: 5'-GAAGGTGATCCAGACTCTGACCTT-3'. Pygo2 forward: 5'-GTTTGGGCTGTCCTGAAAGTCTG3'; Pygo2 reverse: 5'-ATAAGGGCGCCGAAAGTTGA-3'. 18S RNA forward: 5'-GCGGCTTAATTTGACTCAACAC -3'; 18S RNA reverse:

5'-GGCCTCACTAAACCATCCAATC-3'. Expression levels of c-Myc and Pygo2 were normalized against 18sRNA levels and the results were calculated by the ΔΔCT method.

Luciferase assay. HEK 293T cells in 24-well plates were transfected at 50% to 60% confluency using a calcium-phosphate method. Both SuperTopFlash(50 ng) and CMV-β-galactosidase (25 ng) reporter plasmids were co-transfected with one or more of the following expression plasmids: human Wnt1 (25 ng), hPygo2 (50-200 ng) , hBCL9 (100 ng). The total amount of plasmid DNA transfected was made equivalent by adding empty vector. Cells were harvested after 24 hours and processed for luciferase and β-galactosidase assays and data were normalized to β-galactosidase levels.

Lentivirus production and infection. The lentivirus vectors directing expression of HA-Pygo2, HA-Pygo2NLSΔ1-47, Pygo2-specific or LacZ-specific shRNA were co-transfected with packaging vectors PHR and pVSVG into 293T cells using lipofectamine 2000 (Invitrogen). 24-48 hours post transfection, viral supernatants were collected. Cells at 50% to 70% confluency were infected with viral supernatants containing 10 ug/ml polybrene for 24 h, after which fresh media was added to the infected cells and selected with puromycin.

FACS analysis. Confluent cells were trypsinized into single cell suspension, washed with FACS buffer (2% FBS in PBS), counted and stained with fluorophore-conjugated antibodies against two human cell surface markers: CD24-FITC and CD44-PECy5 (eBioscience). A total of 10^6 cells in 100 ul FACS buffer were incubated with antibodies for 30 min at 4 ℃. Unbound antibody was washed off and cells were sorted using Beckman EPICS XL.

Mammosphere culture. Mammospheres were cultured using a previously described protocol (45). Briefly, cells (10^4 cells/ml) were cultured in ultra-low attachment plates in serum-free DMEM/F12 (Invitrogen) supplemented with B27 (1:50, Invitrogen), 20 ng/mg EGF (BD Biosciences), 20 ng/ml bFGF (BD Biosciences) and 4 ug/ml insulin (Sigma), and fed every 3 days. Mammospheres were photographed at day 6, 12 and 18, and the mammosphere number and size were measured at day 12.

RESULTS

Pygo2 associates with the MLL2 HMT complex via its N-terminal domain of 47 amino acids (ND1-47). Previously, we reported a likely β-catenin-independent interaction between Pygo2 and a core component of the SET1-like HMT (referred to as HMT from here on) complexes, namely WDR5 (42). To more directly address the β-catenin dependence of Pygo2-HMT interaction, we generated a GST-Pygo2 N-terminal domain (amino acid 1-99, referred to as ND from here on; includes the previously characterized NHD domain; see Fig. 2A) fusion protein and asked if it pulls down HMT components, reasoning that any interaction via β-catenin would be eliminated when the C-terminal PHD domain is absent. GST pull-down experiments using nuclear extracts from MDA-MB231 cells showed that Pygo2 ND indeed interacted with WDR5, as well as RbBP5 and Ash2L, two other common core components of HMT (Fig.1A). To identify the specific HMT enzyme with which Pygo2 associates, we blotted the pull-down samples with antibodies against SET-1, MLL1, and MLL2. Interaction with Pygo2 ND was detected for MLL2, but not SET1 or MLL1 (Fig. 1B). Consistently, we also detected

menin, a MLL1/MLL2-specific component, in the Pygo2 ND pull-down samples. Results of co-IP experiments confirmed the association between endogenous Pygo2 and the MLL2 HMT complex in MDA-MB231 cells: ~6%, 4%, 3%, and 8% of RbBP5, Ash2L, WDR5, and MLL2, respectively, were found associated with Pygo2 in these cells (Fig. 1C). Co-IP signals were dramatically reduced upon the RNAi knockdown of Pygo2 expression, demonstrating that the observed interaction is indeed Pygo2-dependent. Collectively, these results show that Pygo2 associates with the MLL2 HMT complex in breast cancer cells via its N-terminal domain.

It is possible that β-catenin may interact with HMT and simultaneously with another protein complex that is able to interact with the N-terminus of Pygo2. Alternatively, the N-terminal domain of Pygo2 might be able to dimerize, in which case the minimal domain ND1-47 could bind to full length Pygo2, which in turn interacts with HMT/HAT via β-catenin. To address these possibilities, we repeated the GST pull-down experiment and probed with anti-Pygo2 and anti-β-catenin antibodies. As shown in Fig. 1D, neither GST-ND nor GST-PHD was able to pull down full-length Pygo2, arguing against the possibility of dimerization via either the ND or PHD domain. Moreover, under conditions when an association between PHD and β-catenin was observed, no interaction was detected between ND and β-catenin, arguing against the possibility that the observed ND-HMT/HAT interactions were mediated by β-catenin.

To map the HMT-interacting domain in Pygo2 N-terminal domain, we generated a series of deletions and point mutations including alterations in the previously identified NPF motif required for transcriptional activation (46)(Fig 2A). Interestingly, deletion of, or mutations in the NPF motif did not affect the binding between Pygo2 and the MLL2 HMT complex (Fig 2B，left), indicating a differential involvement of NPF in HMT binding and transcriptional activation. Moreover, deletion of the last 43 (GST-ND1-56) or 52 (GST-ND1-47) amino acids had no

294 significant effect on Pygo2-MLL2 complex interaction. In contrast, deletion of the first 47
295 (GST-ND△1-47) amino acids resulted in dramatically reduced Pygo2 binding to RbBP5, Ash2L,
296 and MLL2 (Fig. 2B, middle). In additional experiment, we confirmed that both the 5' flanking
297 amino acids of the conserved NHD (GST-ND1-34) and the first 13 amino acids of NHD
298 (GST-ND35-47) showed effective binding, indicative of two independent interacting surfaces
299 within ND (Fig. 2B, right). Therefore, region 1-47, which partially overlaps with the conserved
300 NHD, is both necessary and sufficient for binding to the MLL2 HMT complex.

301 To determine if the Pygo2 ND-associated complex has histone methyltransferase activity,
302 we performed enzymatic assays using tritiated methyl-S-adenosyl-methionine (^{3}H-SAM) and
303 unmethylated histones as substrates. Histones incubated with the GST-ND pull-down sample
304 from MDA-MB231 cells showed an elevated level of ^{3}H-SAM incorporation compared with the
305 GST control (Fig. 3A, left), demonstrating that the Pygo2 ND-immunocomplex indeed can
306 catalyze the transfer of methyl group to histones. When the resulting samples were analyzed
307 using SDS-PAGE followed by fluorography, substantial labeling by the Pygo2
308 ND-immunocomplexes was only evident for H3, but not for the other histones present in the
309 reaction (Fig. 3B, left). Overall, the extent of labeling was comparable to that observed using the
310 GST-VP16 pull-down sample as a positive control (47). While the depletion of β-catenin had no
311 effect (data not shown), the depletion of MLL2 caused a decrease in the Pygo2 ND-associated
312 histone methyltransferase activity (Fig. 3A, middle and right). These results demonstrate that the
313 Pygo2 ND-HMT association is β-catenin-independent, and that the associated HMT is MLL2.

314 Using lysine-specific antibodies, we detected H3K4me2 and H3K4me3, but not H3K9me2,
315 in histones that were incubated with the GST-ND pull-down sample (Fig. 3B, right). In contrast
316 to GST-ND, histones that were incubated with the pull-down sample from GST-ND△1-47

showed no remarkable incorporation of ^{3}H-SAM above baseline level (Fig. 3A, left), nor contained appreciable levels of H3K4me2 and H3K4me3 (Fig. 3B, right). Taken together, these results indicate that the Pygo2 ND-immunocomplex is able to specifically methylate H3K4.

The ND1-47 domain of Pygo2 contributes to its association with H3K4me3. We previously reported the interaction of endogenous Pygo2 with H3K4me3 in normal mammary epithelial cells (42), but wondered whether this also occurs in breast cancer cells. Indeed, both anti-H3K4me2 and anti-H3K4me3 antibodies were able to immunoprecipitate endogenous Pygo2 from MDA MB231 cells, whereas anti-H3K9me2 antibody did not (Fig. 3C). To distinguish binding via ND from that via PHD (42), we turned to transfect the Myc-tagged ND and ND△1-47 constructs into MDA-MB231 cells and examine histone interaction of the exogenous proteins. Because ND△1-47 lacks the nuclear localization signal (NLS) of Pygo2, we fused a SV40 NLS in-frame at its 5' end. Results of indirect immunofluorescence experiments indicated that Myc-NLS-ND△1-47 was localized to the nucleus (Fig. 3Da). Western blotting showed similar levels of Myc-ND and Myc-NLS-ND△1-47 were produced in the cell nuclei (Fig. 3Db). Anti-H3K4me2 and anti-H3K4me3 antibodies immunoprecipitated Myc-ND but not Myc-NLS-ND△1-47 from transfected cells, suggesting that the observed interaction in MDA MB231 cells is at least in part mediated by the first 47 amino acid of Pygo2 ND.

To assess the importance of ND1-47 in H3K4me3 association in the context of full-length protein, we transfected into MDA-MB231 cells an HA-tagged, full-length Pygo2 expression construct and its △1-47 mutant derivative, and performed co-IP assays. Anti-HA antibody efficiently immunoprecipitated H3K4me3 but not total H3 from cells transfected with full-length Pygo2 (Fig.3E). The level of H3K4me3 immunoprecipitated from cells transfected with the △1-47 mutant was dramatically reduced. Next we compared the relative binding strength of ND

vs. PHD to methylated-histone. Higher H3K4me3 levels were detected in GST-ND pull-down samples than in GST-PHD pull-down samples (Fig. 3F). Together, these results suggest that in the full-length Pygo2 protein, both the ND and the PHD domains contribute quantitatively to the ability of Pygo2 to associate with H3K4me3.

