Oncogene (2006) 25, 6336–6344

www.nature.com/onc

Nuclear–mitochondrial genomic profiling reveals a pattern of evolution in epithelial ovarian tumor stem cells

AA Wani[1], N Sharma[1], YS Shouche and SA Bapat

National Centre for Cell Science, Ganeshkhind, Pune, Maharashtra, India

Analyses of genome orthologs in cancer on the background of tumor heterogeneity, coupled with the recent identification that the tumor propagating capacity resides within a very small fraction of cells (the tumor stem cells-TSCs), has not been achieved. Here, we describe a strategy to explore genetic drift in the mitochondrial genome accompanying varying stem cell dynamics in epithelial ovarian cancer. A major and novel outcome is the identification of a specific mutant mitochondrial DNA profile associated with the TSC lineage that is drastically different from the germ line profile. This profile, however, is often camouflaged in the primary tumor, and sometimes may not be detected even after metastases, questioning the validity of whole tumor profiling towards determining individual prognosis. Continuing mutagenesis in subsets with a mutant mitochondrial genome could result in transformation through a cooperative effect with nuclear genes – a representative example in our study is a tumor suppressor gene viz. cAMP responsive element binding binding protein. This specific profile could be a critical predisposing step undertaken by a normal stem cell to overcome a tightly regulated mutation rate and DNA repair in its evolution towards tumorigenesis. Our findings suggest that varying stem cell dynamics and mutagenesis define TSC progression that may clinically translate into increasing tumor aggression with serious implications for prognosis.
Oncogene (2006) **25,** 6336–6344. doi:10.1038/sj.onc.1209649; published online 29 May 2006

Keywords: ovarian cancer; tumor stem cells; mtDNA; *CREBBP*; mutational profiling

Introduction

Division and differentiation of a small number of stem cells in healthy tissues ensures a continuous turnover of cells and optimal organ functioning (Michor *et al*., 2004). Genetic heterogeneity exists even at a state of homeostasis within an organ, and is attributed to the clonal evolution of normal stem cell lineages (Shin *et al*., 2004, Calabrese *et al*., 2004). Recent evidence indicates that tumorigenesis is an aberrant process initiated by a rare population of transformed stem cells termed as tumor stem cells (TSCs) that maintains/reacquires the capacity for indefinite proliferation along with tumor-forming capabilities (Bonnet and Dick, 1997). Current reports of the putative isolation of TSCs (Al-Hajj *et al*., 2003, Singh *et al*., 2003) are based on the differential expression of histologic or surface markers and self-renewal mechanisms during tumor maintenance and progression. Although their critical importance is realized, much remains to be learned about the genetic mechanisms leading to the emergence of a TSC from a normal stem cell, yet accounting for tumor heterogeneity (Florian *et al*., 2006).

On this background, we asked whether it would be possible to (i) trace the genetic heterogeneity between various stem cell lineages within a tumor, and (ii), understand the molecular differences between normal and tumor cells that could possibly define specific events within a stem cell lineage that places it on a trajectory towards tumorigenesis. Specifically, we postulated that if the heterogeneity within tumors were to be a cumulative effect of genetic drift and amplification of specific stem cell lineages, one would need to define a mechanism for tracing all these in order to resolve their effects at the molecular and cellular level. Mitochondrial genome (mtDNA) analyses, currently applied to study evolutionary history, population migration, forensic medicine (Pakendorf and Stoneking, 2005) and human disease (O'Brien *et al*., 2005), have been recently applied to model stem cell turnover rates and clonal evolution in normal tissues; hence was thought to be an ideal tool for this study. Among the nuclear genes, the cyclic AMP response element-binding protein (*CREBBP*) is involved in multiple cellular processes, functions as a transcriptional cofactor and is also a histone acetyltransferase (HAT) (Petrij *et al*., 1995). Germline mutations in *CREBBP* result in Rubinstein–Taybi syndrome (Kitabayashi *et al*., 2001), that is characterized by an increased predisposition to cancer; further validated by the observation that $CREBBP^{+/-}$ mice express an increased frequency of hematopoietic malignancies (Kung *et al*., 1999). Several truncating mutations have been identified in *CREBBP* in breast, colon and pancreatic cancer cell lines and primary

Correspondence: Dr SA Bapat, Lab 4, National Centre for Cell Science, NCCS Complex, Pune University Campus, Ganeshkhind, Pune, Maharashtra 411007, India.
E-mail: sabapat@nccs.res.in
[1]These authors contributed equally to this work.
Received 7 February 2006; revised 16 March 2006; accepted 16 March 2006; published online 29 May 2006

tumors, leading to loss or mutation of one allele (Costanzo *et al*., 2002).

In the present study, we report the mutational profiling of the mitochondrial genome and nuclear *CREBBP* gene that led to the identification of a unique mitochondrial–nuclear profile defining the TSC population within a tumor.

Results

Epithelial ovarian cancer is associated with a high incidence of mitochondrial mutations of which three show a statistically significant association

The study was initiated with amplification and sequencing of the entire mitochondrial genome from 12 primary epithelial ovarian tumor samples (Supplementary Table 1), ensuring identification of all occurring nucleotide changes. Further, large polymerase chain reaction (PCR) products excluded the possibility that nuclear pseudogenes could complicate the analysis (Parfait *et al*., 1998). The analyses of these tumors led to the identification of a total of 170 nucleotide variations (Supplementary Table 2) that showed interesting distribution patterns. Their distribution was throughout the mitochondrial genome, viz. in the non-coding (control) as well as protein-coding regions; in the latter case, both synonymous (silent) as well as non-synonymous changes were evident. Of these variations, 30 were entirely novel. Very strikingly, the mitochondrial sequences in this study (tumor samples as well as controls) with the Indian population revealed a unique profile of eight sequence variants, viz. A73G, A263G, A1438G, A2706G, A4769G, C7028T, A8860G and A15326G appeared at high frequencies in all samples and could be of evolutionary significance. Although the individual polymorphisms within this profile are described elsewhere (MITOMAP: A Human Mitochondrial Genome Database & Human Mitochondrial Genome Database), based on their high recurrence, we derived that the profile could be specific to the Indian population.

