Accumulation of mutations over the entire mitochondrial genome of breast cancer cells obtained by tissue microdissection

Liane Fendt, Harald Niederstätter, Gabriela Huber, Bettina Zelger, Martina Dünser, Christof Seifarth, Alexander Röck, Georg Schäfer, Helmut Klocker, Walther Parson

Accumulation of mutations over the entire mitochondrial genome of breast cancer cells obtained by tissue microdissection

Liane Fendt[1], Harald Niederstätter[1], Gabriela Huber[1], Bettina Zelger[2], Martina Dünser[4], Christof Seifarth[2], Alexander Röck[5], Georg Schäfer[3], Helmut Klocker[3], Walther Parson[1§]

[1] Institute of Legal Medicine, Innsbruck Medical University, Austria

[2] Institute of Pathology, Innsbruck Medical University, Austria

[3] Experimental Urology, Innsbruck Medical University, Austria

[4] Department of operative Medicine, Innsbruck Medical University, Austria

[5] Institute of Mathematics, University of Innsbruck, Austria

§Corresponding author

walther.parson@i-med.ac.at; Tel: 0043 512 9003 70640; Fax: 0043 512 9003 73640

Abstract

Purpose: The occurrence of heteroplasmy and mixtures is technically challenging for the analysis of mitochondrial DNA. More than that, observed mutations need to be carefully interpreted in the light of the phylogeny as mitochondrial DNA is a uniparental marker reflecting human evolution. Earlier attempts to explain the role of mtDNA in cancerous tissues led to substantial confusion in medical genetics mainly due to the presentation of low sequence data quality and misinterpretation of mutations representing a particular haplogroup background rather than being cancer specific. The focus of this study was to characterize the extent and level of mutations in breast cancer samples obtained by tissue microdissection by application of an evaluated full mtDNA genome sequencing protocol. **Methods:** We amplified and sequenced the complete mitochondrial genomes of microdissected breast cancer cells of 15 patients and compared the results to those obtained from paired non-cancerous breast tissue derived from the same patients. **Results**: We observed differences in the heteroplasmic states of substitutions between cancerous and normal cells, one of which was affecting a position that has been previously reported in lung cancer and another one that has been identified in 16 epithelial ovarian tumors, possibly indicating functional relevance. In the coding region we found full transitions in two cancerous mitochondrial genomes and 12 heteroplasmic substitutions as compared to the non-cancerous breast cells. **Conclusions**: We identified somatic mutations over the entire mtDNA of human breast cancer cells potentially impairing the mitochondrial OXPHOS system.

Abbreviations:

mtDNA - mitochondrial DNA

LCM - laser capture microdissection

ND - NADH dehydrogenase

CO – cytochrom oxidase

cytb – cytochrome b

OXPHOS – oxidative phosphorylation

Background / Introduction

The involvement of mitochondria in carcinogenesis is based on the Warburg theory of an impaired energy metabolism in cancer cells triggered by mitochondrial dysfunction [1-3]. The role of mtDNA mutations in cancer has been discussed since the reporting of several somatic mutations in various types of cancer [4]. Whereas the majority of the initial studies focused on the non-coding part of the mtDNA, namely the control region, the analysis of the coding region was targeted in detail more recently [5]. Mitochondrial DNA sequence analysis is technically challenging due to complex phenomena such as heteroplasmy and mixtures of mtDNA types that overlay when characterized by direct sequence analysis, but also due to the fact that observed mutations - with respect to control or reference sequences - need to be carefully interpreted in the light of the phylogeny, particularly in the medical context. Hence, investigations on mtDNA mutations in cancer differed profoundly in data quality and their interpretation as discussed in [6]. Still, the literature displays studies that need to be re-evaluated [7]. A recent study by He et al. [8] which is the first study comparing tumor to normal mtDNA with a second generation sequencing approach uncovers widespread heteroplasmy in normal cells as well as heteroplasmy and homoplasmy in cancer cells. Very surprisingly, the mtDNA of an individual that belongs to mitochondrial haplogroup J1c3a1 obviously lacks signature mutations on positions 15326, 2706, 4216, 3010 and 13934. By all means this is an unusual finding as these positions are yet considered to be stable in the phylogeny [9] and should be re-evaluated in that respect. Despite this fact, there is convincing evidence that a general feature of cancer cells. is the accumulation of somatic mutations scattered throughout the entire mitochondrial genome [10]. Theoretically, the high frequency of mtDNA variation in cancer may be explained by neutral random drift in clonally expanding cell populations [11]. However, the selective advantage of mtDNA changes in the development of tumors cannot be excluded either [12]. Recent findings by Beerenwinkel et al. [13] support the theory that tumors contain both neutral mutations (passengers) as well as a smaller fraction of positively-selected mutations (drivers) promoting tumor progression.