Pygo2 associates with the GCN5-containing STAGA HAT complex and this interaction is mediated by its ND1-47 domain. Our previous work has shown that Pygo2 associates with a HAT activity in male germ cells, and that compromised Pygo2 function drastically reduces histone H3 K9/K14 acetylation during spermatogenesis (44). We next asked if Pygo2 also associates with a HAT in breast cancer cells, focusing particularly on GCN5, a HAT that is known to interact with β-catenin (26, 48). As shown in Fig. 4A, endogenous GCN5 and Pygo2 proteins co-immunoprecipitated in MDA-MB231 cells.. Quantitative analysis revealed that ~10% of GCN5 in these cells is Pygo2-bound. When Pygo2 was depleted using RNAi, anti-Pygo2 antibody immunoprecipitated much less GCN5, indicating a Pygo2 dependence. Interaction was also seen in cells transfected with HA- or Flag-tagged Pygo2 ND and Myc-tagged GCN5 (Fig. 4B), indicating that the interaction is mediated at least in part by the ND domain.

Human cells have two distinct GCN5-containing HAT complexes, namely TFTC and STAGA, with TRRAP and SPT3 as common subunits. In GST pull-down experiments, we observed interaction of Pygo2 ND with GCN5, TRRAP and SPT3, but not TAF5, which is a TFTC-specific subunit (Fig. 4C, left). Moreover, the ND1-47 region was both necessary and sufficient for the ND-GCN5/TRRAP interaction (Fig. 4C, middle). Similar to the HMT interaction, both the ND1-34 and the ND35-47 showed effective binding with GCN5/TRRAP (Fig. 4C, right). Finally, we analyzed the ability of the pull-down samples to acetylate core

363 histones in vitro, and found the GST-ND but not the ND△1-47 pull-down to contain high levels
364 of histone H3 HAT activity (Fig. 4D, left). As it was previously shown that CBP, another HAT,
365 interacts with both Pygo2 and β-catenin in SW480 colon cancer cells, the ND-associated HAT
366 activity we observed could arise from CBP. However, we could not detect any Pygo2-CBP
367 interaction in both GST pull-down and endogenous co-IP experiments using MDA-MB231 cells
368 when CBP-β–catenin interaction was readily detectable (data not shown), suggesting a cell type
369 difference in Pygo2-HAT interaction. Importantly, when GCN5 expression was depleted, HAT
370 activity of the GST-ND pull-down sample was remarkably reduced (Fig. 4D, middle and right).
371 While we cannot completely exclude the possibility that Pygo2 also associates with other HAT(s),
372 our results suggest that the Pygo2 ND-associated HAT activity is at least in-part
373 GCN5-dependent. Taken together, our findings indicate that Pygo2 specifically associates with
374 the STAGA HAT complex via the first 47 amino acids at its N-terminus.

375 **The HMT/HAT-interacting domain of Pygo2 is important for maximal**
376 **β-catenin-HMT/HAT association and Wnt1-induced SuperTOPflash reporter gene**
377 **activation.** Our discovery of the Pygo2 ND-HMT/HAT interaction as described above, plus the
378 known β-catenin-HMT/HAT interaction (26, 48), led us to hypothesize that Pygo2, by virtue of
379 its ability to bind both β-catenin (via PHD) and HMT and/or HAT complexes (via ND), may
380 provide additional interacting forces to enhance the β-catenin-HMT/HAT association . To test
381 this, we used RNAi knockdown to deplete Pygo2, and asked whether β-catenin interaction with
382 RbBP5 or GCN5 is affected. Co-IP experiments using β-catenin antibody revealed reduced levels
383 of both RbBP5 and GCN5 in Pygo2-depleted MDA-MB231 cells (Fig. 5A). Moreover,
384 anti-β-catenin antibody immunoprecipitated much less RbBP5 and GCN5 from MDA-MB231
385 cells infected with HA-Pygo2NLSΔ1-47 lentiviruses compared with those from cells infected

386 with HA-tagged Pygo2 lentiviruses (Fig. 5B). These results demonstrate that maximal β-catenin
387 association with RbBP5 or GCN5 requires Pygo2, and that this Pygo2 function likely depends on
388 the presence of its first 47 amino acids involved in HMT/HAT interaction. .

389 We next asked whether the RbBP5/GCN5-mediated Wnt pathway output is dependent on
390 Pygo2 using the widely used SuperTOPflash reporter system in 293T cells. Consistent with
391 previous reports (46, 49), we did not observe any enhancement of SuperTOPflash activity by
392 Pygo2 even when cells were stimulated with LiCl or overexpressing β-catenin (data not shown).
393 Interestingly, when cells were co-transfected with mouse Wnt1, overexpression of Pygo2 now
394 significantly enhanced SuperTOPflash activity in a dose-dependent manner (Fig. 5C, 5D). This
395 effect appeared to be further enhanced by the co-transfection of human BCL9-2, which when
396 added alone had no effect (Fig. 5C). In contrast to wild-type Pygo2, Pygo2NLSΔ1-47 failed to
397 elevate Wnt1-stimulated SuperTOPflash activity (Fig. 5D). These results underscore the
398 importance of the HMT/HAT-interacting domain of Pygo2 in Wnt1-stimulated target gene
399 transcription.

400 **The HMT/HAT interaction of Pygo2 is important for the expression of, and HMT/HAT**
401 **binding to, endogenous Wnt target genes c-Myc and Lef1 in breast cancer cells.** To elucidate
402 the biological relevance of the Pygo2-HMT/HAT interaction, we next asked whether this
403 interaction is required for Pygo2's role in expression of an endogenous Wnt target, c-Myc (50).
404 Depletion of Pygo2 using siRNA knockdown resulted in significantly reduced levels of c-Myc
405 mRNAs in MDA-MB231, T-47D and MCF7 breast cancer cell lines as revealed by quantitative
406 RT-PCR experiments (Fig. 6A). Knockdown of several components of the MLL2 HMT or
407 STAGA HAT complexes also led to a significant reduction in c-Myc mRNA levels in
408 MDA-MB231 cells (Fig. 6B). In contrast, the expression of GAPDH, a RNA pol II-dependent

and Wnt-independent gene, was unaffected by Pygo2, MLL2 and GCN5 depletion (data not shown). The consistent style of c-Myc regulation by Pygo2 and components of the MLL2 HMT/STAGA HAT complexes implies a functional relevance of their physical association. Moreover, over-expression of full-length Pygo2 but not its Δ1-47 mutant derivative resulted in elevated c-Myc mRNA levels (Fig. 6C), providing direct evidence that the MLL2 HMT/STAGA HAT-interacting domain is required for Pygo2-augmented c-Myc expression in breast cancer cells.

To address whether Pygo2 facilitates the binding of MLL2 HMT and STAGA HAT complexes to target chromatin at the c-Myc loci, we performed ChIP assays using MDA-MB231 cells. Co-occupancy of the c-Myc enhancer, but not a control upstream region, by Pygo2(42), RbBP5, MLL2, GCN5, TRRAP and ADA3 was observed (Fig. 6D). Importantly, siRNA depletion of Pygo2 led to reduced occupancy by all these proteins (Fig. 6E). In keeping with the activities of MLL2 HMT and STAGA HAT complexes in histone H3 trimethylation and acetylation, respectively, we observed significantly reduced levels of H3K4me3 and H3K9/K14Ac at the c-Myc enhancer, whereas the level of total histone H3 was unaffected (Fig. 6F). Similar results were achieved on another Wnt target gene, Lef1(data not shown), indicating the general applicability of the c-Myc findings. Together, these results support a model that Pygo2 uses its ND to recruit HMT/HAT to Wnt target genes to activate their transcription in breast cancer cells.

Pygo2 expands breast cancer stem-like cells and this function requires its HMT/HAT-interacting domain. Although we have previously reported that Pygo2 facilitates the expansion of normal mammary stem/progenitor cells (42), its possible involvement in the expansion of breast cancer cells, particularly the so-called cancer stem-like or cancer-initiating

cells has not been addressed. Such cells can be enriched from established cancer cell lines using either mammosphere culture or FACS sorting for a $CD44^+CD24^-$ population (51, 52). Using quantitative RT-PCR, we detected higher levels of Pygo2 mRNAs in mammospheres of MDA-MB231, T-47D and MCF7 cell lines compared with their corresponding adherent cultures (Fig. 7A). Knockdown of Pygo2 expression in MDA-MB231 cells using lentivirally expressed siRNA resulted in reduced number as well as size of mammosphere (Fig 7B). Conversely, increased mammosphere number and size were observed from MDA-MB231 cells infected with HA-Pygo2-expressing lentiviruses as compared to those infected with GFP-expressing lentiviruses, or HA-Pygo2Δ1-47-expressing lentiviruses (Fig.7C). We also examined whether Pygo2 affects the size of the $CD44^+CD24^-$ population in breast cancer cells. T-47D cells were used for this analysis as MDA-MB-231 cells contain too high a percentage (~90%) of such cells, making it difficult to score any potential increase. As shown in Fig. 7D, knockdown of Pygo2 expression in T-47D cells resulted in a smaller $CD44^+CD24^-$ pool. On the other hand, enforced overexpression of HA-Pygo2 in these cells vial lentiviral infection yielded an increased $CD44^+CD24^-$ population, whereas overexpression of HA-Pygo2Δ1-47 failed to do so (Fig.7E). Similarly, depletion of MLL2 or GCN5 expression in T-47D cells also resulted in a reduced $CD44^+CD24^-$ pool (Fig. 7F), and the overexpressed HA-Pygo2 was no longer able to induced an increase in the $CD44^+CD24^-$ population in these MLL2 or GCN5-depleted cells (Fig.7G).

Next we tested whether the Pygo2 ND is also required in MCF10A cells, an in vitro model of normal mammary stem/progenitor cells. As expected, overexpression of HA-Pygo2 resulted in increased size and number of colonies formed by MCF10A cells plated at a clonal density. In contrast, overexpression of HA-Pygo2Δ1-47 did not. Moreover, overexpression of HA-Pygo2 led to an increased $CD44^+CD24^-$ population, whereas overexpression of HA-Pygo2Δ1-47 did not

(data not shown). Collectively, these results demonstrate that Pygo2 expands breast cancer stem-like cells as well as normal mammary stem/progenitor cells in a manner that is dependent on its HMT/HAT-interacting domain.

DISCUSSION

Building upon our previously reported findings that Pygo2 interacts with HAT in male germ cells and HMT in normal mammary progenitor cells, here we provide biochemical evidence of Pygo2-HMT and Pygo2-HAT interactions in human breast cancer cell lines. Moreover, we have identified the specific HMT and HAT enzymes with which Pygo2 associates. Finally, our data point to a critical involvement of the first 47 amino acids at the N-terminus of Pygo2, encompassing 1-13 amino acids of the previously identified NHD domain, in HMT/HAT interaction, Wnt1-induced transcriptional activation, Wnt target expression, and breast cancer stem-like cell expansion, suggesting a causal link between these molecular and cellular processes.