Subjecting the mutational data to a two-tailed χ^2 analyses indicated 64 of the total mitochondrial nucleotide variations to be of moderate significance (Supplementary Table 3), of which, three mutations, viz. A10398G, 523insCA and 523delA (Table 1) expressed high degree of correlation with the primary tumors ($n=12$) than the control samples ($n=106$) (Palanichamy *et al*., 2004), thereby qualifying as putative ovarian TSC markers.

Table 1 Putative mtDNA mutation candidates in ovarian cancer

mtDNA mutation	*A10398G*	*523ins CA*	*523delA*
Present (healthy)	27	4	8
Present (Cancer)	8	3	4
Absent (healthy)	79	102	98
Absent (Cancer)	4	9	8
χ^2	8.768218	8.70365	7.84613
Proportion present (healthy)	0.254717	0.037736	0.075472
Proportion present (cancer)	0.666667	0.25	0.333333
P-value	0.003	0.003	0.005

Abbreviation: mtDNA, mitochondrial DNA.

Analyses of single-cell clones reveal the true nature of mitochondrial profiles within a tumor

The above results suggested that a tumor may contain a multitude of stem cell lineages; yet, whether the dominant expression in the tumor represents the TSC population is still uncertain. To resolve this, our further studies were with a unique model developed by us earlier (Bapat *et al*., 2005 – described in Materials and methods). This mtDNA analyses, indeed, provided a magnified insight into the system. The primary tumor cells and ascites-derived cells expressed identical mtDNA profiles (Table 2), consisting of 37 variations distributed throughout the genome in the protein coding as well as non-coding genes; with synonymous as well as non-synonymous changes being identified. Six of these changes were found to be novel. The same mutational profile was also associated with 14 (of 19) of the immortalized clones as well as the unsorted M1 and M2 cultures (Figure 1a). We further conducted a chase–back study towards tracing the origin of these variations. Normal tissue sample of the same patient was unavailable; hence, a peripheral blood sample from the son was obtained – valid as mtDNA is almost exclusively maternally inherited (Hayashida *et al*., 2005). Comparison of profiles revealed only one mtDNA sequence variation between the son and the tumor cells of the patient (310 insC instead of 310 insCC – Table 2). This suggests that all the variations detected in these samples could be germ line polymorphisms. A marked feature in the immortalized cells from this profile was the acquisition of two mutations (G7393A in *CoxI* gene: amino acid (a.a.) change from glycine to glutamic acid and C16147T in D-loop region) that were absent in the primary cells. These variations are indicative of mtDNA mutagenesis resulting from adaptation in culture, also known to be associated with chromosomal instability and altered *in vitro* ploidy levels (Schmid *et al*., 2004).

Identification of a specific mitochondrial mutation profile associated with the tumor stem cell population

A major finding of this study is the identification of a variant mtDNA profile in a small group of five clones (viz. A2, A3, A4, B2 and C4; Figure 1a). This group includes the two tumorigenic clones A2 and A4, indicating that the profile defines the TSC lineage as a distinct identity among others within the tumor. The mutant profile is characterized by 12 distinct sequence variations and seven common ones within the germ line profile (Figure 1b). A majority of nucleotide variations in this profile were in the non-coding regions (D-loop:8; 16sRNA:1; tRNA cysteine:1), the one in the *ATPase synthase* gene (C8410T) was silent (no change in a.a.), whereas that in the *cytochrome b* gene (T14766C) was non-synonymous (isoleucine to threonine). Acquisition of the mutant profile was accompanied by reversal of 32

Table 2 mtDNA variations in OT29 germline profile

S. No	*Position*	*Gene*	*a.a. change*	*Interspecies conservation (%)*	*Novel*
1	T152C	D-loop	Non-coding	NA	No
2	A263G	D-loop	Non-coding	NA	No
3	A1438G	12S RNA	Non-coding	NA	No
4	A2706G	16S	Non-coding	NA	No
5	A4769G	ND2	Silent	NA	No
6	A8860G	ATP Synthase 6	Threonine–Alanine	50	No
7	A15326G	Cytochrome *b*	Threonine–Alanine	37.50	No
8	A73G	D-loop	Non-coding	NA	No
9	310 (Ins-CC)[a]	D-loop	Non-coding	NA	Yes
10	G329A	D-loop	Non-coding	NA	No
11	T489C	D-loop	Non-coding	NA	No
12	573 (Ins-C)	D-loop	Non-coding	NA	Yes
13	G3483A	ND1	Silent	NA	No
14	C4020T	ND1	Silent	NA	Yes
15	T4216C	ND1	Tyrosine–Histidine	25	No
16	C5354T	ND2	Silent	NA	Yes
17	C7028T	COXI	Silent	NA	No
18	T7278G	COXI	Phenylalanine–Valine	75	Yes
19	T7678C	COXII	Silent	NA	No
20	A8701G	ATP Synthase 6	Threonine–Alanine	62.50	No
21	A9446G	COXIII	Silent	NA	Yes
22	T9540C	COXIII	Silent	NA	No
23	G9932A	COXIII	Silent	NA	No
24	A10398G	ND3	Threonine–Alanine	62.50	No
25	C10400T	ND3	Silent	NA	No
26	T10873C	ND4	Silent	NA	No
27	A11101G	ND4	Silent	NA	No
28	G11719A	ND4	Silent	NA	No
29	C12705T	ND5	Silent	NA	No
30	A13651G	ND5	Threonine–Alanine	50	No
31	T14783C	Cytochrome *b*	Silent	NA	No
32	G15043A	Cytochrome *b*	Silent	NA	No
33	G15301A	Cytochrome *b*	Silent	NA	No
34	T15313C	Cytochrome *b*	Silent	NA	No
35	C16111T	D-loop	Non-coding	NA	No
36	C16223T	D-loop	Non-coding	NA	No
37	T16311C	D-loop	Non-coding	NA	No