Assessment of tumor development and progression by sequence analysis gets technically challenging when heteroplasmic mixtures contain a minor contribution of about 10% of the total signal [6;7;14]

highlighting the paramount role of using appropriate quality of mtDNA raw data to prevent misinterpretation of the data. We have successfully designed a strategy for the amplification and Sanger sequencing of full mitochondrial genomes with reproducible high data quality [15]. Only a systematic investigation of a representative number of tumor cells and corresponding normal cells allows for a meaningful evaluation of the role of mtDNA mutations for the development of cancer and the discussion of functional consequences. However, for the time being, there is insufficient reliable mtDNA data available addressing the comparison between neoplastic versus normal breast cancer cells.

In this study we performed a full mitochondrial genome analysis for somatic mutations in paired cancerous and non-cancerous tissues from 15 patients with diagnosed breast cancer using our evaluated protocols. The main aim of this study was to characterize the number and proportion in heteroplasmic mixtures of somatic mutations in primary cancerous tissues in comparison to non-cancerous cells of the same individual.

Material and Methods

Clinical samples

Samples derived from patients with diagnosed invasive mamma carcinoma upon mastectomy or tumorectomy under informed and written consent. Criteria for exclusion were tumors which were smaller than 0.5 cm in diameter as well as non-invasive tumors. The study received declaration of no objection by the ethical committee of the Innsbruck Medical University.

Tissue specimens

Frozen and paraffin-embedded tissue samples were obtained directly from breast cancer patients who had undergone tumorectomy or mastectomy after fine needle biopsy diagnosis of breast cancer. The specimens were cooled in ice/water immediately and brought to the pathologist who performed a rapid section and isolated a frozen tissue slice that was embedded in Tissue-Tek OCT Compound (Sakura Finetek Germany GmbH, Staufen, Germany), snap frozen in liquid nitrogen and stored at -80°C until use. The rest of the tumorectomy was fixed and paraffin-embedded according to standard procedures.

Laser capture tissue microdissection

Frozen sections of the tumor specimens were stained with hematoxylin and eosin for pathological analysis and exact localization of the tumor regions. Parallel, unstained 8 µm frozen sections were used for laser capture microdissection (LCM). Slides were pretreated for 1 min in each of the following cooled solutions: 75% ethanol, DEPC-Water (nuclease free water) 100% ethanol (twice), xylene (twice) and then air dried. LCM was performed using a PixCell II LCM System (Arcturus Bioscience Inc., Mountain View, CA, USA) with 1,000 – 3,000 laser impulses for each sample corresponding to about 10,000 to 30,000 captured cells.

DNA was extracted from single breast cancer and corresponding normal gland cells derived from laser capture microdissection of the frozen tissue blocks using EZ1 Robotic Workstation and the GenoPrep DNA from Tissue Kit (both Qiagen, Hilden Germany) according to the manufacturer's manual. DNA was resolved in water and quantified via real-time PCR [16].