1. Pygo2, a multifaceted nuclear protein that facilitates β-catenin-HMT/HAT interaction

So far, studies of the *Drosophila* Pygo protein suggest two modes of action: 1) the NHD domain of Pygo is brought to LEF/TCF cognate DNA sites via a linear 'chain of adaptors' comprised of Legless/BCL9 and β-catenin to activate transcription(39); 2) Pygo anchors β-catenin in the nucleus(53). We found that human Pygo2 interacts with HMT/HAT and facilitates β-catenin-HMT/HAT association via its NHD-containing N-terminal domain. This is in keeping with a previous speculation that the Pygo NHD domain might help capture transactivating complexes that bind to β- catenin (54). Given its multiple ND- (this work, and (38, 40, 41)) and PHD-mediated interactions (42, 43) (Fig. 8), Pygo2 has the potential to act as a

scaffolding protein to bring together β-catenin, HMT, HAT and the chromatin. Supporting this view is our finding that Pygo2 facilitates β-catenin-HMT/HAT interaction and transcriptional activation. The context-dependent nature of how a scaffold protein functions would help to reconcile the seemingly contradicting findings in the Pygopus literature, including whether or not Pygo proteins possess transcription-activating ability (39, 46, 49, 54-56) and whether Pygo functions are Wnt-dependent or independent (reviewed by (37)).

Another plausible alternative is that Pygo2 does not act as an active scaffold, but instead its N-terminal domain-mediated interactions with multiple coactivator complexes are permissive to its positive role in Wnt-dependent transcription. In this context, we note that the Pygo2 ND contains several runs of proline and glycine as well as positively charged amino acids that are prone to protein-protein interactions(57-59) . Consistent with the presence of such potential protein interaction motifs, both regions 1-34 (which is not conserved in non-mammalian Pygo homologs) and 35-47 (which is evolutionarily conserved) are required for Pygo2 association with HMT and HAT complexes (Figures 2 and 4). In interesting contrast, the evolutionarily conserved NPFxD motif (76-80 in Pygo2), which in *Drosophila* Pygopus is absolutely essential for its function during fly development, is not required for Pygo2's ability to interact with MLL2 HMT, GCN5 HAT (this work), CBP HAT(41), and TAF4 complexes(38) . Our data support the model that Pygo2 provides a β-catenin-independent recruiting platform for HMT/HAT. This said, there are likely cross-talks between the two parallel HMT/HAT recruiting pathways (Pygo2 and β-catenin) that result in recruitment synergy at Wnt target genes. Supporting this notion is our finding that Pygo2 facilitates β-catenin-HMT/HAT interaction. Whether or not β-catenin facilitates Pygo2-HMT/HAT interaction remains to be determined. Moreover, it will be interesting to experimentally assess the relative contribution of Pygo2 and β-catenin to

HMT/HAT recruitment and target gene activation. The data presented here do not provide insight into whether Pygo2 makes direct contact with any of the identified components of the MLL2 HAT and GCN5/HAT complexes. Future work to address this will help improve our understanding of how transcriptional activator/chromatin modifying complexes are assembled at the Wnt target genes.

Previous studies have provided strong evidence for the PHD domain of mammalian Pygo proteins to directly bind H3K4me2/3(42, 43), Since Pygo2 ND does not contain any known chromatin reader domains, our result showing Pygo2 ND binding to H3K4me2/3 is somewhat surprising. This is likely an indirect effect mediated by Pygo2 ND's association with the HMT complexes. WDR5, an obligatory component of HMT complexes, is known to directly associate with H3K4me2 and H3K4me3(60), could potentially serve as a bridge for the Pygo2 ND-H3K4me2/3 association. Overall, the ND of Pygo2 allows it to engage in multiple interactions with HMT/HAT as well as histone substrates. While these interactions might be low-affinity in nature, they collectively and together with the PHD-mediated interactions could exert a strong promoting force in transcriptional activation.

2. Transcriptional activation by Pygo2 requires the Wnt1 ligand

Despite extensive efforts, we failed to detect any enhancement of SuperTOPflash activity by exogenous Pygo2 in multiple mammalian cell lines including 293T and MCF10A even when cells were either stimulated by LiCl, inhibitor of GSK3β, or transfected with a β-catenin expression construct. A priori, this could be due to a number of factors including (but not limited to): 1) Pygo2 activity is promoter dependent; 2) transiently transfected reporter construct is assembled into an imperfect chromatin structure and its activity does not require chromatin-activating events such as Pygo2 function; 3) Pygo2 protein or its function requires

modification and/or other factors induced by additional signaling events. Our data that Pygo2 enhances SuperTOPflash activity in cells transfected with a Wnt1-expression vector supports the last notion. Moreover, these results suggest that the putative modification of Pygo2 function likely occurs upstream of GSK3β and β-catenin. In the future, it will be interesting to investigate the molecular mechanism of this Wnt1-mediated functional modification.

3. Pygo2 in the breast cancer stem-like cells

The recent advances in characterization of stem cells in mammary epithelium and breast cancer cells have opened the door for understanding the epigenetic and transcriptional program underlying both the development/homeostasis of normal mammary stem cells and proliferation/differentiation of their malignant counterparts(51, 52, 61, 62). We have previously uncovered a role for Pygo2 in normal mammary progenitors cells through the promotion of Wnt signaling(42). Perhaps not surprisingly, here we found Pygo2 to be enriched in breast cancer stem-like cells. Moreover, depletion of Pygo2 appeared to decrease the overall size of this population, as assayed by mammasphere formation and FACS profiling using known surface markers. Therefore, Pygo2 plays an important role in both normal and malignant mammary stem-like cells.

At a molecular level, Pygo2's function in breast cancer stem-like cells ties with its ability to interact with the MLL2-HMT and STAGA-HAT complexes. This model is strongly supported by 1) the converging effects of Pygo2 and components of these complexes on c-Myc and Lef1 transcription, 2) the dependence of chromatin binding activity of these complexes on Pygo2, and 3) the specific requirement of ND1-47 for Pygo2's role in breast cancer stem-like cell expansion. As such, our study presents the first mechanistic characterization of Pygo2 function in breast cancer stem-like cells, and points to Pygo2 as an important epigenetic and transcriptional

regulator in these cells. As breast cancer stem-like cells are thought to be fundamental contributors of tumor initiation as well as recurrence, our study paves the way for developing future cancer therapeutics to target this cell population. Future work should address whether Pygo2 is also involved in the regulation of other cancer cell types in a similar manner.

ACKNOWLEDGEMENTS

We thank Dr. Randall Moon (University of Washington) for SuperTOPflash and SuperFOPflash reporter constructs, and Dr. Ge Kai (National Institutes of Health) for Flag-RbBP5 construct, and anonymous reviewers for many valuable insights and suggestions. This work was supported by grants from the National Natural Science Foundation of China (30871279, 90919037, and 30728005), the "973" Project of the Ministry of Science and Technology (2009CB52220), the Natural Science Foundation of Fujian Province (2008J0007), the Science Planning Program of Fujian Province (2009J1010), the "Project 111" sponsored by the State Bureau of Foreign Experts and Ministry of Education (B06016) (to B. L.), DOD grant W81XWH-04-1-0516 and NIH Grants K02-AR51482 and R01-GM083089 (to X.D.).

REFERANCES

1. **Sims, R. J., 3rd, K. Nishioka, and D. Reinberg.** 2003. Histone lysine methylation: a signature for chromatin function. Trends Genet **19:**629-39.
2. **Cheung, W. L., S. D. Briggs, and C. D. Allis.** 2000. Acetylation and chromosomal functions. Curr Opin Cell Biol **12:**326-33.
3. **Narlikar, G. J., H. Y. Fan, and R. E. Kingston.** 2002. Cooperation between complexes that regulate chromatin structure and transcription. Cell **108:**475-87.
4. **Briggs, S. D., M. Bryk, B. D. Strahl, W. L. Cheung, J. K. Davie, S. Y. Dent, F. Winston, and C. D. Allis.** 2001. Histone H3 lysine 4 methylation is mediated by Set1 and required for cell growth and rDNA silencing in Saccharomyces cerevisiae. Genes Dev **15:**3286-95.
5. **Crawford, B. D., and J. L. Hess.** 2006. MLL core components give the green light to histone methylation.

ACS Chem Biol **1:**495-8.

6. **Milne, T. A., S. D. Briggs, H. W. Brock, M. E. Martin, D. Gibbs, C. D. Allis, and J. L. Hess.** 2002. MLL targets SET domain methyltransferase activity to Hox gene promoters. Mol Cell **10:**1107-17.

7. **Hughes, C. M., O. Rozenblatt-Rosen, T. A. Milne, T. D. Copeland, S. S. Levine, J. C. Lee, D. N. Hayes, K. S. Shanmugam, A. Bhattacharjee, C. A. Biondi, G. F. Kay, N. K. Hayward, J. L. Hess, and M. Meyerson.** 2004. Menin associates with a trithorax family histone methyltransferase complex and with the hoxc8 locus. Mol Cell **13:**587-97.

8. **Glaser, S., J. Schaft, S. Lubitz, K. Vintersten, F. van der Hoeven, K. R. Tufteland, R. Aasland, K. Anastassiadis, S. L. Ang, and A. F. Stewart.** 2006. Multiple epigenetic maintenance factors implicated by the loss of Mll2 in mouse development. Development **133:**1423-32.

9. **Prasad, R., A. B. Zhadanov, Y. Sedkov, F. Bullrich, T. Druck, R. Rallapalli, T. Yano, H. Alder, C. M. Croce, K. Huebner, A. Mazo, and E. Canaani.** 1997. Structure and expression pattern of human ALR, a novel gene with strong homology to ALL-1 involved in acute leukemia and to Drosophila trithorax. Oncogene **15:**549-60.

10. **Lee, S., D. K. Lee, Y. Dou, J. Lee, B. Lee, E. Kwak, Y. Y. Kong, S. K. Lee, R. G. Roeder, and J. W. Lee.** 2006. Coactivator as a target gene specificity determinant for histone H3 lysine 4 methyltransferases. Proc Natl Acad Sci U S A **103:**15392-7.

11. **Allis, C. D., S. L. Berger, J. Cote, S. Dent, T. Jenuwien, T. Kouzarides, L. Pillus, D. Reinberg, Y. Shi, R. Shiekhattar, A. Shilatifard, J. Workman, and Y. Zhang.** 2007. New nomenclature for chromatin-modifying enzymes. Cell **131:**633-6.