Abbreviations: a.a., amino acid; NA, not applicable. [a]-310 ins-C - present in the son.

variations expressed in the germ line profile (to reported sequences in the CRS – Cambridge Reference Sequence). The heteroplasmic state of some of these variations (data not shown) suggests that acquisition of the mutant profile and loss of some of the wild-type polymorphisms is a continuous ongoing process that ultimately culminates in fixing of mutations to a homoplasmic state (Coller *et al*., 2005).

Sublineage demarcation within the germ line and mutant mitochondrial profiles

Within the wild-type/germ line profile clones, further lineage demarcation was evident. Six clones (E1, F3, F4, G5, G7 and H1) express the germ line profile that may be surmised to define at least one stem cell lineage, six others (A5, B7, E2, F1, F2 and F5) show a variance through acquisition of an additional mutation – G6888R in the *CoxI* gene resulting in a termination codon. The heteroplasmic expression of this mutation probably ensures clone survival; it can be speculated that further mutation fixing to a homoplasmic state could be lethal and result in extinction of the clone. The remaining two wild-type clones (D4 and H4) express another heteroplasmic mutation in the tRNA proline gene – G15959R, defining yet another lineage/subline emerging from the common, wild-type profile.

Within the mutant profile lineage, demarcation into three different sublineages is also possible despite their similar mitochondrial profiles. A2 is functionally distinct owing to its tumorigenic potential. A4 cells acquired an additional heteroplasmic, non-synonymous mutation (C12700T) in the *ND5* gene: leucine to isoleucine, that coincided with acquisition of tumorigenecity, whereas the remaining three clones (A3, B2 and C4) retained the profile with no change either in mtDNA sequence or tumorigenecity. An observation with the heteroplasmic C12700T mutation coinciding with acquisition of tumorigenecity was that it was not fixed to homoplasmy even after propagating the tumor sequentially in nude mice for two further generations.

cAMP responsive element binding binding protein exon mutations in the mutant mitochondrial DNA profile

From the above analyses in the mutant mtDNA profile clones, it became quite evident that acquisition of a particular mutant mitochondrial profile, although

