© 1999 Oxford University Press
Nucleic Acids Research, 1999, Vol. 27, No. 4 **1079–1083**

cDNA fingerprinting of osteoprogenitor cells to isolate differentiation stage-specific genes

G. A. Candeliere*, Y. Rao, A. Floh, S. D. Sandler and J. E. Aubin

Department of Anatomy and Cell Biology, Room 6230 Medical Sciences Building, University of Toronto, 1 King's College Circle, Toronto, Ontario M5S 1A8, Canada

Received October 5, 1998; Revised and Accepted December 15, 1998

ABSTRACT

A cDNA fingerprinting strategy was developed to identify genes based on their differential expression pattern during osteoblast development. Preliminary biological and molecular staging of cDNA pools prepared by global amplification PCR allowed discriminating choices to be made in selection of expressed sequence tags (ESTs) to be isolated. Sequencing of selected ESTs confirmed that both known and novel genes can be isolated from any developmental stage of interest, e.g. from primitive progenitors, intermediate precursors or mature osteoblasts. EST expression provides insight into possible interrelated physiological functions and putative interacting molecules during differentiation. This method offers a functional genomics approach to isolate differentiation stage-specific genes in samples as small as a single cell.

INTRODUCTION

Knowledge of the molecular and cellular events characterizing osteoblast development is growing as new markers, including important classes of regulatory molecules such as transcription factors (e.g. the Cbfa-1 family; 1), are elucidated. Nevertheless, a paucity of definitive and specific markers, especially for the more primitive progenitors and stem cells, slows advancement in the field in comparison with other lineages such as the hematopoietic lineages (2,3). One useful model, however, has been populations of freshly isolated cells derived from 21 day fetal rat calvaria (RC); when these are grown long term (~3 weeks) in medium supplemented with ascorbic acid and β-glycerophosphate, the low frequency (<1%) of osteoprogenitor cells present divide and differentiate to form mineralized bone nodules. Morphological (4), immunohistochemical (5) and molecular (6) analyses have confirmed that formation of bone nodules reproducibly recapitulates a proliferation–differentiation sequence from an early precursor cell to a mature osteoblast and in which expression of cell cycle-related and osteoblast-associated macromolecules is acquired and/or lost in a temporal sequence as cells differentiate (3,7). However, because osteoprogenitors comprise such a low fraction of cells in these populations and their differentiation does not occur synchronously, it can be difficult to ascertain with certainty the expressed gene repertoires in osteoprogenitors, particularly during very early events in the maturational sequence, without confounding contributions from other cells present in the cultures.

Recently, we isolated the osteoprogenitors by replica plating, created cDNA pools by using a global amplification poly(A) PCR technique (8) which faithfully amplifies mRNA from single cells or small colonies of cells while maintaining relative mRNA abundances (9), and established molecular profiles for multiple transitional stages as these cells matured to fully functional osteoblasts (10). Since these cDNA pools comprise the expressed gene repertoires of cells at different maturational stages, we have now developed and used a cDNA fingerprinting approach that uses a representative subset of these as the starting point for isolation of differentiation stage-specific genes.

MATERIALS AND METHODS

Selection of cDNA pools for fingerprinting

Twenty-eight poly(A) PCR cDNA pools representative of five transitional stages in osteoblast lineage progression were selected from more than 100 available amplified colonies (10). Stage A are replica-plated monolayer colonies committed to differentiate to the osteoblast lineage but not yet expressing type I α1 collagen or alkaline phosphatase, both early markers of osteoprogenitor cells. Stage B and C colonies are progressively more mature, i.e. expressing type I α1 collagen or both type I α1 collagen and alkaline phosphatase, respectively. Stage D colonies represent multilayered cells and contain histologically identifiable cuboidal osteoblasts. Stage E colonies comprise terminal differentiation stages, with multilayered cells and mineralized bone matrix. A sample of four to seven representative colonies were chosen from each category for comparison. Relative amounts of total cDNA were determined by Southern hybridization and were used to calculate equal amounts of template for cDNA fingerprinting from each colony.