MtDNA amplification and sequencing

The entire human mitochondrial genome was amplified and sequenced as previously described [15]. In brief, full mtDNA genomes were amplified as 2 overlapping fragments of 8.5 kb each in length using aliquots of 5,000 mitochondrial genomes for PCR. Subsequent chain termination sequencing was performed using 96 primers covering the coding region and 10 primers covering the control region [17]. Laser induced fluorescence capillary electrophoretic separation of purified sequencing products was performed on a 3100 DNA Analyzer (Applied Biosystems, Foster City, CA).

Cloning and sequencing of selected samples

For 4 patients displaying point-heteroplasmy in the tumor tissue at positions 12875 (patient no. 1), 2998 (patient no. 8), 12131(patient no. 11), and 1632 (patient no. 13), the corresponding mtDNA segments from the non-cancerous normal tissue samples were amplified with the proofreading KOD hot start DNA polymerase (Novagen, Madison, WI, USA) and unlabelled primers: for 5'-TTGCTCATCAGTTGATGATACG-3' and rev 5'-AGCGGATGAGTAAGAAGATTCC-3' flanking position 12875, for 5'-CCCAACCTCCGAGCAGTACATG-3' and rev 5'-CGAACCTTTAATAGCGGCTGCACCAT-3' flanking position 2998, for 5'-AACCACGTTCTCCTGATCAAA-3' and rev 5'-GGTCGTAAGCCTCTGTTGTCAG-3' flanking position 12131 and for 5'-GTGGCAAGAAATGGGCTAC-3' and rev 5'-GCCAGGTTTCAATTTCTATCG-3' flanking position 1632. Ligation of the PCR products into the pCR 4-Blunt-TOPO vector and subsequent transformation into chemically competent One Shot TOP10 *E. coli* were performed using the Zero Blunt TOPO PCR cloning kit for sequencing (Invitrogen Life Technologies, Carlsbad, CA, USA) according to the manufacturer's protocol. Transformed cells were incubated at 37° C over night on LB agar plates containing 50μg/mL Kanamycin. Small amounts of bacterial colonies were picked using disposable pipette tips, transferred directly to 1.5 mL LB medium supplemented with 50 μg/mL Kanamycin in 96-well deep well plates, and incubated for 18 hours at 37 °C with constant shaking (200 rpm). Plasmid DNA (pDNA) minipreps from over-night suspension cultures were performed in the 96-well format with the montáge

Plasmid Miniprep$_{96}$ Kit (Millipore, Billerica, MA, USA) according to the manufacturer's recommendations. 1 μL aliquots of pDNA minipreps were subjected to cycle sequencing (30 cycles comprising 95 °C for 15 s, 50 °C for 10 s and 60 °C for 4 min after an initial denaturation step of 2 min at 95 °C) using the primers M13 for (GTAAAACGACGGCCAGTGA) and M13 rev (GGAAACAGCTATGACCATG) and BigDye terminator chemistry (v1.1, Applied Biosystems). Electrophoresis of purified sequencing products was performed using a 3100 DNA Analyzer (Applied Biosystems).

Data analysis and quality assurance

Upon analysis of the raw data the sequences were aligned and the base-calls were reviewed independently by two scientists using the Sequencher software (v.4.8, Gene Codes Corporation, Ann Arbour, MI, USA). Consensus sequences were reported relative to the revised Cambridge Reference Sequence (rCRS) [18;19] following nomenclature guidelines for mtDNA typing suggested by [20-22].

Results

Somatic mutations and differences in the control region between breast tumor cells and corresponding normal breast tissue

In the 15 investigated sample pairs of cancerous and corresponding normal breast tissues we detected six differences in a total of five patients within the mtDNA control region (Table I). Four of the observed heteroplasmic mutations at positions 152, 215, 16304 and 16390 in the control region were also observed by Irwin et al. [23] who investigated the occurrence of point-heteroplasmy in the control region of over 5,000 global population samples. Mutations at positions 152, 16304 and 16390 occur in multiple macro-haplogroup backgrounds (haplogroups L, M, N, and R) [24] whereas position 215 (8 observations of heteroplasmy in more than 5000 individuals) is not reported as evolutionary fast site [25;26] . The remaining two mutations were 16106G/$_{A}$ and 16391G/$_{A}$. 16106A variant was described for haplotypes belonging to European haplogroup T [27], and 16391A is common in clades I and A2g of haplogroup N [9]. Note that positions 152, 215 and 16391 harboured the heteroplasmic mixture in the normal tissue but were homoplasmic in the tumor.