12. **Yokoyama, A., Z. Wang, J. Wysocka, M. Sanyal, D. J. Aufiero, I. Kitabayashi, W. Herr, and M. L. Cleary.** 2004. Leukemia proto-oncoprotein MLL forms a SET1-like histone methyltransferase complex with menin to regulate Hox gene expression. Mol Cell Biol **24:**5639-49.

13. **Brownell, J. E., J. Zhou, T. Ranalli, R. Kobayashi, D. G. Edmondson, S. Y. Roth, and C. D. Allis.** 1996. Tetrahymena histone acetyltransferase A: a homolog to yeast Gcn5p linking histone acetylation to gene activation. Cell **84:**843-51.

14. **Martinez, E., V. B. Palhan, A. Tjernberg, E. S. Lymar, A. M. Gamper, T. K. Kundu, B. T. Chait, and R. G. Roeder.** 2001. Human STAGA complex is a chromatin-acetylating transcription coactivator that interacts with pre-mRNA splicing and DNA damage-binding factors in vivo. Mol Cell Biol **21:**6782-95.

15. **Wieczorek, E., M. Brand, X. Jacq, and L. Tora.** 1998. Function of TAF(II)-containing complex without TBP in transcription by RNA polymerase II. Nature **393:**187-91.

16. **Nagy, Z., and L. Tora.** 2007. Distinct GCN5/PCAF-containing complexes function as co-activators and are involved in transcription factor and global histone acetylation. Oncogene **26:**5341-57.

17. **Kuo, M. H., J. E. Brownell, R. E. Sobel, T. A. Ranalli, R. G. Cook, D. G. Edmondson, S. Y. Roth, and C. D. Allis.** 1996. Transcription-linked acetylation by Gcn5p of histones H3 and H4 at specific lysines. Nature **383:**269-72.

18. **Grant, P. A., A. Eberharter, S. John, R. G. Cook, B. M. Turner, and J. L. Workman.** 1999. Expanded lysine acetylation specificity of Gcn5 in native complexes. J Biol Chem **274:**5895-900.

19. **Wodarz, A., and R. Nusse.** 1998. Mechanisms of Wnt signaling in development. Annu Rev Cell Dev Biol **14:**59-88.

20. **Moon, R. T., and J. R. Miller.** 1997. The APC tumor suppressor protein in development and cancer. Trends Genet **13:**256-8.

21. **Polakis, P.** 2000. Wnt signaling and cancer. Genes Dev **14:**1837-51.

22. **Logan, C. Y., and R. Nusse.** 2004. The Wnt signaling pathway in development and disease. Annu Rev Cell Dev Biol **20:**781-810.

23. **Liu, C., Y. Li, M. Semenov, C. Han, G. H. Baeg, Y. Tan, Z. Zhang, X. Lin, and X. He.** 2002. Control of beta-catenin phosphorylation/degradation by a dual-kinase mechanism. Cell **108:**837-47.

24. **Amit, S., A. Hatzubai, Y. Birman, J. S. Andersen, E. Ben-Shushan, M. Mann, Y. Ben-Neriah, and I. Alkalay.** 2002. Axin-mediated CKI phosphorylation of beta-catenin at Ser 45: a molecular switch for the Wnt pathway. Genes Dev **16:**1066-76.

25. **Mosimann, C., G. Hausmann, and K. Basler.** 2009. Beta-catenin hits chromatin: regulation of Wnt target gene activation. Nat Rev Mol Cell Biol **10:**276-86.

26. **Sierra, J., T. Yoshida, C. A. Joazeiro, and K. A. Jones.** 2006. The APC tumor suppressor counteracts beta-catenin activation and H3K4 methylation at Wnt target genes. Genes Dev **20:**586-600.

27. **Brown, A. M.** 2001. Wnt signaling in breast cancer: have we come full circle? Breast Cancer Res **3:**351-5.

28. **Taipale, J., and P. A. Beachy.** 2001. The Hedgehog and Wnt signalling pathways in cancer. Nature

411:349-54.

29. **Lustig, B., and J. Behrens.** 2003. The Wnt signaling pathway and its role in tumor development. J Cancer Res Clin Oncol **129:**199-221.
30. **Lane, T. F., and P. Leder.** 1997. Wnt-10b directs hypermorphic development and transformation in mammary glands of male and female mice. Oncogene **15:**2133-44.
31. **Li, Y., B. Welm, K. Podsypanina, S. Huang, M. Chamorro, X. Zhang, T. Rowlands, M. Egeblad, P. Cowin, Z. Werb, L. K. Tan, J. M. Rosen, and H. E. Varmus.** 2003. Evidence that transgenes encoding components of the Wnt signaling pathway preferentially induce mammary cancers from progenitor cells. Proc Natl Acad Sci U S A **100:**15853-8.
32. **Lindvall, C., W. Bu, B. O. Williams, and Y. Li.** 2007. Wnt signaling, stem cells, and the cellular origin of breast cancer. Stem Cell Rev **3:**157-68.
33. **Belenkaya, T. Y., C. Han, H. J. Standley, X. Lin, D. W. Houston, J. Heasman, and X. Lin.** 2002. pygopus Encodes a nuclear protein essential for wingless/Wnt signaling. Development **129:**4089-101.
34. **Parker, D. S., J. Jemison, and K. M. Cadigan.** 2002. Pygopus, a nuclear PHD-finger protein required for Wingless signaling in Drosophila. Development **129:**2565-76.
35. **Thompson, B., F. Townsley, R. Rosin-Arbesfeld, H. Musisi, and M. Bienz.** 2002. A new nuclear component of the Wnt signalling pathway. Nat Cell Biol **4:**367-73.
36. **Kramps, T., O. Peter, E. Brunner, D. Nellen, B. Froesch, S. Chatterjee, M. Murone, S. Zullig, and K. Basler.** 2002. Wnt/wingless signaling requires BCL9/legless-mediated recruitment of pygopus to the nuclear beta-catenin-TCF complex. Cell **109:**47-60.
37. **Jessen, S., B. Gu, and X. Dai.** 2008. Pygopus and the Wnt signaling pathway: a diverse set of connections. Bioessays **30:**448-56.
38. **Wright, K. J., and R. Tjian.** 2009. Wnt signaling targets ETO coactivation domain of TAF4/TFIID in vivo. Proc Natl Acad Sci U S A **106:**55-60.
39. **Stadeli, R., and K. Basler.** 2005. Dissecting nuclear Wingless signalling: recruitment of the transcriptional co-activator Pygopus by a chain of adaptor proteins. Mech Dev **122:**1171-82.
40. **Carrera, I., F. Janody, N. Leeds, F. Duveau, and J. E. Treisman.** 2008. Pygopus activates Wingless target gene transcription through the mediator complex subunits Med12 and Med13. Proc Natl Acad Sci U S A **105:**6644-9.
41. **Andrews, P. G., Z. He, C. Popadiuk, and K. R. Kao.** 2009. The transcriptional activity of Pygopus is enhanced by its interaction with cAMP-response-element-binding protein (CREB)-binding protein. Biochem J **422:**493-501.
42. **Gu, B., P. Sun, Y. Yuan, R. C. Moraes, A. Li, A. Teng, A. Agrawal, C. Rheaume, V. Bilanchone, J. M. Veltmaat, K. Takemaru, S. Millar, E. Y. Lee, M. T. Lewis, B. Li, and X. Dai.** 2009. Pygo2 expands mammary progenitor cells by facilitating histone H3 K4 methylation. J Cell Biol **185:**811-26.
43. **Fiedler, M., M. J. Sanchez-Barrena, M. Nekrasov, J. Mieszczanek, V. Rybin, J. Muller, P. Evans, and M. Bienz.** 2008. Decoding of methylated histone H3 tail by the Pygo-BCL9 Wnt signaling complex. Mol Cell **30:**507-18.
44. **Nair, M., I. Nagamori, P. Sun, D. P. Mishra, C. Rheaume, B. Li, P. Sassone-Corsi, and X. Dai.** 2008. Nuclear regulator Pygo2 controls spermiogenesis and histone H3 acetylation. Dev Biol **320:**446-55.
45. **Dontu, G., W. M. Abdallah, J. M. Foley, K. W. Jackson, M. F. Clarke, M. J. Kawamura, and M. S. Wicha.** 2003. In vitro propagation and transcriptional profiling of human mammary stem/progenitor cells. Genes Dev **17:**1253-70.
46. **Townsley, F. M., B. Thompson, and M. Bienz.** 2004. Pygopus residues required for its binding to Legless are critical for transcription and development. J Biol Chem **279:**5177-83.
47. **Wysocka, J., M. P. Myers, C. D. Laherty, R. N. Eisenman, and W. Herr.** 2003. Human Sin3 deacetylase and trithorax-related Set1/Ash2 histone H3-K4 methyltransferase are tethered together selectively by the cell-proliferation factor HCF-1. Genes Dev **17:**896-911.
48. **Sustmann, C., H. Flach, H. Ebert, Q. Eastman, and R. Grosschedl.** 2008. Cell-type-specific function of BCL9 involves a transcriptional activation domain that synergizes with beta-catenin. Mol Cell Biol **28:**3526-37.
49. **de la Roche, M., and M. Bienz.** 2007. Wingless-independent association of Pygopus with dTCF target genes. Curr Biol **17:**556-61.
50. **He, T. C., A. B. Sparks, C. Rago, H. Hermeking, L. Zawel, L. T. da Costa, P. J. Morin, B. Vogelstein, and K. W. Kinzler.** 1998. Identification of c-MYC as a target of the APC pathway. Science **281:**1509-12.
51. **Ponti, D., A. Costa, N. Zaffaroni, G. Pratesi, G. Petrangolini, D. Coradini, S. Pilotti, M. A. Pierotti,**

and M. G. Daidone. 2005. Isolation and in vitro propagation of tumorigenic breast cancer cells with stem/progenitor cell properties. Cancer Res **65:**5506-11.

52. **Al-Hajj, M., M. S. Wicha, A. Benito-Hernandez, S. J. Morrison, and M. F. Clarke.** 2003. Prospective identification of tumorigenic breast cancer cells. Proc Natl Acad Sci U S A **100:**3983-8.

53. **Krieghoff, E., J. Behrens, and B. Mayr.** 2006. Nucleo-cytoplasmic distribution of beta-catenin is regulated by retention. J Cell Sci **119:**1453-63.

54. **Kessler, R., G. Hausmann, and K. Basler.** 2009. The PHD domain is required to link Drosophila Pygopus to Legless/beta-catenin and not to histone H3. Mech Dev **126:**752-9.

