SUPPORTING METHODS

Array comparative genomic hybridization analysis

Mouse 105k CGH array analysis was done using CGH Analytics 3.514 (Agilent). Aberration detection was accomplished using the ADM-1 algorithm (Lipson et al., 2006), with a threshold of 6. Centralization was used with a bin size of 10, and fuzzy zero correction was applied to minimize detection of segments with low absolute mean ratios. Recurrent aberrations in mouse data were defined using the T Test common aberration algorithm, utilizing an overlap threshold of 0.9. Genome build used for these analyses was mouse NCBI 36.

Gene expression array analysis

Microarray normalization and expression analysis was done using Genespring 7.3 (Agilent). For Agilent mouse whole genome arrays, values below 0.01 were set to 0.01 and each measurement was divided by the 50th percentile of all measurements in that sample. Each gene was divided by the median of its measurements in all samples and if the median of the raw values was below 10 then each measurement for that gene was divided by 10 if the numerator was above 10, otherwise the measurement was thrown out. For Affymetrix HGU133A arrays, raw CEL files were preprocessed using robust multi-array average (RMA) method (Irizarry et al., 2003) and batch correction using distance weighted discrimination (Benito M et al 2004). Further, each gene was divided by the median of its measurements in all samples. If the median of the raw values was below 0.01 then each measurement for that gene was divided by 0.01 if the numerator was above 0.01, otherwise the measurement was thrown out. For response to chemotherapy, samples were ranked low to high according to the mean probe measurements for each gene in question, and then divided into groups at the 1st quartile or median.

Gene expression signatures were obtained using a welsh T Test, and the False Discovery Rate (Benjamini and Hochberg, 1995) was employed to adjust for multiple hypotheses, using a cutoff of 5%. Genes within each signature were counted on the basis of official gene symbols.

Integrative copy number and expression analysis

For integrative analysis, multiple probes for each gene within the mouse 11qE1 deletion region were averaged and the correlation between these values and the mean log2ratio over the copy number interval that each gene occupies was determined using spearman's rank correlation coefficient.

Unsupervised clustering

Hierarchical clustering (Pearson's correlation coefficient and average linkage) was performed using the 500 genes with the highest coefficient of variation. Genes with no data in half the starting conditions were discarded. Cluster stability: We first scaled the data by dividing the expression profiles of each sample by their root-mean-square. The root-mean-square is obtained by computing the square-root of the sum-of-squares of the non-missing values in the sample divided by the number of non-missing values minus one. We evaluated the cluster stability by repeatedly subsampling the data and comparing the cluster membership of the subsamples to the original clusters (Ben-Hur et al. 2002; Smolkin and Ghosh, 2003). We used "ward" as the agglomeration method in hierarchical clustering (see hclust in R (http://www.r-project.org/)). We did 5,000 permutations and 85% of genes were used in each subsample.

Class prediction

A gene expression signature was obtained by comparing classes within the training samples (MOTO PKA^+ vs. PKA^-) and this gene set was used for class prediction. K-means cluster analysis was done in R. First the training samples were clustered into k=2 clusters, and then the kmeans.predict function in R was used to assign the test samples to a group. Principle component analyses was done using mean centering and scaling. Hierarchical clustering used Pearson's correlation coefficient and average linkage.

Functional annotation of gene signatures

Gene set enrichment was calculated using the hypergeometric test, or using EASE (DAVID bioinformatics database; Huang da et al., 2007) for annotation of over and underexpressed genes and for construction of functional clusters.

Databases

Gene sets were collected from the GO database (www.geneontology.org) and mSIG (http://www.broad.mit.edu/gsea/msigdb/index.jsp). Mouse-human orthologs were obtained from the Ensembl (53) Genes database. Human OSA gene expression data is available from NCBI Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo/) under the accession numbers GSE16088 and GSE16091.

Supplemental methods Table S8. Applied Biosystems Gene Expression Assay catalogue numbers and sequences

Gene Name	Catalogue number
Prkaca	Mm00660092_m1
Prkar1a	Mm00660315_m1
Rankl	Mm01313944_g1
Opg	Mm01205928_m1
Osteonectin	Mm00486332_m1
Osteopontin	Mm00436767_m1
Runx2	Mm00501578_m1
Col1a-1	Mm01302047_g1
Eukaryotic 18s rRNA	Hs99999901_s1
Custom Primers / Probes	
Gene Name	Sequence
SV40 large T-antigen	
Probe	5’ - AAG CAA CTC CAG CCA TCC ATT CTT CTA TGT C - 3'
Forward	5’ - TTT GGG CAA CAA ACA GTG TAG C - 3’
Reverse	5’ - AAT GTT TGG TTC TAC AGG CTC TGC - 3’
Osteocalcin	
Probe	5’ - AGG AGG GCA ATA AG - 3’
Forward	5’ - CTG ACA AAG CCT TCA TGT CCA - 3’
Reverse	5’ - GCG GGC GAG TCT GTT CAC TA - 3’

Alkaline phosphatase

Probe	5’ - TAC TGG CGA CAG CAA G - 3’
Forward	5’ - TTG TGC CAG AGA AAG AGA GAG A - 3’
Reverse	5’ - GTT TCA GGG CAT TTT TCA AGG T - 3’

References

Ben-Hur, A. Elisseeff and I. Guyon. A stability based method for discovering structure in clustered data. Pacific Symposium on Biocomputing, 2002.