Somatic mutations and differences in the coding region between breast tumor cells and corresponding normal breast tissue

In the mitochondrial coding region of the 15 analyzed patients a total of 14 mutations was detected between cancerous and corresponding non-cancerous reference tissue, five (35.7%) of which affected ND genes, two (14.3%) were located in the CO genes, four (28.6%) were found in mtRNA genes, and one hit (7.1%) was observed in the cytb gene (Table II). Two mutations in the cancerous cells (12875C, 12131C) were present in homoplasmic state. Eight positions harboured wild type and variant within the cancer cells but only wild type within the normal tissue (7379G➔A/$_{G}$, 5703G➔G/$_{A}$, 15341T➔T/$_{C}$, 2998T➔C/$_{T}$, 2145G➔A/$_{G}$, 12803G➔A/$_{G}$, 1632T➔C/$_{T}$, 1132T➔C/$_{T}$). Four heteroplasmic substitutions observed within the normal mtDNA were not detected in the cancer cells (9966A/$_{G}$➔A, 5102A/$_{G}$➔A, 5390A/$_{G}$➔A, 1578A/$_{G}$➔A). It has to be noted that we also found (low level) heteroplasmy within four non-cancerous DNA samples (Figure 1, 4a-d).

For non-synonymous base substitutions (2145A/$_G$, 2998C/$_T$, 12803A/$_G$, 12875C, 12131C, 15341T/$_C$) as well as for the mutations found in the rRNA and tRNA genes an evolutionary conservation analysis was performed. Positions 2998, 15341, 12803 and 12875 were conserved in an inter-vertebrate species comparison while position 12131 was not. (Figure 2).

The sensitivity of direct cycle sequencing to reliably display low level mixtures depends on numerous factors including the base composition of the adjacent nucleotide sequence and the uneven ddNTP incorporation rate. To clarify if the apparently fixed somatic mutations found in the tumor cells were actually present at a proportion below the detection limit of sequencing analysis in the normal tissue, clones of PCR products obtained from normal distant tissue samples were sequenced (Table III). Therefore we selected two patients where the corresponding cancerous samples harboured full transitions 12875T (patient 1) and 12131C (patient 11), one patient where the corresponding cancer sample displayed point heteroplasmy (1632C/$_T$, sample no. 13) and patient 8 whose corresponding cancerous sample harboured a partial transition at the evolutionarily well conserved position 2998T. None of the analyzed clones displayed any variants that would indicate low level contributions potentially undetected by sequence analysis. The 95% confidence intervals further strengthened the assumption that there was no sequence variation in the corresponding reference tissue samples at positions 12875 (sample no. 1), 12131 (sample no. 11), 1632 (sample no.13) and 2998 (sample no. 8).

Discussion and Conclusions

In this study we sequenced the entire mitochondrial genomes of 15 primary breast cancer patients and observed in total 20 differences between the mtDNA of breast cancer cells and the corresponding distant normal tissue. Our results are based on an evaluated sequencing strategy for full mtDNA genome sequencing analysis [15], yielding high quality sequencing electropherograms that are characterized by very low background noise levels, multiple sequence coverage, electronic data handling and hence unequivocal base calls. In addition, the data was checked for plausibility by constructing a phylogenetic tree of the detected lineages (Figure 3). All of the observed full mtDNA sequences could be assigned to known European and Asian mtDNA lineages. Haplotypes belonging to superhaplogroup H were found most commonly (53.3%) which is in agreement with the expected prevalence of this haplogroup in Europe [28]. In addition, we observed two haplotypes from clusters HV and J (each 13.3%) and one representative of clusters K and R0a (6.7%). One sequence (6.7%) belonged to an Asian lineage (M35b).