Arbitrarily primed PCR

Conditions for PCR were chosen so as to minimize specificity while preserving reproducibility, thus AmpliTaq at a high concentration was used. The final concentrations of reagents in a 2.5 µl reaction volume were 1× AmpliTaq PCR buffer, 4 mM $MgCl_2$, 250 µM each of the four dNTPs, 0.1 µCi/µl [^{32}P]dCTP,

*To whom correspondence should be addressed. Tel: +1 416 978 6917; Fax: +1 416 978 3954; Email: tony.candeliere@utoronto.ca

0.1 U/µl Perkin Elmer Cetus AmpliTaq DNA polymerase, 10 µM primer and a 1:100 dilution of the first poly(A) PCR amplification. An initial priming step consisting of two cycles of 94°C for 1 min, 35°C for 5 min and 72°C for 5 min was followed by 25 cycles of 94°C for 1 min, 50°C for 1 min, 72°C for 1 min. Twenty-five cycles were used because relative mRNA abundance in different samples is maintained with this protocol as shown in the original poly(A) PCR protocol (used for amplification steps; 11). The entire 2.5 µl reaction mix was then run on a 5% denaturing gel (12). In the representative figure shown (Fig. 3), an arbitrary primer with the sequence TGTAGGAGCCAGAGGTGGTG was used. These PCR conditions produced a complex banding pattern for all primers tested to date, although varying film exposures are required for different primers (data not shown).

Expressed sequence tag (EST) identification

EST expression patterns were identified on films exposed to dried-down gels and bands containing DNA of interest were isolated for reamplification as described (12). Briefly, the piece of dried gel containing the DNA fragment of interest was placed in 500 µl of 50 mM NaOH and boiled to elute DNA. After 30 min, the solution was neutralized with 50 µl of 1 M Tris, pH 8.0. An aliquot (1/10 vol of PCR reaction) was then used for reamplification with standard PCR conditions (12) for up to 40 cycles at 94°C for 1 min, 50°C for 1 min, 72°C for 1 min. PCR fragments were then subcloned into the pBluescript II SK– cloning vector for sequencing and for use as templates for making DNA probes. Sequenced ESTs were screened against public databases by using the GCG Wisconsin package.

Confirmation of expression patterns

Expression patterns were initially confirmed on Southern blots containing poly(A) PCR-amplified material from more than 100 colonies at various transitional stages of osteoblast development (10). Results from the Southerns were also shown to correlate with expected patterns of expression obtained by northern hybridization to RNA isolated from differentiating primary RC cell cultures (10) and *in situ* hybridization (not shown). Southern hybridization signal intensity was quantified by densitometry (from phosphor-images using the IP Lab Gel analysis program) and corrected for loading against total cDNA. Colonies were grouped according to the above criteria into five stages of differentiation. Statistical significance was assessed with a statistics program (InStat) that calculated standard deviation and ANOVA.

RESULTS

Selection of amplified cDNA pools and the cDNA fingerprinting method

A schematic of the cDNA fingerprinting approach is outlined in Figure 1. The cDNA pools prepared with global amplification poly(A) PCR (Fig. 1A and B) were characterized by establishing their molecular profiles by Southern hybridization with 3′-end probes for known osteoblast markers (e.g. type I collagen, alkaline phosphatase, bone sialoprotein, osteocalcin and osteopontin) (Fig. 1C; 10). This allowed their alignment into pools of cDNAs representative of cells at different transitional stages from primitive progenitor to mature osteoblast; we selected 28 such cDNA pools from the more than 100 available on the basis of their