Benjamini, Yoav; Hochberg, Yosef (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society, Series B (Methodological) 57 (1): 289–300.

Huang da, W., Sherman, B.T., Tan, Q., Collins, J.R., Alvord, W.G., Roayaei, J., Stephens, R., Baseler, M.W., Lane, H.C., and Lempicki, R.A. (2007). The DAVID Gene Functional Classification Tool: a novel biological module-centric algorithm to functionally analyze large gene lists. Genome biology 8, R183.

Irizarry, R.A., Bolstad, B.M., Collin, F., Cope, L.M., Hobbs, B., and Speed, T.P. (2003). Summaries of Affymetrix GeneChip probe level data Nucleic acids research 31, e15.

Lipson, D., Aumann, Y., Ben-Dor, A., Linial, N., and Yakhini, Z. (2006). Efficient calculation of interval scores for DNA copy number data analysis. J Comput Biol 13, 215-228.

Smoklin M., Ghosh D. (2003) Cluster stability scores for microarray data in cancer studies. BMC Bioinformatics, 4:36.

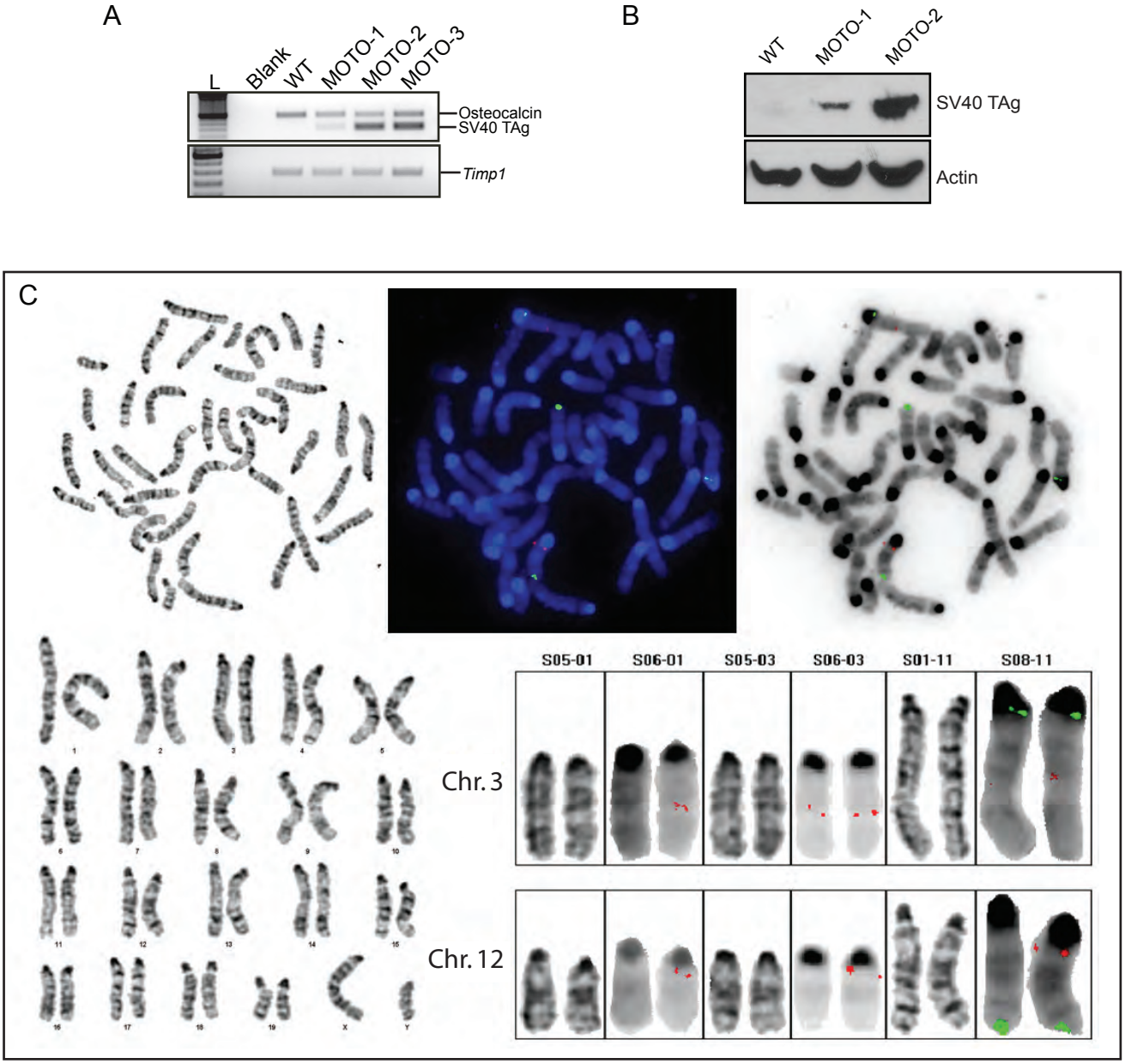

Figure S1. SV40-T-antigen Transgene Copy Number, Expression and Localization in MOTO