When comparing our observations to previous studies with a similar setup, we noticed a variety of different findings, one being the number of observed mutations in the various tumor tissues. While Kassauei et al. [29] found 71 mutations in 15 complete mtDNA sequences of pancreatic ductal adenocarcinoma cells, Zhu et al. [30] found 45 somatic mutations in 15 complete mtDNA genomes of breast cancer tissues and Rosson et al. [31] observed various sets of somatic mutations in the mtDNA control region when comparing breast cancer cells and corresponding normal tissues. In a study by Tan et al. [32] fourteen of the 19 analyzed tumors displayed at least one somatic mtDNA mutation. Parella et al. [33] found 12 somatic mutations in 18 primary breast tumors as compared to the corresponding mtDNAs. In contrast, Wang et al. [10] found only two heteroplasmic point substitutions within 10 sample pairs when analyzing the coding region.

The reasons for the different extent of mutations within tumor mtDNA remain unclear so far but it is conceivable that multiple parameters influence the induction of mtDNA mutations. Wang [10] compared mtDNA mutations in early stage breast cancer and suggested that the observed mutations resulted from somatic mutational events in primary cancerous or paracancerous tissues. Sigh et al. [34] investigated the influence of mutations within nuclear genes on the mutation rate of mtDNA.

They concluded that mutations in the polymerase domain of the POLG gene decrease mitochondrial activity and increase oxidative stress resulting in an increased leakage of ROS which in turn induce mutations and promote tumorigenesis.

The reporting of significantly different numbers of mutations in tumor cells might be also due to the misinterpretation of the resulting electropherograms based on ambiguous data [6].

Also, the nature of observed mutations in our study differed from other reports. While we found – and confirmed by cloning - 2 full transitions and 12 heteroplasmic mutations in the coding region (Table II), the majority of other studies observed largely apparent homoplasmic somatic mutations, for discussion see [5]. Tzen et al. [35] as well as Kassauei et al. [29] observed that several mutations are present in heteroplasmy in healthy counterparts and became homoplasmic in cancer tissues.

This is an interesting fact as the level of heteroplasmy is supposed to influence the phenotype in a way that if the pathogenic threshold is surpassed cellular dysfunction caused by defective mitochondria becomes apparent. This also explains the time-related and tissue-specific variability of clinical features displayed in mtDNA-related disorders. Hence, it is remarkable that Park et al. [36] found that a cell line carrying a heteroplasmic ND5 mtDNA mutation showed significantly enhanced tumor growth, while cells with the homoplasmic mutant variant inhibited tumor formation. In contrast, in a study of Petros et al. [37], the introduction of the homoplasmic variant 8993G in a prostate cancer cell line led to a 7 times larger tumor growth in the resulting cybrid cell line as compared to the wild type-cybrids.

It is further worth mentioning that one cancer sample in our study harboured a mutation at the same position as observed by Aikhionbare et al. [38] in 16 out of 102 epithelial ovarian tumors. In our case position 1632 ($tRNA^{Val}$) was affected by a heteroplasmic transition to C, while Aikhionbare et al. [38] reported transversions to G in all of the observed tumor samples. A phylogenetic analysis of different species suggested that site 1632 is conserved (Figure 2) but might mean that the transversion has a more deleterious effect than the transition. Another interesting finding is the co-occurrence of mutational events at position 2998 (16s rRNA) in one cancer sample of this study and in a lung cancer sample published by Lorenc et al. [39]. While we found a heteroplasmic transition to C at that

position, Lorenc et al. reported an insertion. Evolutionary comparisons suggested this position being a highly conserved site and therefore, an alteration of this position might have functional consequences.