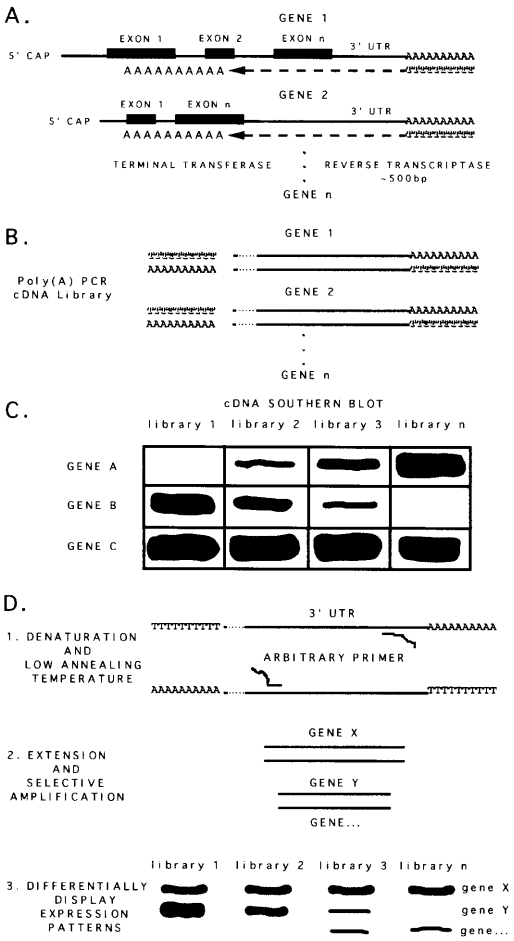

Figure 1. The cDNA fingerprinting strategy combines poly(A) or global amplification PCR on developmentally staged colonies of osteoprogenitor cells (10) with an arbitrarily primed PCR technique (21). Global amplification of mRNA from samples with limiting cell number (50–100 cells) was done by reverse transcription and addition of a second poly(A) tail to cDNA (**A**). Amplification of cDNA with preservation of relative cDNA abundance was achieved by the amplification of ~600 bp of 3′-end with an oligo(dT) primer (11) (**B**). The developmental status of the cells used to create the cDNA pools was confirmed by Southern blotting with 3′-ends of cDNA probes for known osteoblast markers to obtain relative gene expression profiles (**C**). Appropriately selected cDNA pools were amplified with arbitrary primers to generate a fingerprint of ESTs tags which differentially display patterns of gene expression (**D**) across the developmental sequence.

marker expression profiles (Fig. 2). The cDNA pools covered five transitional stages, labeled A–E; multiple examples for each stage were chosen for analysis to avoid spurious bands that might occur in only one colony. When an arbitrary primer was used to prime amplification in these cDNA pools (Fig. 1D), clearly different patterns of gene expression or fingerprints were evident in cells during lineage progression; differences were evident both in intensity of individual ESTs and in how many cDNA pools within a transitional stage expressed detectable levels of that EST (Fig. 3). Six ESTs from one particular fingerprint gel were selected (Fig. 3). The two largest tags were EST 1 and 2, of 750 and 700 bp, respectively. EST 1 corresponds to a relatively abundant mRNA, which is present in progenitor colonies but is up-regulated during intermediate differentiation stages and maintains this level through cell maturation. In contrast, EST 2