(A) OG2-TAg PCR of genomic DNA shows the relative transgenic copy number in male MOTO mice; 575bp, WT osteocalcin; 315bp transgene; 400bp, *Timp1* (endogenous copy number control). (B) Western blot of SV40 TAg in 15.4 week WT, MOTO-1 and 10.3 week MOTO-2 femurs. (C) OG2-SV40 transgene mapping using FISH on splenocytes from MOTO-1 mice identifies transgene insertion site at Chromosome 12qA-B1 and the endogenous OG2 promoter (*Bglap2*) at Chr3qE3-F1 (transgene probe = red signal). The identity of chromosomes 12 and 3 were confirmed using control probes specific for Chr 3qA1 and 12qF2 (green).

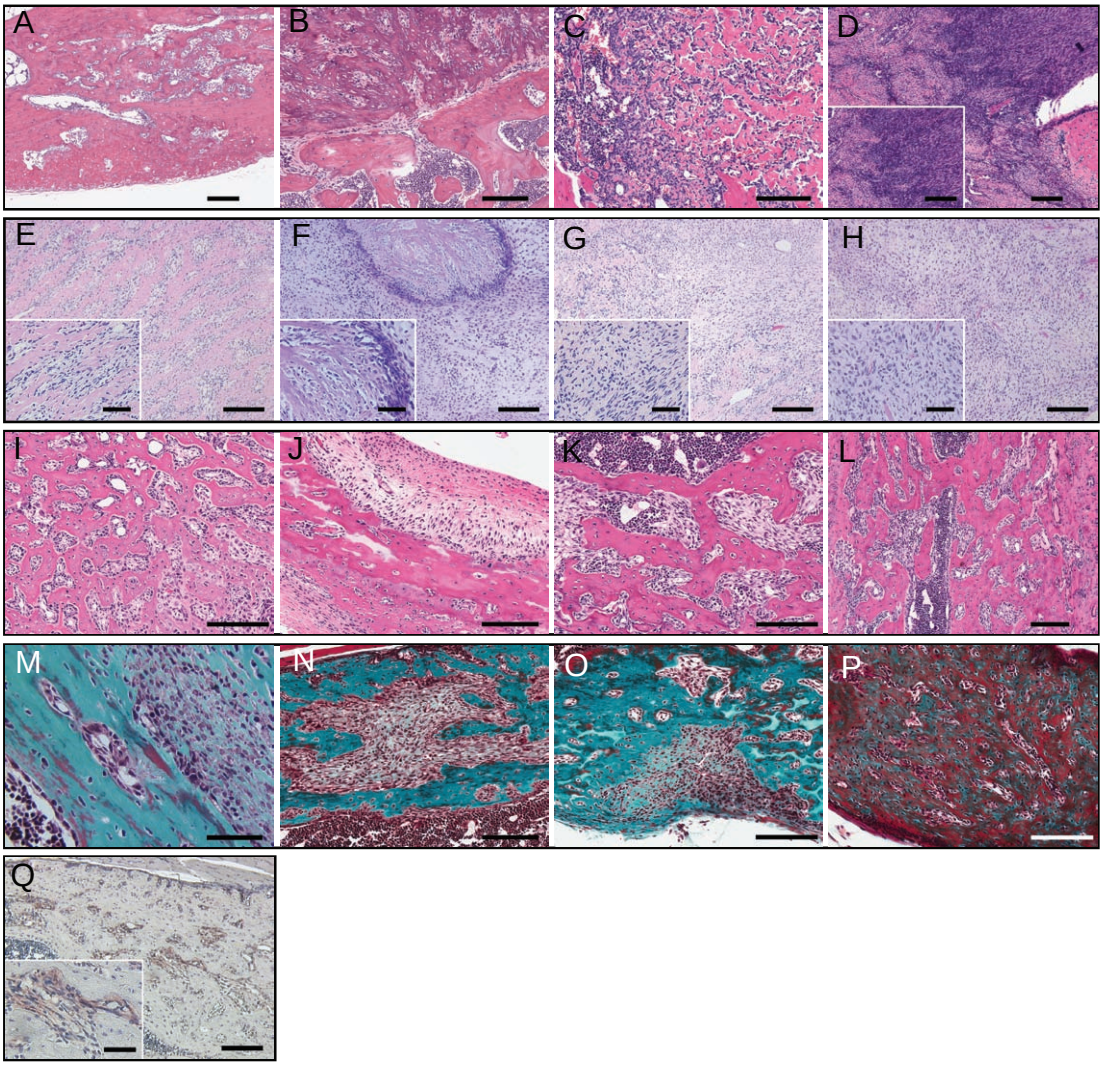

Figure S2. OSA Characterization in MOTO and *Prkar1a*$^{\Delta OB/+}$/MOTO^{+} Mice