The main yet unexplained question is: when do these mutations actually arise? Do the detected mutations induce cancer development – thus being present at very low levels in pre-cancer states, or does cancer induce mutations? In a study of Wang et al. [10] one of the two detected cancer variants (8601G) also occurred in the paracancerous tissue, and the homoplasmic variant in the tumor arose either through random drift or selection, while the other detected cancer variant (2275C) arose only in the tumor, which was confirmed by subsequent cloning. In our study, tumor specificity of two apparently homoplasmic mutations in the ND4 and ND5 genes at positions 12875 and 12131 was confirmed by cloning and subsequent sequencing of the corresponding mtDNA segments from the normal tissue samples, for which only the wild type alleles were found. These findings support the model of induction of mutations within the tumor. Contrary, a study of Tzeng et al. [35] showed that mutations were found in 13 out of 53 analyzed ND4 genes of tumor cells but could also be identified in 11 samples of the corresponding normal tissues by DHPLC, indicating that most mtDNA mutations identified in tumors pre-existed already as minor components in heteroplasmic mixtures. These results suggested that mtDNA mutation occurred before tumorigenesis and became apparent in cancer cells.

MtDNA mutations in tumors obviously do not emerge randomly over the entire genome but rather affect phylogenetic informative positions as well as evolutionary fast sites, at least in the control region [40]. This is also evident in our data, as transitions affecting positions 16390 and 16391 are found across the entire phylogeny and positions 152 as well as 16304 are known mutational hotspots [9]. Four of the observed mutations on positions 152, 215, 16304, 16390 were also determined as frequent hits at heteroplasmic sites in the control region of more than 5,000 global population samples [23]. Only position 215 was not reported as evolutionary fast site [25;26] and therefore selection against it cannot be excluded, in particular as position 215 is situated close to the origin of heavy strand replication and a mutation there might influence mtDNA replication [5;23]. On the contrary, 215T can be found in sublineages of all 4 major macro-haplogroups M, N, R and L [9]. 16106A is a

rare variant which has been observed in seven of 115 haplotypes of an alpine population [24;27]. It has been shown in a cell culture system that the rate of ROS induced lesions was 5 fold higher in the control region as compared to the coding region [41]. In the coding part of the mtDNA the literature displays the most affected regions to code for Complex I in the electron transport chain [5;42;43] which is consistent with our results (35.7 % of detected mutations were found in genes coding for Complex I). Functional consequences of these mutations are conceivable as they cause amino acid changes from serine to proline (12131C) and from isoleucine to threonine (12875C). It is possible that a small fraction of mutations act as driver mutations having functional consequences. This assumption goes with the findings of Zhidkov et al. [44] providing evidence for similar selective constraints for the patterns in mtDNA mutations in tumors and human evolution reflecting a response to positive selection pressures in cancer, rather than a random adaptive process.

On the contrary, mutations can become homoplasmic during the process of clonal expansion of tumor cells, representing a model of "rapid evolution". This theory is in agreement with the findings of Coller et al. [11] who explained the existence of homoplasmy in tumors solely as a consequence of tumor kinetics without any needs for positive selection, although they do not exclude the possibility of other mechanisms.

In conclusion, we identified somatic mutations over the entire mtDNA of human breast cancer cells potentially impairing the mitochondrial OXPHOS system. Two mutations affected positions that have already been reported by others. Our results suggest that mtDNA mutations may play an important role in the progression of breast cancer, but the mechanism by which mtDNA mutations contribute to cancer development remain unclear. Clearly, there is still more data needed from full mitochondrial genomes of cancer and corresponding non-cancerous tissues to get more insight in the processes involved.

Reference List

1 Warburg O (1956) On the origin of cancer cells. Science **123** 309-314

2 Wallace DC (2005) A mitochondrial paradigm of metabolic and degenerative diseases aging and cancer: a dawn for evolutionary medicine. Annu Rev Genet **39** 359-407

3 Brandon M, Baldi P, Wallace DC (2006) Mitochondrial mutations in cancer. Oncogene **25** 4647-4662

4 Bianchi NO, Bianchi MS, Richard SM (2001) Mitochondrial genome instability in human cancers. Mutat Res **488** 9-23

5 Santos C, Martínez M, Lima M, Hao YJ, Simoes N, Montiel R (2008) Mitochondrial DNA mutations in cancer: a review. Curr Top Med Chem **8** 1351-1366