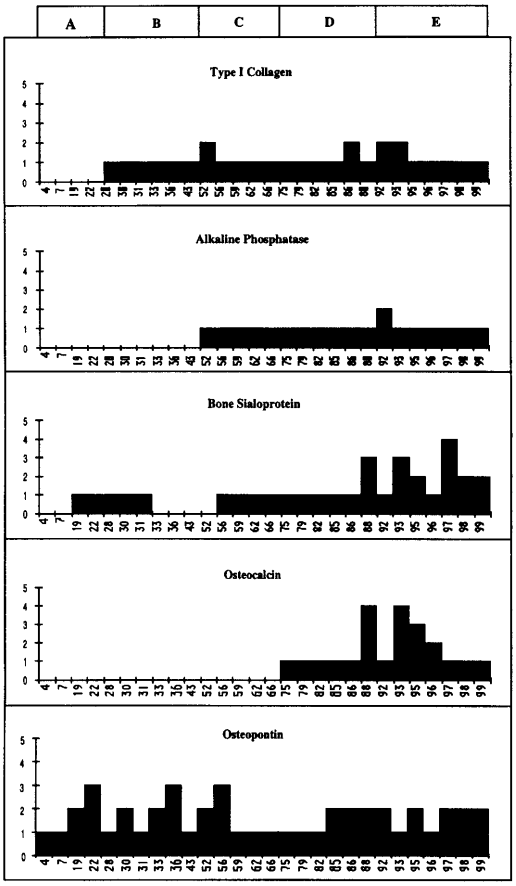

Figure 2. Samples for fingerprinting were selected based on the basic molecular phenotype of the cDNA pools made from discrete isolated colonies at different stages of osteoblast differentiation and bone development. Gene expression profiles of colonies were determined by analyzing expression of several known osteoblast lineage markers (10). These 28 cDNA pools, from a group of more than 100 characterized cDNA pools, were selected for fingerprinting on the basis that they represent several transitional stages: primitive progenitors (**A**), progressively more mature precursors (**B–D**) and terminally differentiated, bone-forming osteoblasts (**E**) (Materials and Methods). While category order is progressive, the order of colonies within each category is random.

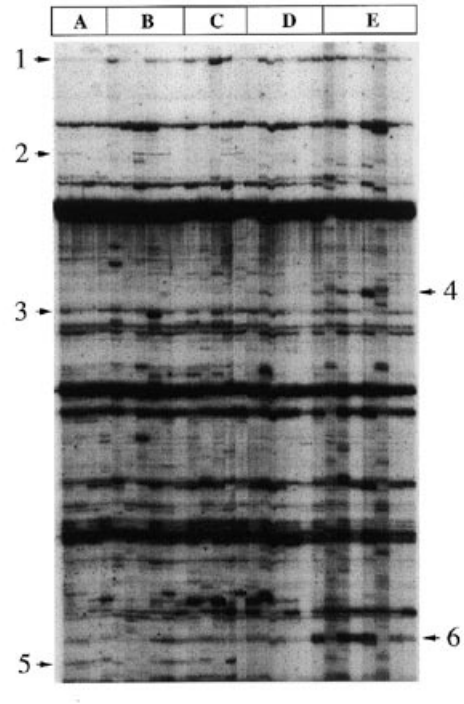

Figure 3. Amplification with arbitrary primers of appropriately selected cDNA pools differentially displays several different ESTs with varied patterns of gene expression during osteoblast development (Materials and Methods). Selected cDNA pools are aligned from left to right according to progression from immature osteoprogenitor to mature osteoblast respectively, as shown in Figure 2. Six different ESTs were selected based on size, relative abundance and expression patterns (text). Selected ESTs were reamplified and sequenced (Materials and Methods). Expression patterns were confirmed on Southern blots containing >100 cDNA pools representative of the entire osteoblast differentiation sequence (Materials and Methods).

corresponds to a much lower abundance mRNA, which is detectable in progenitor cells, but is up-regulated and then down-regulated to undetectable levels late in the differentiation sequence. For comparison, EST 3 (~300 bp) is not differentially expressed in cells during lineage progression, EST 4 (~300 bp) and EST 6 (~150 bp) are up-regulated later in the differentiation sequence and EST 5 (~150 bp) decreases during osteoblast development. The expression patterns of these six ESTs were confirmed by probing the full array of lineage Southern blots (>100 cDNA pools) representative of cells at various transitional stages of osteoblast development (Fig. 4); for some ESTs, northern blots of differentiating primary RC cell cultures were also done (data not shown). Quantitative analysis of the Southern blots revealed statistically significant changes in their expression levels over the five differentiation stages represented by the cDNA pools (Fig. 4), with the exception of EST 3, whose relatively uniform level of expression was confirmed.