(A-D) Representative H&E stained OSA variants: MOTO mice exhibit a spectrum of bone neoplasms that vary in malignant appearance and site. (A) Benign well differentiated cortical growth in distal tibia. (B) Parosteal osteosarcoma: extracortical with only focal attachment to the periosteum of the tibia and no evidence of invasion. (C) Classical osteoblastic osteosarcoma originating in the femoral medulla that invades and replaces bone marrow and deposits osteid. (D) Poorly differentiated fibroblastic osteosarcoma in femur that has destroyed regular bone structure but still producs some osteoid. (E-H) PKA^{+} and PKA^{-} mouse OSA subclasses show can not be distinguished histologically. Shown are H&E stained osteoblastic (E, F) and fibroblastic (G, H) osteosarcoma variants from representative end-stage (>20 weeks of age) tumours in the PKA^{-} (E, G) and PKA^{+} (F, H) subclasses. (I-L) H&E and (M-P) Masson's trichrome stained *Prkar1a*$^{\Delta OB/+}$/MOTO^{+} tumors. Medullary tibia (I) humeral (L) and cortical femoral (K) tumors displaying the features of osteoblastic OSA with disorganized trabecular-like structures interspersed with neoplastic cells depositing osteoid. Tibial (N) and femural (O) tumors having cellular areas that are more fibroblastic in appearance but that still deposit osteoid as evidenced by pale gree trichrome staining (N). (J) Tumour of periosteal origin invading into cortex of distal tibia. Tibial (M, O) and humoral (P) tumours with osteoid deposition. Mice are 3.4-4.1 weeks of age. (Q) Alkaline phosphatase immunostaining of *Prkar1a*$^{\Delta OB/+}$/MOTO^{+} humerus tumor showing positive osteoblasts (red).