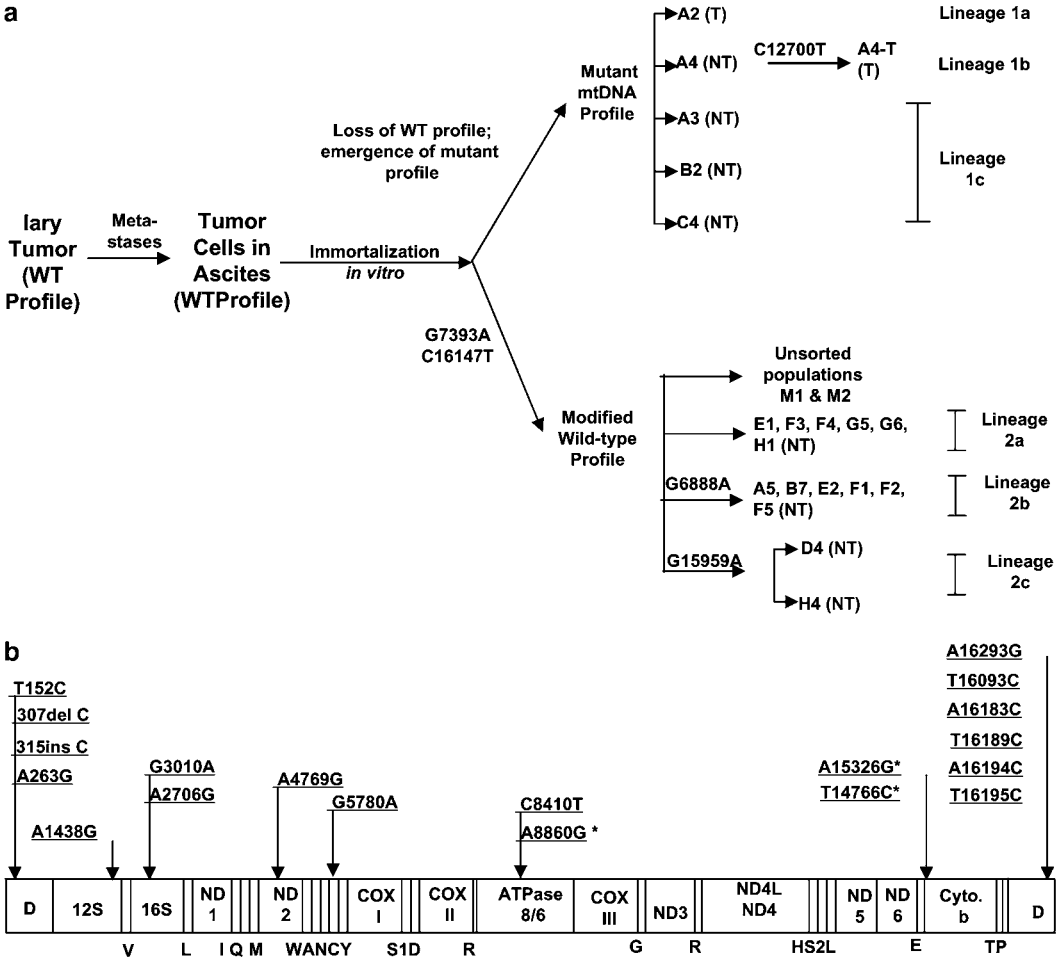

Figure 1 (**a**) Schematic representation of tumor evolution in OT29 samples. The primary and metastatic ascites both express wild-type mitochondrial DNA profiles. The 19 immortalized single-cell clones isolated from a primary ovarian tumor separate into two major groups: Group 1 consists of five clones, of which one (A2) was tumorogenic (Lineage 1a); another (A4) underwent a spontaneous transformation and acquired one heteroplasmic mutation (12700C > T) in *ND5* gene (Lineage 1b) whereas the other remained untransformed (A3, B2 and C4 – Lineage 1c). Group2: too diverged into three lineages – Lineage 2a consisted of six clones (wild-type profile); Lineage 2b expressed one additional heteroplasmic mutation (G6888A) in the *CoxI* gene and Lineage 2c another heteroplasmic transition (G15959A) in the *tRNA proline* gene (see Supplementary Figure 1). (**b**) Distribution of (12 + 7) mutant sequence variations marked on a representative linearized mitochondrial genome; *Sequence variations causing non-synonymous amino-acid changes in coding genes.

highly implicative of a signatorial association with the TSCs, cannot be the sole determinant of tumorigenecity. It has been earlier suggested that complementation with specific mutations in nuclear genes (Singh *et al*., 2005) or epigenetic regulation (Feinberg *et al*., 2006) could be the next step required to propel a stem cell towards a tumorigenic program. To resolve this, we carried out mutational analyses *CREBBP* in the five clones expressing the mutant mtDNA profile, the primary tumor and ascites-derived cells and one representative clone expressing the germ line mtDNA profile, viz. H4. This led to the detection of 23 novel mutations in the *CREBBP* exon sequences, none of which were common to all the samples, although some overlap was evident within smaller groups (Figure 2a). Significantly, no exon mutations were evident in the H4 genome that retained its wild-type *CREBBP* profile, whereas both the primary samples (tumor and ascites-derived cells) expressed three identical mutations. Within the five mtDNA mutant profile clones, four mutations in the nuclear hormone receptor domain (NHRD) of the gene, eight in the CREB binding domain (KIX), three in exons 12–17 (which have no known functional domains), three in the bromodomain and two in the HAT domain, were identified (Figure 2a). The tumorigenic clones A2 and A4-T showed one identical non-synonymous mutation in the HAT domain (asparagine to isoleucine); an identical non-synonymous mutation (aspartate to asparagine) was also detected in this domain in the non-tumorigenic clones B2 and C4. These mutations could be significant as the HAT domain mediates a key acetylation function of the protein and triggers off several downstream signal transduction pathways (Goodman and Smolik, 2000). In this particular