6 Salas A, Yao YG, Macaulay V, Vega A, Carracedo A, Bandelt HJ (2005) A critical reassessment of the role of mitochondria in tumorigenesis. PLoS Med **2** e296

7 Chen D and Zhan H (2009) Study on the mutations in the D-loop region of mitochondrial DNA in cervical carcinoma. J Cancer Res Clin Oncol 135 291-295

8 HeY, Wu J, Dressman DC, Iacobuzio-Donahue C, Markowitz SD, Velculescu VE, Diaz Jr LA, Kinzler KW, Vogelstein B, Papadopoulos N (2010) Heteroplasmic mitochondrial DNA mutations in normal and tumour cells. Nature doi:10.1038/nature08802

9 van Oven M, Kayser M (2008) Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. Hum Mutat **1039** e386-e394

10 Wang CY, Wang HW, Yao YG, Kong QP, Zhang YP (2007) Somatic mutations of mitochondrial genome in early stage breast cancer. Int J Cancer 121 1253-1256

11 Coller HA, Khrapko K, Bodyak ND, Nekhaeva E, Herrero-Jimenez P, Thilly WG (2001) High frequency of homoplasmic mitochondrial DNA mutations in human tumors can be explained without selection. Nat Genet 28 147-150

12 Maximo V, Lima J, Soares P, Sobrinho-Simoes M (2009) Mitochondria and cancer. Virchows Arch **454** 481-495

13 Beerenwinkel N, Antal T, Dingli D, Traulsen A, Kinzler KW, Velculescu VE, Vogelstein B, Nowak MA (2007) Genetic progression and the waiting time to cancer. PLoS Comput Biol **3** e225

14 Bandelt HJ, Salas A, Bravi CM (2006) What is a 'novel' mtDNA mutation--and does 'novelty' really matter? J Hum Genet **51** 1073-1082

15 Fendt L, Zimmermann B, Daniaux M, Parson W (2009) Sequencing strategy for the whole mitochondrial genome resulting in high quality sequences. BMC Genomics **10** 139

16 Niederstätter H, Köchl S, Grubwieser P, Pavlic M, Steinlechner M, Parson W (2007) A modular real-time PCR concept for determining the quantity and quality of human nuclear and mitochondrial DNA. Forensic Sci Int: Genet **1** 29-34

17 Parson W, Bandelt HJ (2007) Extended guidelines of mtDNA typing of population data in forensic science. Forensic Sci Int: Genetics **1** 13-19

18 Anderson S, Bankier AT, Barrell BG, de Bruijn MH, Coulson AR, Drouin J, Eperon IC, Nierlich DP, Roe BA, Sanger F, Schreier PH, Smith AJ, Staden R, Young IG (1981) Sequence and organization of the human mitochondrial genome. Nature **290** 457-465

19 Andrews RM, Kubacka I, Chinnery PF, Lightowlers RN, Turnbull DM, Howell N (1999) Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. Nat Genet **23** 147

20 Carracedo A, Bär W, Lincoln P, Mayr W, Morling N, Olaisen B, Schneider P, Budowle B, Brinkmann B, Gill P, Holland M, Tully G, Wilson M (2000) DNA Commission of the International Society for Forensic Genetics: guidelines for mitochondrial DNA typing. Forensic Sci Int **110** 79-85

21 Tully G, Bär W, Brinkmann B, Carracedo A, Gill P, Morling N, Parson W, Schneider P (2001) Considerations by the European DNA profiling (EDNAP) group on the working practices nomenclature and interpretation of mitochondrial DNA profiles. Forensic Sci Int **124** 83-91

22 Bandelt HJ, Parson W (2007) Consistent treatment of length variants in the human mtDNA control region: a reappraisal. Int J Legal Med **122** 11-21