55. **Hoffmans, R., R. Stadeli, and K. Basler.** 2005. Pygopus and legless provide essential transcriptional coactivator functions to armadillo/beta-catenin. Curr Biol **15:**1207-11.

56. **Hoffmans, R., and K. Basler.** 2004. Identification and in vivo role of the Armadillo-Legless interaction. Development **131:**4393-400.

57. **Van Dusen, C. M., L. Yee, L. M. McNally, and M. T. McNally.** A glycine-rich domain of hnRNP H/F promotes nucleocytoplasmic shuttling and nuclear import through an interaction with transportin 1. Mol Cell Biol **30:**2552-62.

58. **Kay, B. K., M. P. Williamson, and M. Sudol.** 2000. The importance of being proline: the interaction of proline-rich motifs in signaling proteins with their cognate domains. Faseb J **14:**231-41.

59. **Balendra, S., C. Lesieur, T. J. Smith, and H. Dalton.** 2002. Positively charged amino acids are essential for electron transfer and protein-protein interactions in the soluble methane monooxygenase complex from Methylococcus capsulatus (Bath). Biochemistry **41:**2571-9.

60. **Wysocka, J., T. Swigut, T. A. Milne, Y. Dou, X. Zhang, A. L. Burlingame, R. G. Roeder, A. H. Brivanlou, and C. D. Allis.** 2005. WDR5 associates with histone H3 methylated at K4 and is essential for H3 K4 methylation and vertebrate development. Cell **121:**859-72.

61. **Shackleton, M., F. Vaillant, K. J. Simpson, J. Stingl, G. K. Smyth, M. L. Asselin-Labat, L. Wu, G. J. Lindeman, and J. E. Visvader.** 2006. Generation of a functional mammary gland from a single stem cell. Nature **439:**84-8.

62. **Stingl, J., P. Eirew, I. Ricketson, M. Shackleton, F. Vaillant, D. Choi, H. I. Li, and C. J. Eaves.** 2006. Purification and unique properties of mammary epithelial stem cells. Nature **439:**993-7.

FIGURE LEGENDS

Figure 1. Pygo2 interaction with the MLL2-HMT complex via its N-terminal Domain. (A-B) GST pull-down of nuclear extracts from MDA-MB231 cells showing Pygo2 ND interaction with the common core components RbBP5, Ashl2, WDR5 (A), and with MLL2, but not SET1 or MLL1 (B). Equal amounts of GST, as indicated by Commassie blue staining at the bottom, were used as a negative control. (C) Co-IP experiments detecting interaction between endogenous Pygo2 and SET1-like HMT core-components and the MLL2 enzyme in MDA-MB231 cells. Note knockdown of Pygo2 resulted in reduced Pygo2-HMT association. Input and IP samples were run on the same SDS-PAGE gel. (D) Pygo2 ND does not dimerize with full-length Pygo2 or associate with β-catenin. Note: for Ash2L, in addition to the major protein band, a smaller isoform is also recognized; for MLL1, the cleaved C-terminal fragment is recognized and the

lower band should be caused by degradation, for the density varies with different experiments.

Figure 2. Domain 1-47 is required for Pygo2-HMT interaction. (A) Schematic diagram of deletions and mutations of Pygo2 ND. (B) Results of GST pull-down of RbBP5, Ash2L, and MLL2 using mutant constructs in (A). Equal amounts of wild-type and mutant recombinant proteins were used, as visualized by Commassie blue staining at the bottom.

Figure 3. The Pygo2 ND1-47-associated complex possesses histone H3 methylating activity. (A) Histone methylation assay using Pygo2 ND pull-down samples from MDA-MB231 nuclear extract. Left, quantification of ^{3}H-SAM incorporation into unmethylated histone H3 substrates catalyzed by pull-down fractions using GST control, GST-ND, and GST-NDΔ1-47. Two independent experiments were performed, each containing triplicate samples. Shown are results of one experiment. Middle and right panels, ^{3}H-SAM incorporation assays of GST-ND upon MLL2 depletion. Experimental procedure was similar to that described in left. (B) Pygo2 ND-immunocomplex is able to specifically methylate H3K4. Left, fluorography of histones after incubation with different pull-down fractions, indicating that the methyltransferase activity was directed specifically at histone H3. GST-VP16 was used as a positive control. Commassie staining results (middle, bottom) show that equal amounts of histone substrates and GST recombinant proteins were used. Right, Western blot of histones incubated with the pull-down samples were probed with pan H3 and lysine-specific antibodies. The H3K4me2/Commassie blue or H3K4me3/Commassie blue ratio in GST-ND group was arbitrarily set to be 1. (C) Co-IP experiment revealing association of endogenous nuclear Pygo2 with H3K4me2 and H3K4me3, but not H3K9me2 in MBA-MB231 cells. (D) a, Immunofluorescence staining indicated that

751 Myc-NLS-ND△1-47 was localized to the nucleus. b, top, co-IP experiment showing that the
752 association of overexpressed, exogenous Pygo2-ND with H3K4me2/3 in MBA-MB231 cells
753 depends on the presence of domain 1-47. H3K9me2 was used as a negative control. Bottom,
754 Western blotting showing that similar levels of Myc-ND and Myc-NLS-ND□1-47 were detected
755 in nuclear extracts of transfected MDA-MB231 cells. (E) Domain 1-47 is also important for the
756 association of full-length Pygo2 protein with H3K4me3 but not H3. MDA-MB231 cells were
757 infected with lentiviruses expressing HA-Pygo2 or HA-Pygo2NLSΔ1-47. Input panels showed
758 the comparable expression levels between HA-Pygo2 and HA-NLS-Pygo2□1-47 in nuclear
759 extracts. Note the total cellular levels of H3K4me3 were not affected. (F) Contribution of ND vs.
760 PHD domain to methylated-histone binding.
761
762 **Figure 4. Domain 1-47 mediates Pygo2 interaction with the GCN5 HAT complex.** (A) Co-IP
763 experiments detecting interaction between endogenous Pygo2 and GCN5 in MDA-MB231 cells.
764 Note that knockdown of Pygo2 resulted in reduced association between endogenous Pygo2 and
765 GCN5. Input and IP samples were run on the same SDS-PAGE gel. (B) Co-IP experiment
766 between Pygo2 ND and exogenous GCN5 in 293T cells. (C) Pygo2 specifically associates with
767 the STAGA HAT complex and ND1-47 is the critical binding domain. Left, GST pull-down of
768 nuclear extracts from MDA-MB231 cells showing Pygo2 interaction with GCN5, TRRAP and
769 SPT3, but not TAF5. Middle and right, GST pull-down experiment indicating the importance of
770 ND1-47 for this association. (D) Left, comparative HAT activity analysis of ND and NDΔ1-47
771 pull-down fractions. Middle and right, GST-ND-associated HAT activity upon GCN5 depletion.
772 Values obtained without the presence of any pull-down fractions (negative control) were
773 arbitrarily set as 1.

774

775 **Figure 5. Pygo2 facilitates β-catenin interaction with RbBP5 or GCN5.** (A) Co-IP of
776 endogenous β-catenin and RbBP5/GCN5 in control and Pygo2-depleted cells. MDA-MB231
777 cells were infected with lentiviruses expressing scrambled (control, -) or Pygo2 shRNA (+). (B)
778 The first 1-47 amino acids are important for Pygo2 in facilitating β-catenin-RbBP5/GCN5
779 interaction. Shown are results of co-IP experiments of endogenous β-catenin and RbBP5 or
780 GCN5 in the presence of Pygo2 or Pygo2NLSΔ1-47. MDA-MB231 cells were infected with
781 lentiviruses expressing HA-Pygo2 or HA-Pygo2NLSΔ1-47. (C) Pygo2 enhances SuperTOPflash
782 activity in the presence of Wnt1. 293T cells were transiently transfected with the indicated
783 plasmids. SuperFOPflash where the Wnt-responsive elements are mutated was used as a control.
784 (D) Pygo2 augments Wnt1-stimulated transcription in a dose- and ND1-47-dependent manner.
785 Western blot shows comparable expression levels for exogenously introduced HA-Pygo2 and
786 HA- Pygo2NLSΔ1-47.

787

788 **Figure 6. Pygo2 recruits HMT/HAT to the c-Myc enhancer and activates c-Myc gene**
789 **expression.** (A) Reduced c-Myc expression in MDA-MB231, T-47D and MCF7 breast cancer
790 cell lines upon Pygo2 depletion. Total mRNAs were extracted 72 h after transfection of
791 Pygo2-specific siRNA or scrambled control. 18S RNA and c-Myc transcript levels were
792 quantified by real-time PCR. Normalized values were calculated as percentages of transcript
793 levels detected in cells treated with the scrambled control. (B) siRNA knockdown of
794 MLL2-HMT (MLL2, Ash2L and RbBP5) and STAGA-HAT (GCN5, TRRAP and ADA3)
795 components in MDA-MB231 cells leads to reduced c-Myc mRNA levels. Experimental
796 procedure was similar to that described in (A). (C) Overexpression of full-length Pygo2 but not