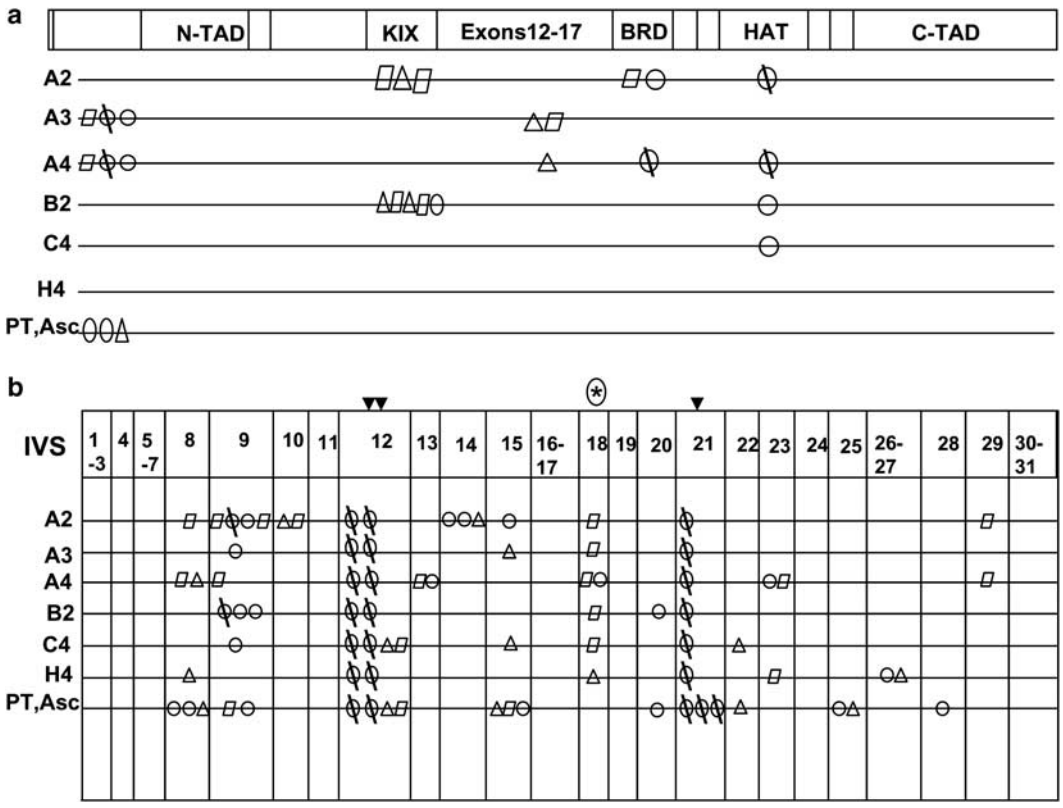

Figure 2 Representation of genetic alterations in the nuclear cAMP responsive element binding binding protein (*CREBBP*) gene. (**a**) Mutations in the coding region of the *CREBBP* gene. Functional domains are represented as boxes and include: NTAD (NH_2-terminal transactivation domain), KIX (CREB-binding domain). BRD (Bromodomain), HAT (histone acetyltrasferase domain) and CTAD (COOH terminal transactivation domain), A2, A3, A4-T, B2 and C4 are the mutant clones; H4 is a representative of the mitochondrial wild-type clone; PT is the primary tumor and Asc is the ascites-derived tumor cells. (**b**) Mutations in the non-coding region of the *CREBBP* gene. Introns are represented in boxes: various mutations are indicated by different icons those common to all samples are indicated as ▼, a putative hotspot common to the mutant clones is represented as *.

instance, the presence of a specific mutation coincidentally in the two tumorigenic/mtDNA mutant group clones is indicative of an altered protein with implied association of a predisposition or definitive role in tumorigenity. Mutations in the bromodomain were also expressed exclusively by the two tumorigenic clones. This conserved domain has a specific role during histone acetylation (Deng *et al*., 2003), binds specifically to p53 (Dhalluin *et al*., 1999) and is responsible for p53 acetylation-dependent co-activator recruitment during signaling (Grossman, 2001). Clones A3 (non-tumorigenic) and A4-T (tumorigenic) expressed two identical and one differential mutation each in the NHRD domain; one differential mutation was seen in each of the A2 (tumorigenic) and B2 (non-tumorigenic) clones in the KIX domain (KIX is the region where CREB binds and mediates the genomic effects of cAMP (Murata *et al*., 2001). The mutations in A2 and A4-T in HAT and bromodomains are suggestive of being responsible for the tumorigenic functions of these clones; the differential mutations between these clones in the NHRD and KIX domain could contribute to the qualitative differences during tumor formation and progression from these clones in animal models (A4-T has been shown earlier to be a more aggressive clone than the A2 (Bapat *et al*., 2005).

cAMP responsive element binding binding protein intron/exon boundary mutation analyses

As in case with exon mutations, both the primary samples showed identical intron-exon boundary (I/Eb) mutation profiles, whereas three nucleotide variations were found to be common to all the samples (Figure 2b), and may be considered as polymorphisms. Each single-cell clone expressed a distinct profile with some mutations in common with other clones – in various permutations and combinations. The primary tumor and tumor-derived ascites samples expressed one variant in the donor site of intron 9 whereas H4 genome expressed a variant in the acceptor splice site of intron 18 (Table 3). These variations may not be significant in view of tumorigenecity as the cells were not seen to be tumorigenic despite the presence of the mutations. The most striking intronic mutation was expressed in the five mutant clones within a polyA tract (A13) located at the tail end of intron 18 (IVS18). A successive deletion was expressed at this site in the five clones, with B2 (non-tumorogenic) having a single adenine deletion (delA), C4 and A3 (non-tumorogenic) showing an AA deletion and the two tumorogenic clones showing a triple AAA deletion (Figure 3a, Table 3). These successive mutations, due to an attribute of being a part of the normal acceptor splice site, could

Table 3 CREBBP Splice Site mutations

Intron	*Mutation*	*Part of splice site*	*Donor/acceptor splice site*	*Clones in which present*
IVS9	IVS9−3A > T	Yes	Donor splice site	Primary tumor and ascites cells
IVS13	IVS13 + 10A > T	Yes	Acceptor splice site	A4
IVS15	IVS15−141G > A	No	−4 bp from Acceptor splice Site	Primary tumor and ascites cells
IVS18	IVS18−3A > T	Yes	Acceptor splice site	A4
IVS18	IVS18−2A > T	Yes	Acceptor splice site	H4
IVS18	IVS18−13DELA	Yes	Acceptor splice site	B2
IVS18	IVS18−13DELAA	Yes	Acceptor splice site	A3,C4
IVS18	IVS18−13DELAAA	Yes	Acceptor splice site	A2,A4
IVS22	IVS22 + 117C > G	No	Acceptor splice site(2 bp downstream)	Primary tumor and ascites cells
IVS22	IVS22 + 120T > G	No	Acceptor splice site(5 bp downstream)	Primary tumor and ascites cells
IVS25	T > G	Yes	Donor splice site(1 bp upstream)	C4

Abbreviation: CREBBP, cAMP responsive element binding binding protein.