Identification of selected ESTs and confirmation of expression patterns

Sequencing revealed that EST 1 was the rat homolog of osteonectin, a marker confirmed by both Southern and northern analysis (not shown) to be up-regulated as osteoprogenitors in RC cell populations differentiate and form bone nodules; osteonectin expression peaked approximately concomitantly with up-regulation of type I collagen (13–15; Fig. 4). Similarly, EST 6 was found to be osteocalcin, an abundant, well-characterized extracelluar matrix molecule in bone and a late marker of osteoblast development (2,3,5,6), consistent with its pattern seen in the fingerprints. EST 4 was identified as cystatin c (16,17), a cysteine protease inhibitor only recently identified in bone; we found it to be more highly expressed at intermediate developmental stages, with peak expression concomitant with the acquisition of alkaline phosphatase expression. EST 2 corresponds to a novel low abundance gene, whose expression also peaks concomitant with up-regulation of alkaline phosphatase expression but decreases to non-detectable levels at late differentiation stages (18). EST 5 is also a novel gene, confirmed to be more highly expressed in osteoprogenitors than in their more differentiated progeny. The sequence of EST 3 also suggests that it is a previously uncharacterized gene.

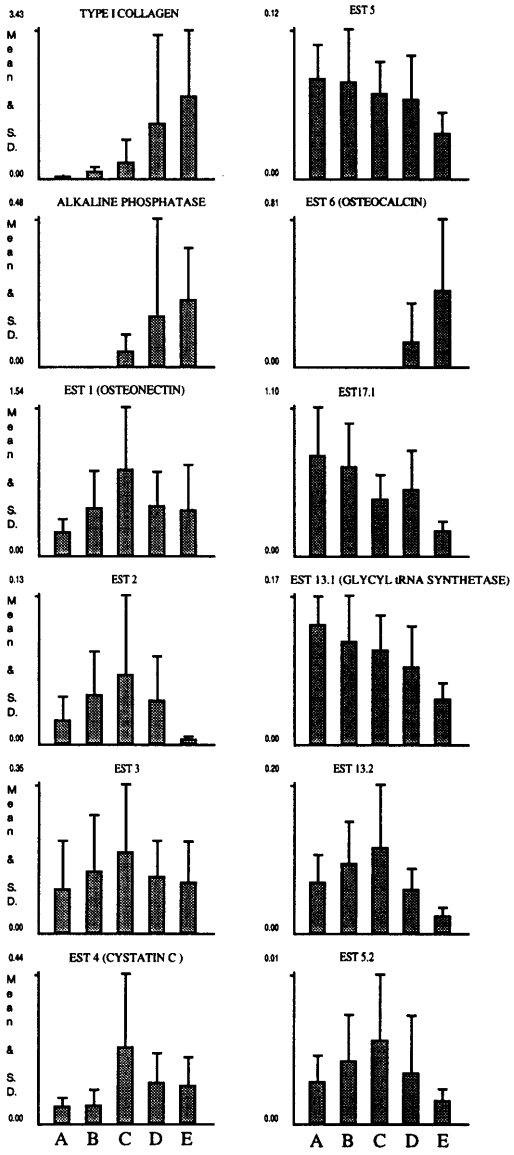

Figure 4. EST identification and confirmation of expression patterns (Materials and Methods). The means and standard deviations of expression levels, relative to the total cDNA quantity as a standard, were determined by phophorimaging; sequencing identified some ESTs as known genes (identified in brackets) or as novel. Type I α1 collagen and alkaline phosphatase are included as known osteoblast marker controls. ANOVA confirmed significant differences between mean expression level at different developmental stages for all tags except EST 3, which we include as an example of a housekeeping gene. The ability to predict expression patterns from a cDNA fingerprint allowed selection of ESTs whose expression is highest in early progenitors (e.g. 17.1 and 13.1), those whose expression peaks in intermediate precursors (e.g. 2, 13.2 and 5.2) or those peaking in mature cells (e.g. EST 6).