797 the Δ1-47 mutant results in elevated c-Myc mRNA levels. MDA-MB231 cells were infected with
798 lentiviruses expressing HA-Pygo2 or HA-Pygo2NLSΔ1-47 at comparable levels (right). The
799 c-Myc transcript levels were quantified by real-time PCR. (D) Co-occupancy of the c-Myc
800 enhancer by Pygo2 , RbBP5, MLL2, GCN5, TRRAP and ADA3. Chromatin fragments from
801 MDA-MB231 cells were immunoprecipitated with anti-Pygo2 , RbBP5, MLL2, GCN5, TRRAP
802 or ADA3 antibody. The immunoprecipitated DNA was analyzed by PCR, using locus-specific
803 primers. Primers amplifying a region that is 10 kb upstream of the c-Myc enhancer were used as
804 a negative control. (E) Pygo2 knockdown results in decreased occupancy of the c-Myc enhancer
805 by the HMT/HAT components tested. Chromatin fragments from MDA-MB231 cells infected
806 with lentiviruses expressing LacZ (control) or Pygo2 shRNA were immunoprecipitated with the
807 indicated antibodies. DNA association was determined by real-time PCR. Western blotting
808 indicated efficient knockdown of Pygo2 expression in MDA-MB231 cells. (F) Pygo2
809 knockdown results in significantly decreased levels of H3K4me3 and acetyl-H3K9/K14 at the
810 c-Myc enhancer. Note that the total H3 levels at these loci were not significantly affected.
811 Experimental procedure was similar to that described in (E). Error bar in all panels represents
812 standard deviation.
813 .
814 **Figure 7. Pygo2 expands breast cancer stem-like cells in a ND1-47-dependent manner.** (A)
815 Pygo2 expression is enriched in breast cancer stem-like cells. Shown are results of real-time PCR
816 analysis of Pygo2 transcripts in mammospheres or corresponding adherent cells of different
817 breast cancer cell lines (B) Pygo2 is required for optimal mammosphere formation by breast
818 cancer cells. MDA-MB231 cells were infected with lentiviruses expressing LacZ (control) or
819 Pygo2 shRNA and then cultured in sphere media. Left, representative photographs of

mammospheres taken at days indicated. Upper right, sphere number was quantified at day 12 in eight independent experiments. *, p=3.91x10^{-6}. Lower right, sphere size was quantified at day 12 by measuring and averaging the diameter of all spheres in one plate. Two independent experiments were performed, and shown are results of one experiment. **, p=1.56x10^{-12}, n=64 and 39 for control and Pygo2-deficient samples, respectively. (C) Pygo2's function in regulation of mammosphere formation is dependent on ND1-47. MDA-MB231 cells were infected with lentiviruses expressing GFP, HA-Pygo2 or HA-Pygo2NLSΔ1-47. Experimental procedure was similar to that described in (B). *, p=1.043x10^{-7}. **, p=6.38x10^{-18}, n=67, 87 and 68, respectively. (D) Pygo2 depletion in T-47D cells resulted in a smaller CD44^{+}CD24^{-} population. T-47D cells were infected with lentiviruses expressing LacZ (control) or Pygo2 shRNA and then the percentage of CD44^{+}CD24^{-} cells was determined by FACS analysis. Representative FACS profiles from a single pair are shown on the left, and mean values from three different pairs are shown on the right. (E) ND1-47 is important for Pygo2 in expanding the CD44^{+}CD24^{-} population. T-47D cells were infected with lentiviruses expressing control vector, HA-Pygo2 or HA-Pygo2NLSΔ1-47. Experimental procedure was similar to that described in (D). Note that WT and mutant Pygo2 proteins were expressed at comparable levels (data not shown). (F) Knockdown of MLL2 or GCN5 in T-47D cells also resulted in a reduced CD44^{+}CD24^{-} pool. Representative FACS profiles from a single pair are shown at the top; mean values from three different pairs are shown on lower left; and Western blot of knockdown samples are shown on lower right. (G) The positive effect of HA-Pygo2 on the size of the CD44^{+}CD24^{-} population in T47D cells requires MLL2 and GCN5. T47D cells were infected with HA-Pygo2 lentiviruses and treated by MLL2 or GCN5 siRNA simultaneously. Representative FACS profiles from a single pair are shown at the top; mean values from three different pairs are shown on lower left;

and Western blot of knockdown samples are shown on lower right. Error bar in all panels represents standard deviation. Bars: 100μm.

Figure 8. A schematic diagram summarizing the multiple interactions mediated by the ND and PHD domains of Pygo2.