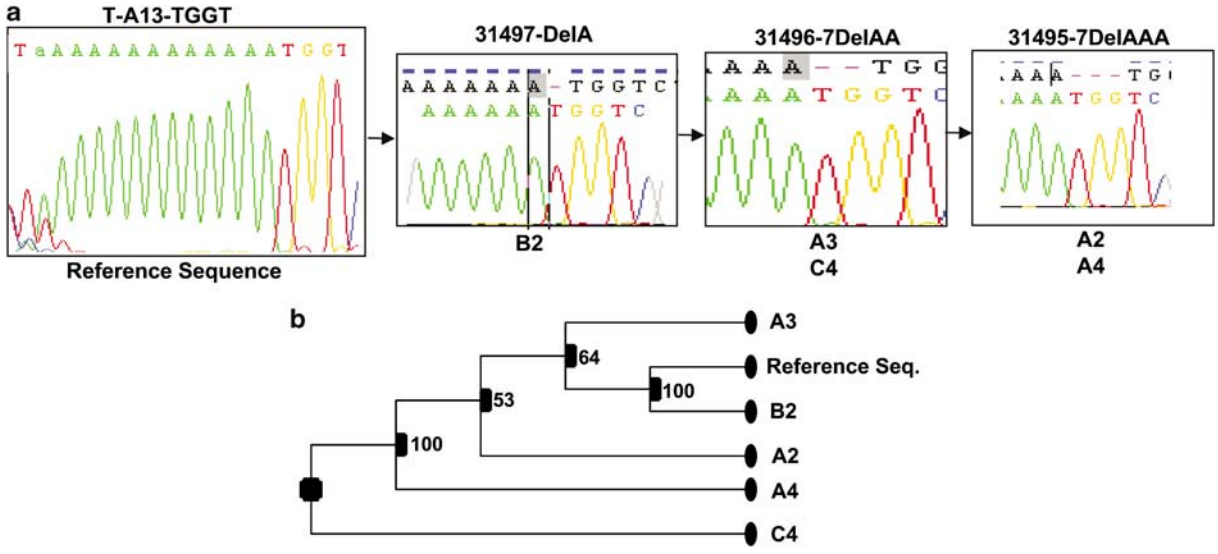

Figure 3 (**a**) Electropheretograms depicting a putative hotspot at the intron 17/exon18 boundary in the mutant clones. (**b**) Phylogenetic tree obtained using PHYLIP version 3.61 software, showing the relationship among the mutant clones viz A2, A3, A4, B2, C4 with reference *CREBBP* sequence obtained from NCBI. Two different methods (KITSCH and UPGMA) were used independently on a distance matrix calculated by DNA-DIST and yielded the same pattern of clustering in both cases. Bootstrap values are placed above branches.

have some significance in splicing; alternatively, they may serve as a definite marker for delineating the mutant clones within the tumor or cause dosage effects of the gene. The tumorigenic clone A4 expressed two more splice site variants in the acceptor site of introns 13 and 18, respectively. The latter may complement the earlier mutation (del AAA in IVS18), contributing further to the differential tumorigenecity between the A2 and A4-T clones. Similar kind of intronic mutations causing splicing defects has been reported earlier in several cancers; a comparable study in leukemia associates such mutations with altered levels of HAT activity of *CREBBP* (Shigeno *et al*., 2004).

Phylogenetic divergence within the mutant clones

The above analyses cumulatively indicate that after the mutant mitochondrial profile group diverges out from a germ line mtDNA profile, each individual clone within the group has the potential to become a TSC. To understand why only a few clones of the mutant group acquire this capability, we performed a phylogenetic analysis to resolve the distribution coexistence within the different clones (Figure 3b). The B2 clone (non-tumorigenic) clustered with the reference *CREBBP* gene at 100% homology; A3 clone branched out from this cluster with 64% homology. A2 and A4-T (tumorigenic) further branched out and expressed 53% similarity with the (B2, reference sequence and A3) cluster, but shared 100% homology with each other. C4 (non-tumorigenic) stood out as an outgroup in the tree. Thus, one can predict that A2 and A4 are highly similar, align on the same branch and are tumorigenic. On the other hand, B2 and A3 have only a partial degree of homology with the tumorigenic clones and are closer to the reference *CREBBP* sequence. Lastly, C4 shows

minimum homology and is least related with any other clone in the group yet has significantly evolved away from the reference sequence.

Discussion

Most tumors, at the stage of detection, are a heterogeneous mixture of several subclones that makes it difficult to establish the order of the genetic insults (van Tilborg *et al*., 2000). The approach in our study aimed at resolving such issues relating to the evolution of TSCs. The three mtDNA mutations identified in the primary samples were suggested to be highly significant in epithelial ovarian cancer; yet were actually absent in the TSC clones – although one (A10398G) was expressed in the normal germ line profile of the patient. Thereby, although TSCs have a distinct genetic identity within the gross tumor as do a multitude of other non-transformed stem cell lineages, the overall dominant expression would be germ line (at initial stages of tumorigenesis) and mixed mutant-germ line at later stages (due to overlap of expressions of dominating clones in the tumor; Sidransky *et al*., 1992). This suggests that genetic profiling of entire tumors could often be a confusing exercise and possibly lead to errors in tumor classification (Alonso *et al*., 2005).

A further finding is that, although TSCs have an evasive yet distinctly identifiable mtDNA profile, acquisition of such a mutant genome cannot be the sole determinant of tumorigenecity. Different nuclear backgrounds and an identical mitochondrial genome in clones with varying levels of tumorigenecity is a distinct indication that mitochondrial mutations do not initiate the tumorigenic process. However, previous observations with cancer cells expressing altered metabolic patterns indicate the possibility that some of the variations in the specific mutant mtDNA profile may alter mitochondrial energy conversion efficiencies in those particular lineages by increasing the expression of genes involved in glycolysis and reducing Kreb's cycle activity (Costanzo *et al*., 2002). Negative clone selection also occurs; some of the germ line profile lineages may thus be eliminated during tumor progression, for example, the lineage that accumulates a mutation leading to a termination codon in the *CoxI* gene. Thus, concurrent to the positive selection, a continuous turnover of stem cell populations through neutral evolution of untransformed stem cell lineages during tumor progression can occur.