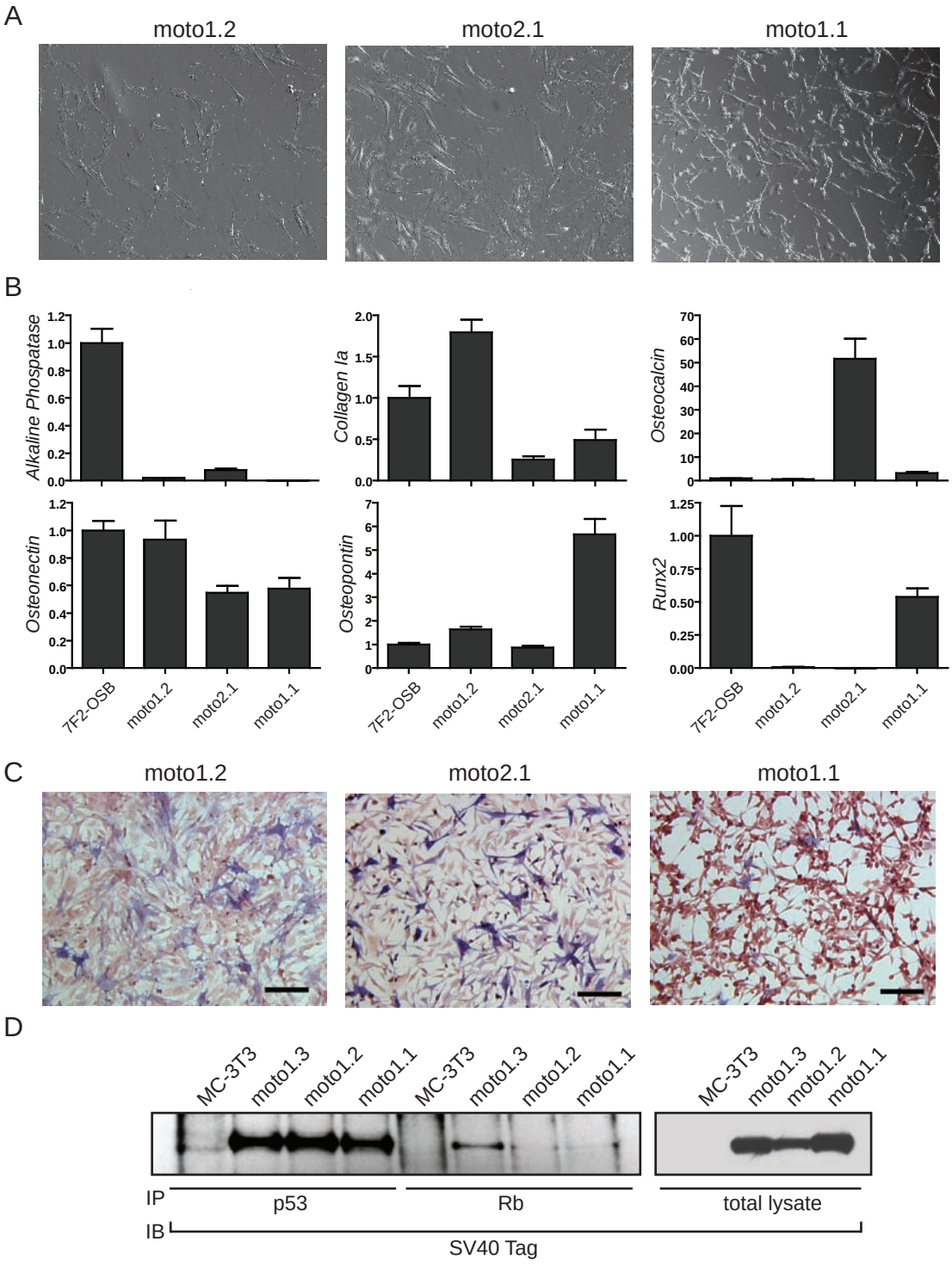

Figure S3. Markers of Osteoblastic Differentiation MOTO Tumor Cell Lines

The cell lines are numbered according to their MOTO founder line of origin followed by their order of derivation. 7F2-OSB is a control osteoblastic cell line. (A) Phase contrast light microscopy images of the three cell lines. The cells exhibit a fibroblastic morphology. (B) Relative qPCR showing expression of osteoblastic markers: alkaline phosphatase, collagen 1a1, osteocalcin, osteonectin, osteopontin and Runx2. Expression levels are normalized to endogenous 18S and are depicted relative to the values for the 7F2 osteoblastic cell line hich was assigned a level of 1. Data are represented as mean±SEM. (C) Alkaline phosphatase histochemical staining of cultured moto1.1, moto1.2, and moto2.1 cell lines. The blue cytoplasmic staining indicates the alkaline phosphatase activity. Scale bars are 30μm. (D) Western blot showing TAg co-immunoprecipitated with p53 or Rb protein, as indicated.

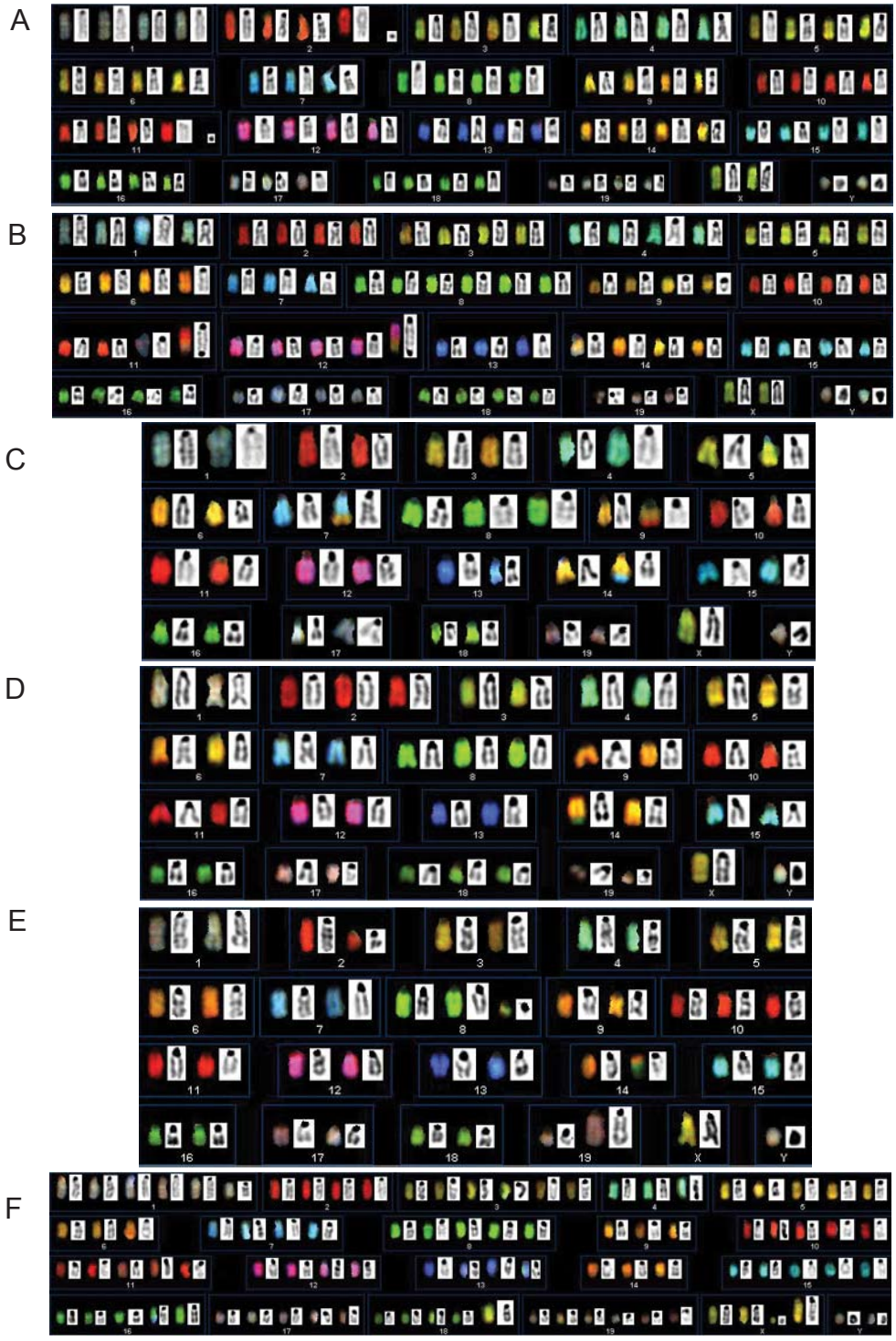

Figure S4. Karyotypic Complexity of OSA in the MOTO Model

Additional SKY analysis of MOTO OSA primary tumors and metastases. (A-B) MOTO-1 primary tumor exhibiting clonal numerical changes:,gain of chr. 8 and high frequency of non-clonal structural changes (0.8/metaphase spread). Cells were diploid (10/20 cells) and tetraploid (10/20 cells). (C-E) MOTO-1 cell line displaying aneuploidy and structural rearrangements. (F) MOTO-1 tumor metastasis displaying aneuploidy and structural rearrangements. Cells were diploid (7/10 cells) and tetraploid (3/10 cells) with 0.3 rearrangements/metaphase spread.