Figure 1

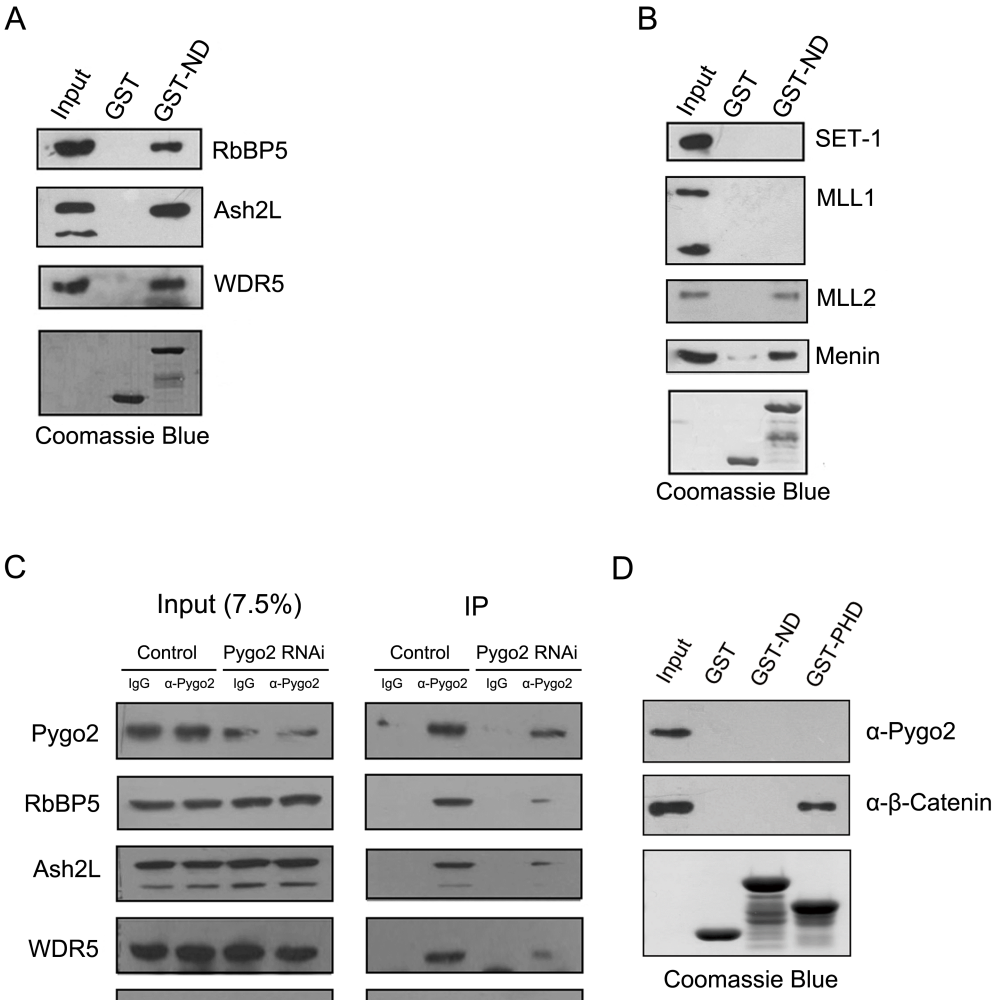

Figure 2

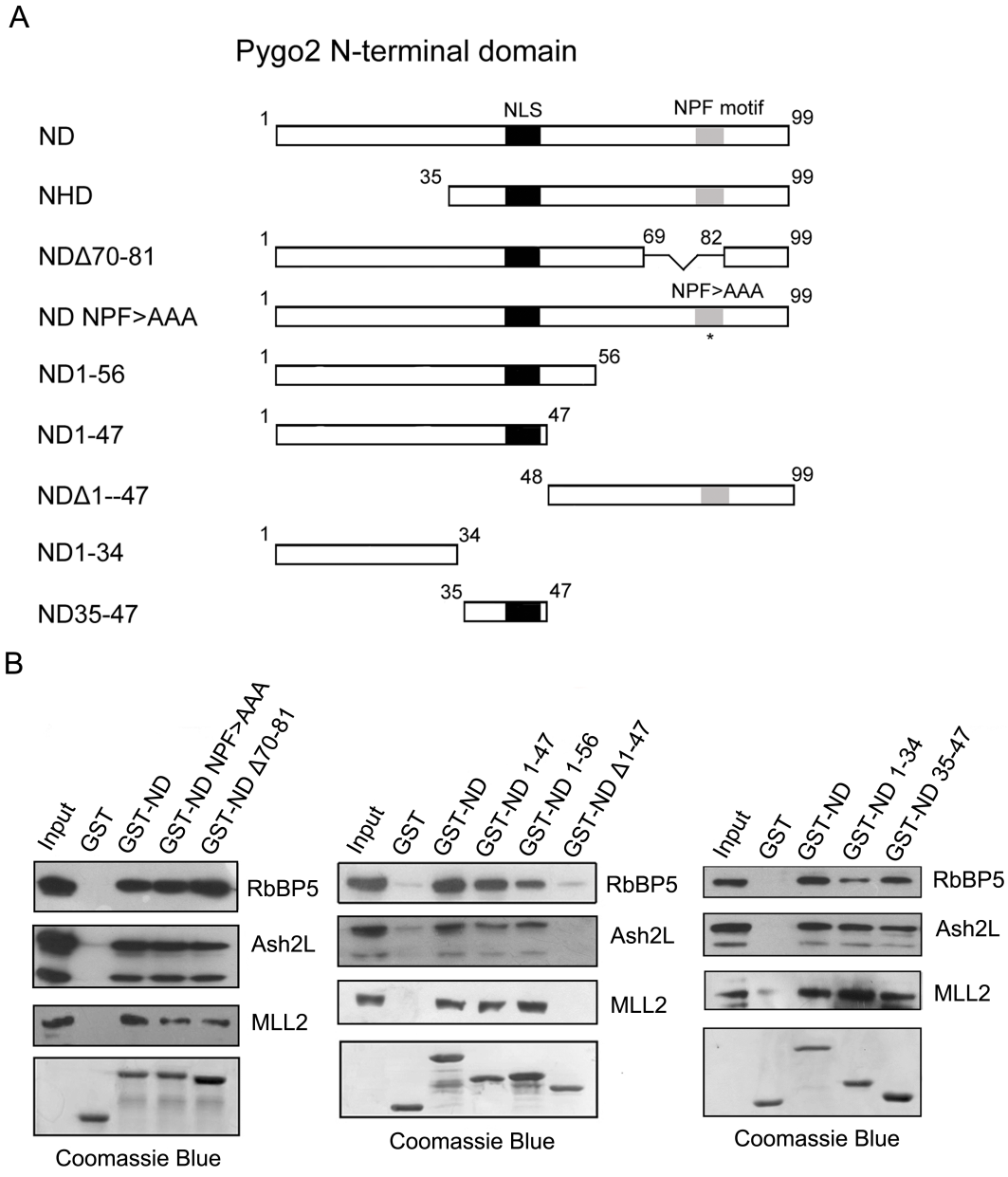

Figure 3

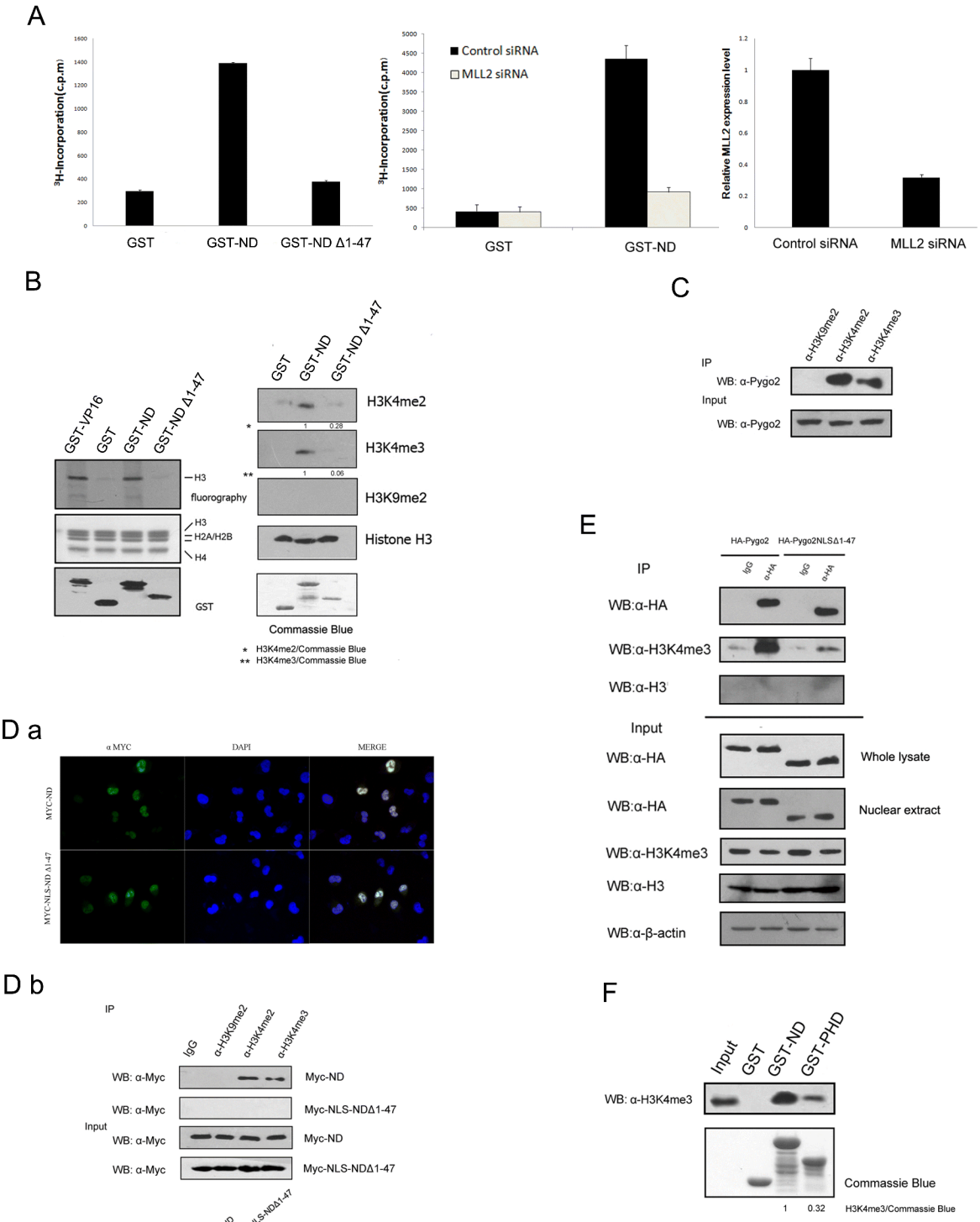

Figure 4

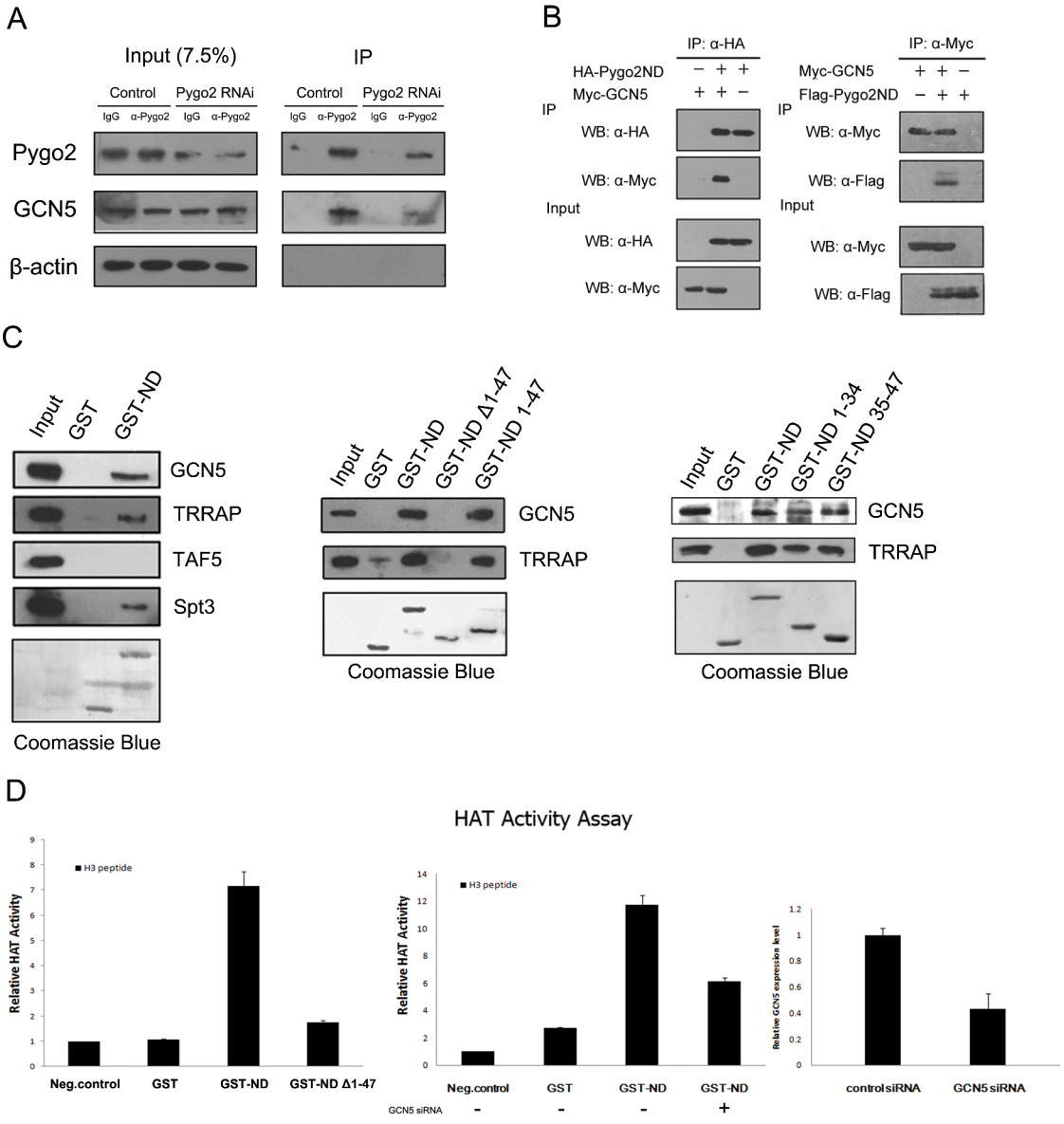