What exactly defines the events within a lineage on a trajectory towards tumorigenesis? Our study clearly shows that TSCs have a distinctive mitochondrial–nuclear signature within a tumor, and gives a rare glimpse of changing mutational patterns accompanying varying stem cell dynamics and turnover within the organ during tumor evolution. It has been realized that the classically defined 'two hits' (first germ line and second somatic) produce only a benign precursor lesion and that additional events are necessary for transformation (Knudson, 1996; Beckman and Loeb, 2005). Based on our findings, the likely cascade of events could be

(i) Alterations in the mitochondrial genome: at present, it is not clear whether this altered pattern gives a selective advantage, or only leads to increased reactive oxygen species production. Further, the emergence of a mutant mitochondrial profile is either as a consequence of increased mutagenesis or differential segregation of mutant mtDNA copies in the cell. Acquisition of the mutant profile however, although not capable of mediating transformation, may represent a mutator phenotype that could in addition to the biochemical effects, accelerate the rate of mutations in the mutant clones, whereas some of the others lead to negative clonal selection.
(ii) The second target is the nuclear genome wherein neutral evolution and mutagenesis in various genes (exons as well as introns) leads to the currently prevalent caveat of tumor suppressor inactivation and oncogene activation leading to the emergence of a TSC.
(iii) Further evolution may not be entirely neutral, but could be determined by extrinsic factors. Chemotherapy can eliminate the initial TSC population along with a majority of the tumor cells. However, the genetic variation at the nuclear level within the mutant group clones may ensure the survival of some of these populations that have a predilection towards transformation. Such a situation could lead to tumor recurrence, which at the clinical level is often more aggressive than the primary disease.

Our findings will thus be critical in addressing several themes such as expression analyses of tumors, surface antigens that may be used for developing targeted antibodies and drug discovery towards the development of chemotherapy for eliminating the entire mutant group including the TSCs. This would lead to better optimization of therapeutic regimes for ovarian cancer. Further developments on the identification of continuing mutagenesis in subsets with the mutant phenotype could also provide a mechanism for monitoring minimal residual disease.

Materials and methods

Primary tissue samples and cell cultures

The present study is approved by the Bio-Ethics Committee of NCCS; informed consent was obtained from all patients. Frozen samples of primary ovarian serous adenocarcinomas (OT6, OT13, OT14, OT17, OT19), papillary adenocarcinoma (OT9) and cystadenomas (OT10, OT22) and three samples of tumor cells from the ascites of patients with metastatic serous adenocarcinomas (OASC1, OASC2 and OASC3) were retrieved from our tissue bank and used for DNA analysis. Thirty normal ethnic controls belonging to the same haplogroup as the patients and 75 downloaded sequences covering almost all known Indian haplogroups (Rieder *et al*., 1998) were also used for comparison (data not shown).

In the second part of the study, we used an *in vitro* model developed earlier by us that comprises 19 single-cell clones isolated from the ascites of an ovarian serous adenocarcinoma patient (Bapat *et al*., 2005). Of these, one clone was tumorigenic (A2); another one (A4) underwent a spontaneous transformation *in vitro*, whereas the remaining 17 clones were non-tumorigenic. In addition, following samples from the same patient were also used

(i) Paired sample of cryopreserved primary tumor and cells isolated from ascites,
(ii) two immortalized, unsorted (heterogenous), non-tumorigenic populations designated as M1 and M2 that were derived from the ascites cultures and
(iii) a control sample that comprised of primary MNCs isolated from a periferal blood sample of the son of the patient.

Polymerase chain reaction amplification, DNA sequencing and mutation analysis

Genomic DNA was extracted using DNeasy tissue kit (Qiagen, Germany). The entire mitochondrial genome was amplified in 24 separate reactions using overlapping primer pairs as described earlier (Rieder *et al*., 1998). The use of large PCR products excluded the possibility that nuclear pseudogenes will complicate the analysis (Choi *et al*., 2000). The *CREBBP* gene was amplified as above, using 26 sets of primers designed at the exon–intron boundaries from the five clones expressing a mutant profile, one representative clone of the germ line profile, viz. H4 and the primary cells derived from the tumor and the ascites. For more details see Supplementary Information.

Mitochondrial gene and a.a. maps for the mtDNA-encoded proteins used in the analysis were from the MITOMAP: A Human Mitochondrial Genome Database; with additional data for the mutation and polymorphism analysis from the Human Mitochondrial Genome Database (mtDB) website maintained by Uppsala University. Comparison of mutations found in the different segmentation categories was with the use of Bayesian confidence intervals (CIs) and non-informative priors. Confidence intervals including zero indicate the difference not to be statistically significant. *CREBBP* variant sequences obtained were applied for constructing phylogenetic trees based on evolutionary distances using the neighbor joining method implemented through NEIGHBOUR (DNADIST) from the PHYLIP version 3.61 package (Andrews *et al*., 1999). A total of 1000 bootstrap value replicates resembling data sets were generated using the programs SEQBOOT to build a consensus tree. Tree files were viewed in the PHYLODRAW program (Felsenstein, 1993).

Statistical analysis

Two-tailed χ^2 analysis was carried out for determining the significance of mtDNA mutations in the primary tumor samples ($n=12$) in comparison with controls ($n=106$). The α-value for the test was 0.005.

Abbreviations

CREBBP, cAMP responsive element binding binding protein; TSC, tumor stem cells; mtDNA, mitochondrial DNA; ROS, reactive oxygen species; AMP, adenosine mono-phosphate; CRS, Cambridge Reference Sequence; NHRD, nuclear hormone receptor domain; KIX, CREB binding domain; HAT, histone acetyltransferase.

Acknowledgements

We thank Dr GC Mishra, Director, National Center for Cell Science (Pune, India) for encouragement and support. This work is funded by the Department of Biotechnology (DBT). Ms N Sharma receives a research fellowship from the Council of Scientific and Industrial Research (CSIR). We also thank Dr CB Koppikar (Jehangir Hospital, Pune, India) and Dr Sanjay Gupte (Gupte Hospital, Pune) for providing the tissue and tumor samples, Mr Sarang Satoor for the backup in DNA sequencing and Mr AM Mali for excellent technical assistance. Statistical analysis was carried out under the kind guidance of Dr AP Gore (Department of Statistics, Pune University, India).