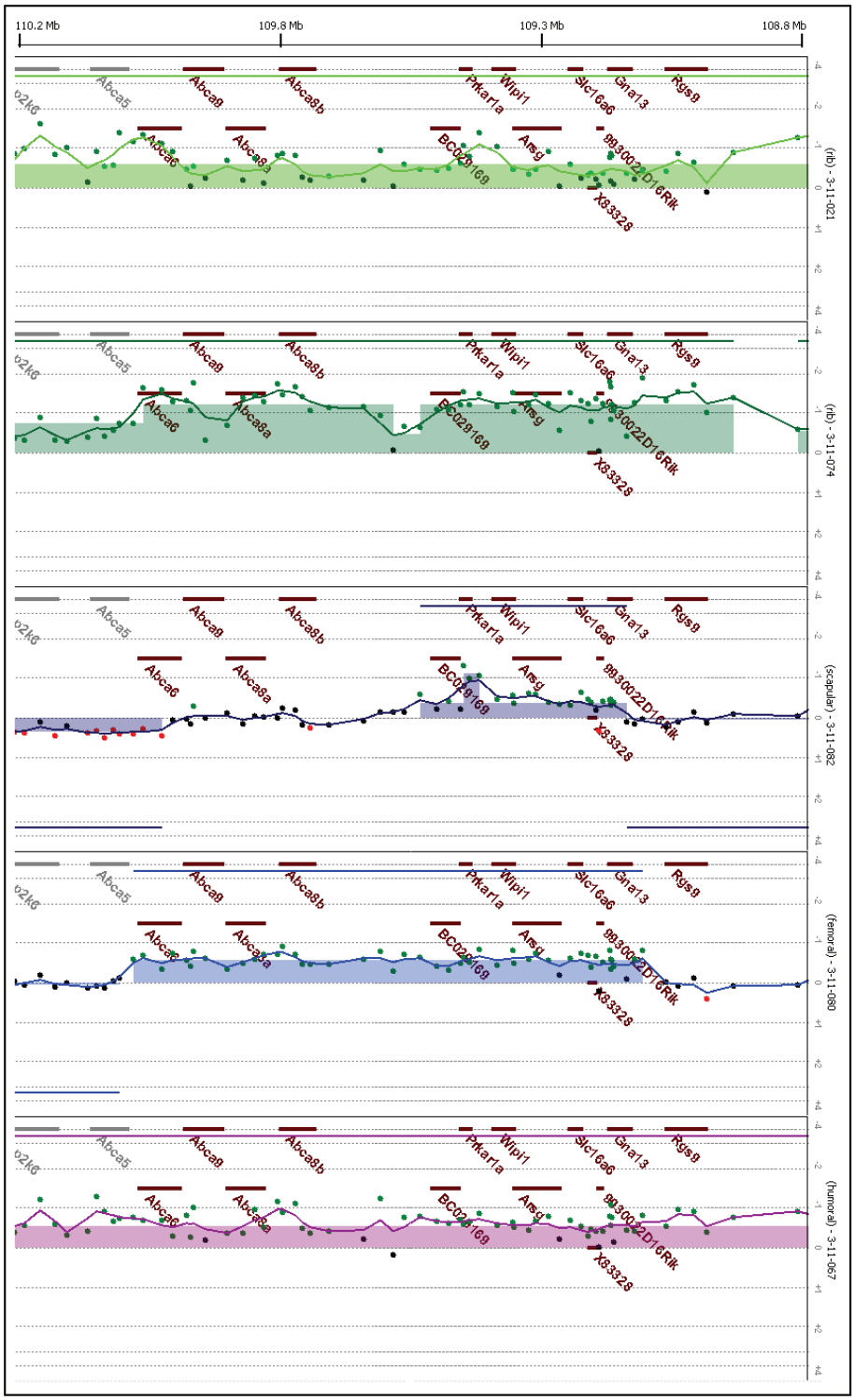

Figure S5. Recurrent Deletions at 11qE1 in MOTO Mouse Tumors

Array CGH plots at 11qE1 for individual MOTO OSA tumors bearing deletions containing *Rgs9*, *Gna13*, *9930022D16Rik*, *X83328*, *Slc16a6*, *Arsg*, *Wipi1*, *Prkar1a*, *BC029169*, *Abc8b*, *Abc8a*, *Abca9*, *Abca6*. Probes are smoothed using a 50kb moving average.