Figure 5

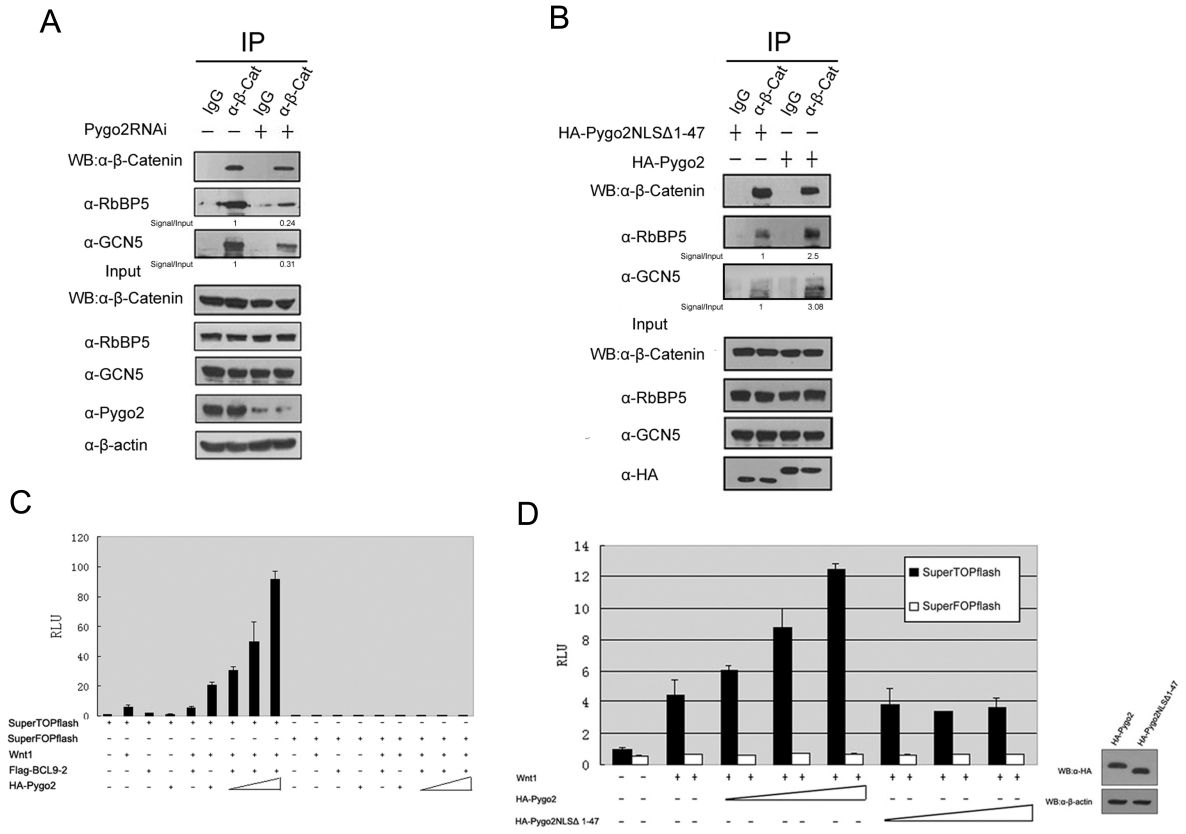

Figure 6

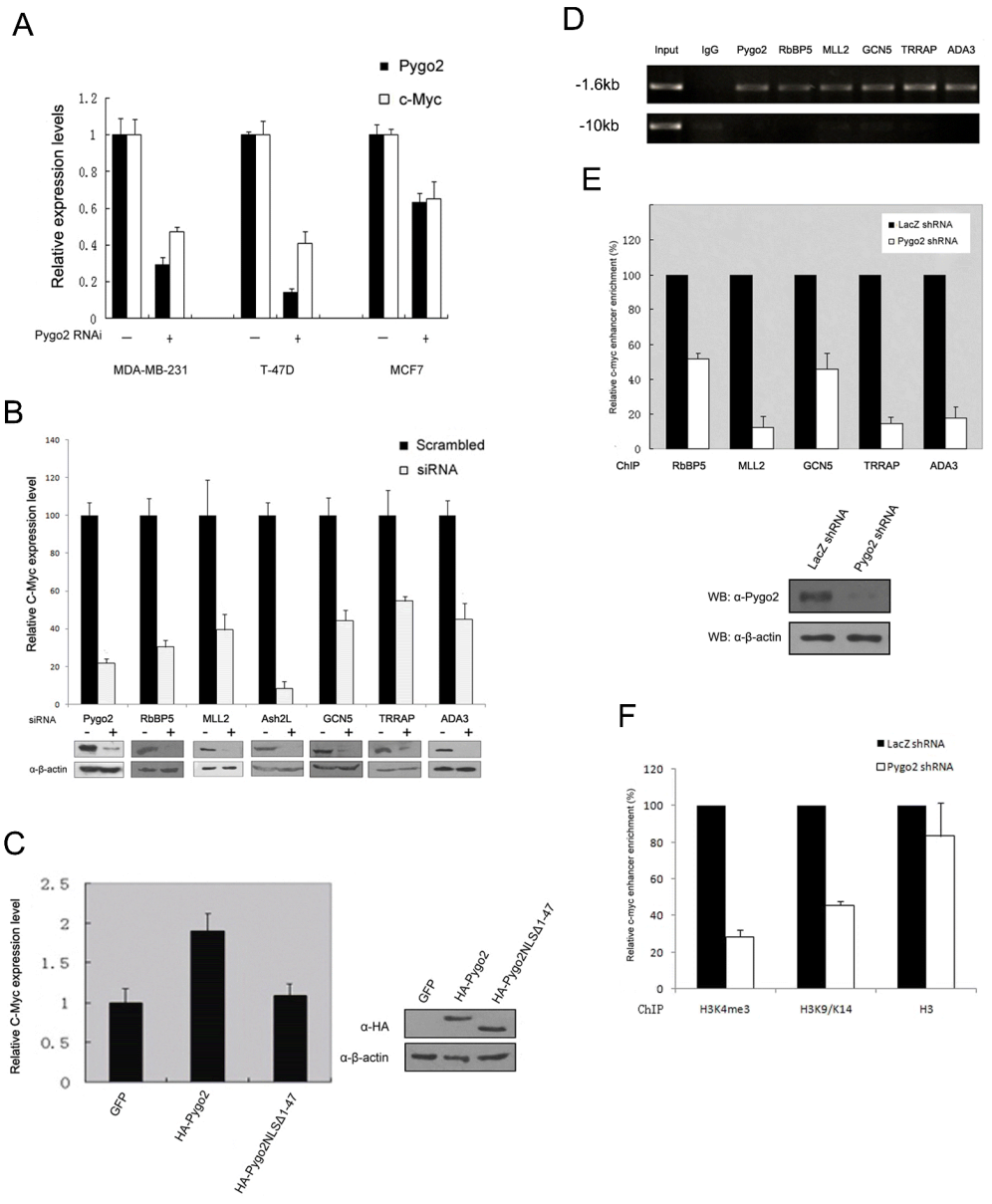

Figure 7

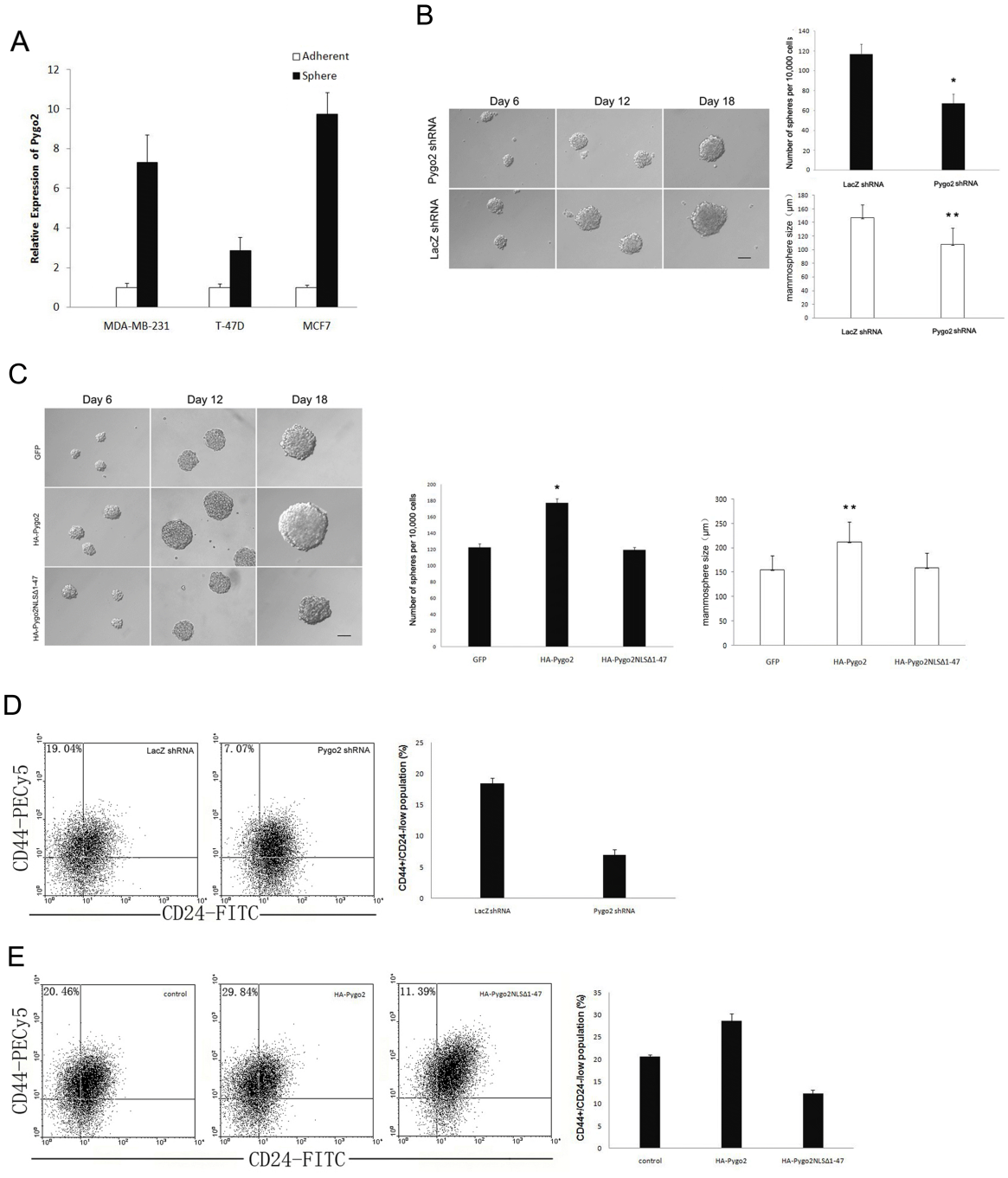

Figure 7

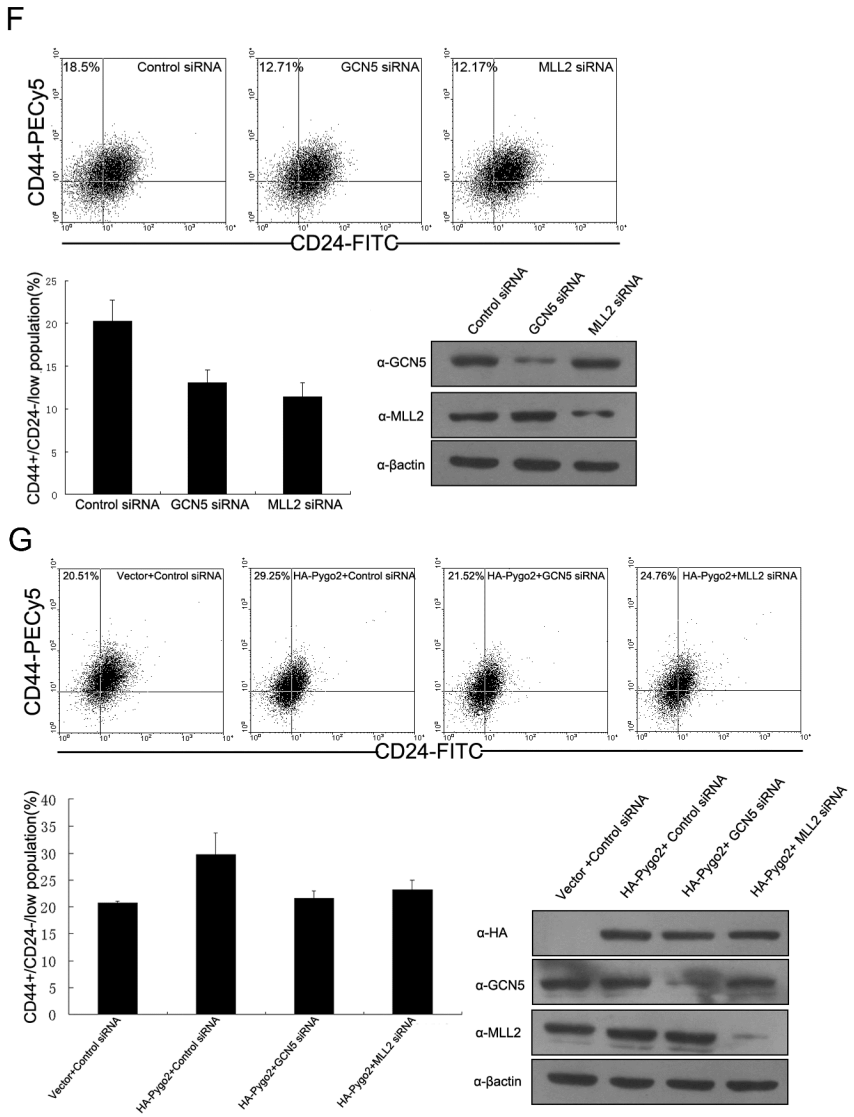

Figure 8