References

Al-Hajj M, Wicha MS, Benito-Hernandez A, Morrison SJ, Clarke MF. (2003). *Proc Natl Acad Sci USA* **100**: 3983–3988.
Alonso A, Alves C, Suarez-Mier MP, Albarran C, Pereira L, Fernandez de Simon L *et al*. (2005). *J Clin Pathol* **58**: 83–86.
Andrews RM, Kubacka I, Chinnery PF, Lightowlers RN, Turnbull DM, Howell N. (1999). *Nat Genet* **23**: 147.
Bapat SA, Mali AM, Koppikar CB, Kurrey NK. (2005). *Cancer Res* **65**: 3025–3029.
Beckman RA, Loeb LA. (2005). *Genetics* **171**: 2123–2131.
Bonnet D, Dick JE. (1997). *Nat Med* **3**: 730–737.
Calabrese P, Tavare S, Shibata D. (2004). *Am J Pathol* **164**: 1337–1346.
Choi J, Jung H, Kim H, Cho H. (2000). *Bioinformatics* **16**: 1056–1058.
Coller HA, Khrapko K, Herrero-Jimenez P, Vatland JA, Li-Sucholeiki XC, Thilly WG. (2005). *Mutat Res* **578**: 256–271.
Costanzo A, Merlo P, Pediconi N, Fulco M, Sartorelli V, Cole PA *et al*. (2002). *Mol Cell* **9**: 175–186.
Deng Z, Chen C, Chamberlin M, Lu F, Blobel GA, Speicher D *et al*. (2003). *Mol Cell Biol* **23**: 2633–2644.
Dhalluin C, Carlson JE, Zeng L, He C, Aggarwal AK, Zhou MM. (1999). *Nature* **399**: 491–496.
Feinberg AP, Ohlsson R, Henikoff S. (2006). *Nat Rev Genet* **7**: 21–33.
Felsenstein J. (1993). *Phylip – Phylogeny Inference Package. Version 3.5*. University of Washington: Seattle, WA.
Florian S, Sonneck K, Hauswirth AW, Krauth MT, Schernthaner GH, Sperr WR *et al*. (2006). *Leukemia Lymphoma* **47**: 207–222.
Goodman RH, Smolik S. (2000). *Genes Dev* **14**: 1553–1577.
Grossman SR. (2001). *Eur J Biochem* **268**: 2773–2778.
Hayashida K, Omagari K, Masuda J, Hazama H, Kadokawa Y, Ohba K *et al*. (2005). *Cell Biol Int* **29**: 472–481.
Human Mitochondrial Genome Database http://www.genpat.uu.se/mtDB.
Kitabayashi Y, Aikawa LA, Nguyen A, Ohki YM. (2001). *EMBO J* **20**: 7184–7196.
Knudson AG. (1996). *J Cancer Res Clin Oncol* **122**: 135–140.
Kung AL, Rebel VI, Bronson RT, Ch'ng LE, Sieff CA, Livingston DM *et al*. (1999). *Genes Dev* **14**: 272–277.
Michor F, Iwasa Y, Nowak MA. (2004). *Nat Rev Cancer* **4**: 197–205.

MITOMAP (2005). *A Human Mitochondrial Genome Database* http://www.mitomap.org.

Murata T, Kurokawa R, Krones A, Tatsumi K, Ishii M, Taki T *et al*. (2001). *Hum Mol Genet* **10**: 1071–1076.

O'Brien TW, O'Brien BJ, Norman RA. (2005). *Gene* **354**: 147–151.

Pakendorf B, Stoneking M. (2005). *Ann Rev Genomics Hum Genet* **6**: 165–183.

Palanichamy MG, Sun C, Agrawal S, Bandelt HJ, Kong QP, Khan F *et al*. (2004). *Am J Hum Genet* **75**: 966–978.

Parfait B, Rustin P, Munnich A, Rotig A. (1998). *Biochem Biophys Res Commun* **247**: 57–59.

Petrij F, Giles RH, Dauwerse HG, Saris JJ, Hennekam RC, Masuno M *et al*. (1995). *Nature* **376**: 348–351.

Rieder MJ, Taylor SL, Tobe VO, Nickerson DA. (1998). *Nucleic Acids Res* **2**: 967–973.

Schmid M, Steinlein C, Haaf T. (2004). *Cytogenet Genome Res* **104**: 277–282.

Shigeno K, Yoshida H, Pana L, Luo JM, Fujisawa S, Naito K *et al*. (2004). *Cancer Lett* **213**: 11–20.

Shin MG, Kajigaya S, McCoy Jr JP, Levin BC, Young N. (2004). *Blood* **103**: 553–561.

Sidransky D, Mikkelsen T, Schwechheimer K, Rosenblum ML, Cavanee W, Vogelstein B. (1992). *Nature* **355**: 846–847.

Singh KK, Kulawiec M, Still I, Desouki MM, Geradts J, Matsui S. (2005). *Gene* **354**: 140–146.

Singh SK, Clarke ID, Terasaki M, Bonn VE, Hawkins C, Squire J *et al*. (2003). *Cancer Res* **63**: 5821–5828.

van Tilborg AAG, de Vries A, de Bont M, Groenfeld LE, van der Kwast TH, Zwarthoff EC. (2000). *Hum Mol Genet* **9**: 2973–2980.

Supplementary Information accompanies the paper on the Oncogene website (http://www.nature.com/onc).