Targeted selection of other ESTs

We next determined whether other random primers would prove equally efficacious. With every primer tried to date, usable fingerprints have resulted. For example, EST 17.1 (a novel gene) and EST 13.1 (identified by sequence analysis to be glycyl tRNA synthetase) are both more highly expressed in more primitive than in more mature osteoblast lineage cells, while expression of ESTs 13.2 and 5.2 peak at intermediate maturational stages (Fig. 4).

DISCUSSION

Our results show that we have developed a powerful method, termed cDNA fingerprinting, to analyze globally amplified poly(A) PCR cDNA pools to differentially display patterns of gene expression during lineage progression. Using these expression patterns, we predicted, and have isolated from these cDNA pools, genes for known osteoblast markers (e.g. osteonectin and osteocalcin), others for molecules recently found in osteoblasts but for which differential expression during differentiation had not been described (e.g. cystatin c and glycyl tRNA synthetase) and still others for novel (e.g. ESTs 2, 3, 5, 17.1, 13.2 and 5.2) molecules that demarcate, and some of which probably regulate, transitions between different maturational stages. With respect to the first class (osteonectin and osteocalcin), our data from fingerprinting and secondary screens on the Southern lineage blots are entirely consistent with other approaches to assess differentiating osteoblasts, e.g. northerns on differentiating cultures or *in situ* hybridization of bone tissue (13–15). In the second class (cystatin c and glycyl tRNA synthetase), we extend information from that previously reported. For example, cystatin c has recently been shown to be expressed by osteoblasts (16,17), but whether it was differentially expressed as osteoblast precursors mature was not analyzed. Given the fact that cystatin c is thought to inhibit cysteine proteinases secreted by osteoclasts, it is interesting that its maximum expression appears to coincide with early events in osteoblast differentiation, prior to peak expression of alkaline phosphatase, followed by a plateau to intermediate levels during late phases of differentiation. This suggests a differentiation-related mechanism by which osteoblast precursors might down-regulate osteoclastic resorptive activity, coupling the latter to osteoblast maturation and in-filling with new bone. It is also notable that the mRNA for glycyl tRNA synthetase achieves highest levels at the outset of differentiation prior to up-regulation of type I collagen expression, since type I collagen, which makes up 90% of the matrix of bone, is high in glycine content (collagen is a long GLY-X-Y repeated amino acid chain). Thus, the approach described here clearly demonstrates gene expression cascades, at least some of which are presumably for interrelated physiological functions, and predicts possible interacting or upstream or downstream molecules in such cascades based on temporally related expression profiles. In this context, the ESTs for novel genes, notably ESTs 2, 5, 17.1, 13.2 and 5.2, are particularly interesting as their maximum expression is in the earliest progenitor cells, i.e. in cells that have not yet reached differentiation stages commonly accessible *in vitro* and *in vivo*, emphasizing the utility of the approach for identifying molecules for the most elusive cells in this or other lineages.

cDNA fingerprinting is a powerful approach to identify markers of interest from biologically and minimally molecularly characterized samples obtained from sources in which the quantity of mRNA is extremely limited. This approach provides a functional genomics strategy which targets any physiological or even pathological differentiation stage of interest for any cell population in tissue culture or in tissue sections in which a differentiation or developmental sequence can be identified. There are multiple approaches that may be used to obtain relevant cells from tissue culture or from tissue sections (19) for cDNA fingerprinting; these include simple micromanipulation to laser capture microdissection (20). The specific example we used comprised cDNA pools prepared from replica-plated differentiating