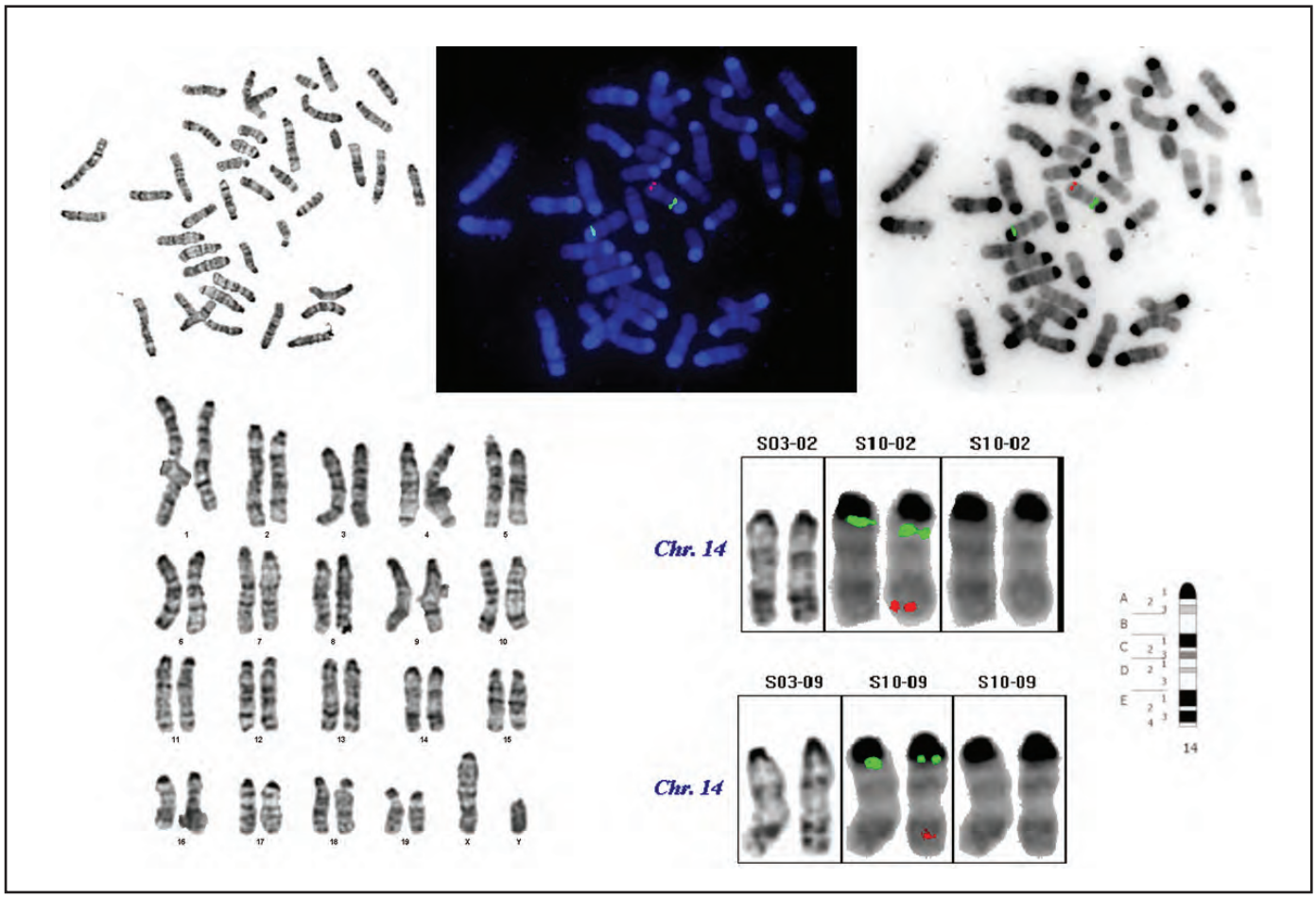

Figure S6. Cytogenetic Mapping of the Colα(I)Cre Transgene

Localization of the Colα(I)Cre transgene to chromosome 14qE2-4 using metaphase FISH analysis of mouse splenocytes and a Cre plasmid probe (red). The identity of Chr. 14 was confirmed by G-to-FISH using a control probe specific for 14qA1 (green).