osteoblast colonies; the replica plating allowed unambiguous retrospective identification of low frequency osteoprogenitors, whose location *in vivo* is uncertain, through multiple transitional steps to the fully functional terminally differentiated osteoblast (10). Combined as it was with a basic molecular characterization of the cells from which the RNA was obtained allowed appropriately developmentally staged populations to be compared. Arbitrarily primed PCR approaches (21) to compare nucleic acid samples include protocols for DNA fingerprinting, designed to identify and trace genetic markers between DNA samples (22), and RNA fingerprinting (23,24), with arbitrary primer variations to selectively amplify mRNA (differential display PCR; 12,25,26), particular gene family members (27) or identifiable 3′-end restriction fragments (28). The poly(A) PCR approach (29–32) globally amplifies mRNA while maintaining relative abundance and provides cDNA pools that can be reamplified indefinitely. Cell type-specific and differentiation stage-specific gene searches have and are being done on poly(A) PCR cDNA pools by subtractive hybridization approaches (11), but the subtraction strategy relies on comparison of only two populations or stages of interest at any one time. Our strategy, on the other hand, allows rapid comparison of numerous samples simultaneously. There are, however, some limitations to our approach. One of these relates to the possibility of isolating multiple ESTs of one size in one band, although strategies exist for overcoming this problem (33). Second, 'false positive' and 'false negative' signals have been seen with some small ESTs (<300 bp) when these have been rescreened in secondary screens on cDNA Southern and northern blots. Third, the strategy may fail to identify particularly low abundance mRNAs, since these may not be detectably amplified in the initial global amplification step. Finally, cDNA fingerprinting may not be appropriate for the selective amplification of particular family members, since the global amplification strategy amplifies relatively short gene tags (~0.6–0.7 kb) at the 3′-end of mRNAs, frequently outside the coding region of genes of interest. Nevertheless, as we have shown, the strategy described here is applicable to many physiological or pathological developmental systems. By virtue of simultaneous analysis of expression profiles over multiple stages of interest, it also identifies and predicts potential temporally relevant interrelationships and interacting molecules in particular differentiation cascades.

ACKNOWLEDGEMENTS

The authors thank B. Ganss, R. Hudson, E. Bonnelye and F. Liu for discussions and U. Bhargava for technical assistance. G.A.C. was funded by an MRC of Canada Post-doctoral Fellowship. This work was funded by MRC grant MT-12390 to J.E.A.