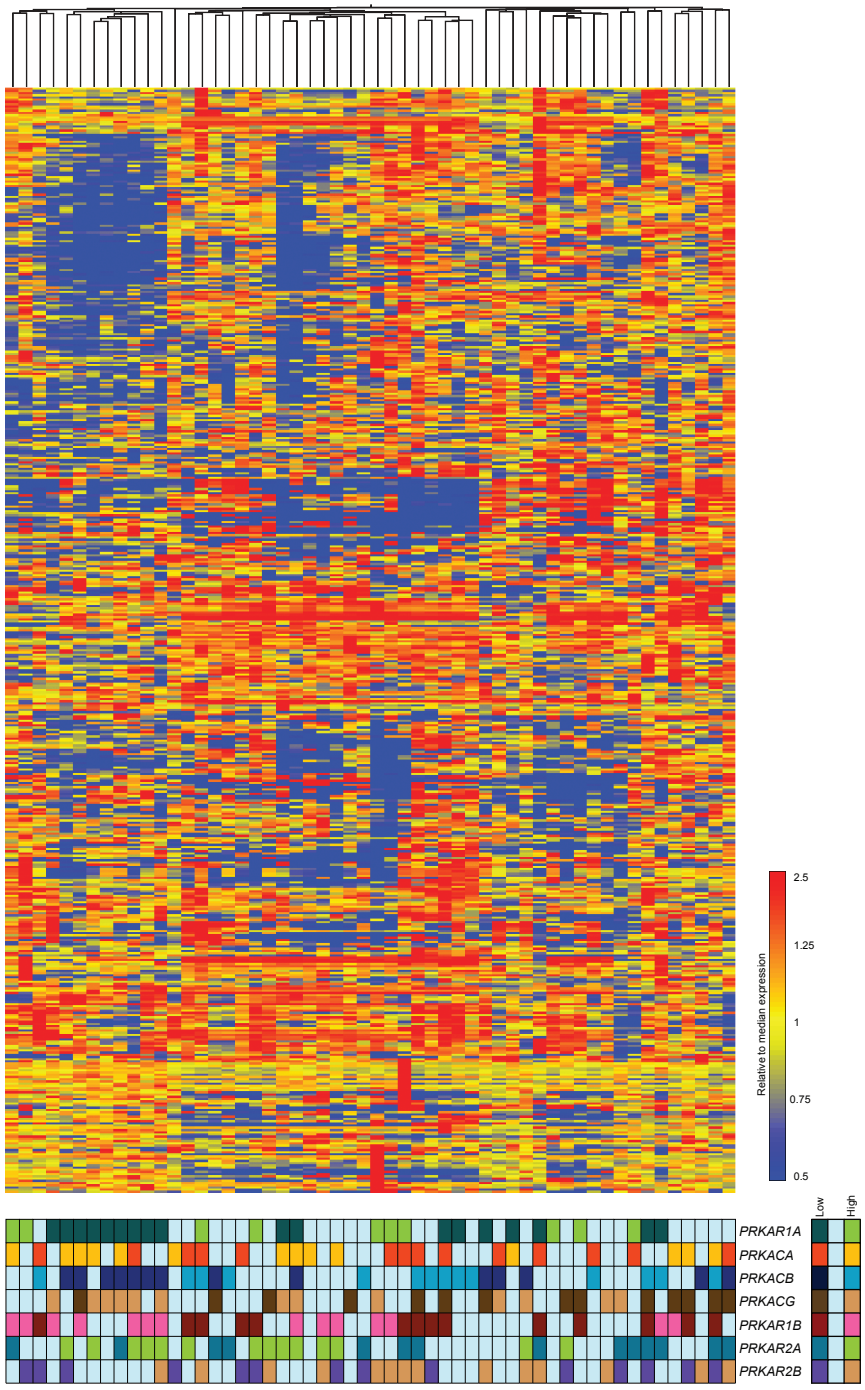

Figure S7. Unsupervised Analysis of Human OSA Expression Profiles

A subset of the cAMP/PKA/skeletal geneset with highly variant expression was used for hierarchical clustering and samples were annotated according to low or high expression for the all PKA holoenzyme gene subunits as indicated (below heatmap).

Supplemental Table 1. Incidence of spontaneous osteosarcoma metastasis in MOTO mice

	MOTO-1	*MOTO-2*
MOTO$^+$ mice (n)	99	95
Metastasis incidence	95/99	95/95
Lung	95	86
Liver	23	56
Kidney	19	19
Spleen	0	2

Supplemental Table 2. Minimal common region analysis of deletions at 11qE1 in MOTO mouse OS

Chr	*Start Index*	*End Index*	*No Of Probes*	*Start Position*	*End Position*	*Size (KB)*	*dir*	*Score*	*-log(p value)*
11	4673	4675	3*	109452262	109481052	28791	-1	-2.87	0.005583*
11	4611	4735	125	108110250	111622519	3512270	-1	-2.75	0.007125
11	4651	4735	85	109156656	111622519	2465864	-1	-2.6	0.009726
11	4611	4679	69	108110250	109559639	1449390	-1	-2.58	0.01006
11	4679	4679	1	109559239	109559639	401	-1	-2.53	0.01118
11	4641	4645	5	108742029	108875414	133386	-1	-2.52	0.01134
11	4611	4716	106	108110250	110466143	2355894	-1	-2.47	0.0126
11	4699	4701	3	110026031	110077335	51305	-1	-2.19	0.02203
11	4660	4660	1	109235585	109235985	401	1	2.17	0.02268
11	4651	4716	66	109156656	110466143	1309488	-1	-2.17	0.02287
11	4651	4701	51	109156656	110077335	920680	-1	-2.16	0.02303
11	4651	4679	29	109156656	109559639	402984	-1	-2.1	0.02617
11	4646	4700	55	108991847	110060584	1068738	-1	-2.08	0.02685
11	4646	4679	34	108991847	109559639	567793	-1	-2.02	0.03015
11	4699	4700	2	110026031	110060584	34554	-1	-1.88	0.03944
11	4702	4723	22	110102627	110825631	723005	-1	-1.84	0.04217
11	4641	4653	13	108742029	109184920	442892	-1	-1.79	0.0459

* deletion contains only Prkar1a