REFERENCES

1 Ducy,P., Zhang,R., Geoffroy,V., Ridall,A.L. and Karsenty,G. (1997) *Cell*, **89**, 747–754.
2 Aubin,J.E. and Liu,F. (1996) In Bilezikian,J.P., Raisz,L.G. and Rodan,G.A. (eds), *The Osteoblast Lineage*. Academic Press, Toronto, Canada, pp. 51–67.
3 Stein,G.S., Lian,J.B., Stein,J.L., van Wijnen,A.J., Frenkel,B. and Montecino,M. (1996) In Bilezikian,J.P., Raisz,L.G. and Rodan,G.A. (eds), *Mechanisms Regulating Osteoblast Proliferation and Differentiation*. Academic Press, Toronto, Canada, pp. 69–86.
4 Bhargava,U., Bar-Lev,M., Bellows,C.G. and Aubin,J.E. (1988) *Bone*, **9**, 155–163.
5 Liu,F., Malaval,L., Gupta,A.K. and Aubin,J.E. (1994) *Dev. Biol.*, **166**, 220–234.
6 Liu,F., Malaval,L. and Aubin,J.E. (1997) *Exp. Cell Res.*, **232**, 97–105.
7 Aubin,J.E., Turksen,K. and Heersche,J.N.M. (1993) In Noda,M. (ed.), *Osteoblastic Cell Lineage*. Academic Press, Toronto, Canada, pp. 1–45.
8 Cumano,A., Paige,C.J., Iscove,N.N. and Brady,G. (1992) *Nature*, **356**, 612–615.
9 Aubin,J.E., Liu,F., Malaval,L. and Gupta,A. (1995) *Calcif. Tissue Int.*, **56**, S54–S56.
10 Liu,F. and Aubin,J.E. (1999) *Dev. Biol.*, submitted for publication.
11 Malaval,L., Modrowski,D., Gupta,A.K. and Aubin,J.E. (1994) *J. Cell. Physiol.*, **158**, 555–572.
12 Yao,K.L., Todescan,R.,Jr and Sodek,J. (1994) *J. Bone Miner. Res.*, **9**, 231–240.
13 Lynch,M.P., Stein,J.L., Stein,G.S. and Lian,J.B. (1995) *Exp. Cell Res.*, **216**, 35–45.
14 Lerner,U.H. and Grubb,A. (1992) *J. Bone Miner. Res.*, **7**, 433–440.
15 Lerner,U.H., Johansson,L., Ranjso,M., Rosenquist,J.B., Reinholt,F.P. and Grubb,A. (1997) *Acta Physiol. Scand.*, **161**, 81–92.
16 Candeliere,G.A. and Aubin,J.E. (1997) *J. Bone Miner. Res.*, **12**, S141.
17 Jensen,R.A., Page,D.L. and Holt,J.T. (1997) *Methods Mol. Biol.*, **85**, 277–283.
18 Emmert-Buck,M.R., Bonner,R.F., Smith,P.D., Chuaqui,R.F., Zhuang,Z., Goldstein,S.R., Weiss,R.A. and Liotta,L.A. (1996) *Science*, **274**, 998–1001.
19 McClelland,M., Mathieu-Daude,F. and Welsh,J. (1995) *Trends. Genet.*, **11**, 242–246.
20 Welsh,J. and McClelland,M. (1990) *Nucleic Acids Res.*, **18**, 7213–7218.
21 McClelland,M. and Welsh,J. (1994) *PCR Methods Appl.*, **4**, S66–S81.
22 McClelland,M., Honeycutt,R., Mathieu-Daude,F., Vogt,T. and Welsh,J. (1997) *Methods Mol. Biol.*, **85**, 13–24.
23 Liang,P. and Pardee,A.B. (1992) *Science*, **257**, 967–971.
24 Liang,P. and Pardee,A.B. (1998) In Ausubel,F.A., Brent,R., Kingston,R.E., Moore,D.D., Seidman,J.A., Smith,J.A. and Struhl,K. (eds), *Differential Display of mRNA by PCR*. John Wiley & Sons, New York, NY, pp. 15.8.1–15.8.2.
25 Liang,P. and Pardee,A.B. (1997) *Methods Mol. Biol.*, **85**, 3–11.
26 Donohue,P.J., Hsu,D.K. and Winkles,J.A. (1997) *Methods Mol. Biol.*, **85**, 25–35.
27 Prashar,Y. and Weissman,S.M. (1996) *Proc. Natl Acad. Sci. USA*, **93**, 659–663.
28 Brady,G., Billia,F., Knox,J., Hoang,T., Kirsh,I.R., Voura,E.B., Hawley,R.G., Cumming,R., Buchwald,M., Siminovitch,K., Miyamoto,N., Boehmelt,G. and Iscove,N.N. (1995) *Curr. Biol.*, **5**, 909–922.
29 Rambhatla,L., Patel,B., Dhanasekaran,N. and Latham,K.E. (1995) *Mol. Reprod. Dev.*, **41**, 314–324.
30 Kato,K., O'Dowd,D.K., Fraser,S.E. and Smith,M.A. (1997) *Dev. Biol.*, **191**, 259–269.
31 Brown,J., Brandwood,C., Hazelhurst,Z., Ashton,B.A., Brady,G., Freemont,A.J. and Hoyland,J.A. (1997) *J. Pathol.*, **181**, 12.
32 Brady,G. and Iscove,N.N. (1993) *Methods Enzymol.*, **225**, 611–623.
33 McClelland,M., Arensdorf,H., Cheng,R. and Welsh,J. (1994) *Nucleic Acids Res.*, **22**, 1770–1771.