23 Irwin JA, Saunier JL, Niederstätter H, Strouss KM, Sturk KA, Diegoli TM, Brandstätter A, Parson W, Parsons TJ (2009) Investigation of Heteroplasmy in the Human Mitochondrial DNA Control Region: A Synthesis of Observations from More Than 5000 Global Population Samples. J Mol Evol **68** 516-527

24 Malyarchuk BA, Derenko MV (1999) Molecular instability of the mitochondrial haplogroup T sequences at nucleotide positions 16292 and 16296. Ann Hum Genet **63** 489-497

25 Bandelt HJ, Kong QP, Richards M, Macaulay V (2006) Estimation of mutation rates and coalescence times: some caveats In Bandelt HJ, Macaulay V, and Richards M (eds) Human mitochondrial DNA and the evolution of Homo sapiens. Springer-Verlag Berlin Germany pp 47-90

26 Bandelt HJ, Quintana-Murci L, Salas A, Macaulay V (2002) The fingerprint of phantom mutations in mitochondrial DNA data. Am J Hum Genet **71** 1150-1160

27 Vernesi C, Fuselli S, Castrí L, Bertorelle G, Barbujani G (2002) Mitochondrial diversity in linguistic isolates of the Alps: a reappraisal. Hum Biol **74** 725-730

28 Achilli A, Rengo C, Magri C, Battaglia V, Olivieri A, Scozzari R, Cruciani F, Zeviani M, Briem E, Carelli V, Moral P, Dugoujon JM, Roostalu U, Loogväli EL, Kivisild T, Bandelt HJ, Richards M, Villems R, Santachiara-Benerecetti AS, Semino O, Torroni A (2004) The molecular dissection of mtDNA haplogroup H confirms that the Franco-Cantabrian glacial refuge was a major source for the European gene pool. Am J Hum Genet **75** 910-918

29 Kassauei K, Habbe N, Mullendore ME, Karikari CA, Maitra A, Feldmann G (2006) Mitochondrial DNA mutations in pancreatic cancer. Int J Gastrointest Cancer **37** 57-64

30 Zhu W, Qin W, Bradley P, Wessel A, Puckett CL, Sauter ER (2005) Mitochondrial DNA mutations in breast cancer tissue and in matched nipple aspirate fluid. Carcinogenesis **26** 145-152

31 Rosson D, Keshgegian AA (2004) Frequent mutations in the mitochondrial control region DNA in breast tissue. Cancer Lett **215** 89-94

32 Tan DJ, Bai RK, Wong LJ (2002) Comprehensive scanning of somatic mitochondrial DNA mutations in breast cancer. Cancer Res **62** 972-976

33 Parrella P, Xiao Y, Fliss M, Sanchez-Cespedes M, Mazzarelli P, Rinaldi M, Nicol T, Gabrielson E, Cuomo C, Cohen D, Pandit S, Spencer M, Rabitti C, Fazio VM, Sidransky D (2001) Detection of mitochondrial DNA mutations in primary breast cancer and fine-needle aspirates. Cancer Res **61** 7623-7626

34 Singh KK, Ayyasamy V, Owens KM, Koul MS, Vujcic M (2009) Mutations in mitochondrial DNA polymerase-gamma promote breast tumorigenesis. J Hum Genet **54** 516-524

35 Tzen CY, Mau BL, Wu TY (2007) ND4 mutation in transitional cell carcinoma: Does mitochondrial mutation occur before tumorigenesis? Mitochondrion **7** 273-278

36 Park JS, Sharma LK, Li H, Xiang R, Holstein D, Wu J, Lechleiter J, Naylor SL, Deng JJ, Lu J, Bai Y (2009) A heteroplasmic not homoplasmic mitochondrial DNA mutation promotes tumorigenesis via alteration in reactive oxygen species generation and apoptosis. Hum Mol Genet **18** 1578-1589

37 Petros JA, Baumann AK, Ruiz-Pesini E, Amin MB, Sun CQ, Hall J, Lim S, Issa MM, Flanders WD, Hosseini SH, Marshall FF, Wallace DC (2005) mtDNA mutations increase tumorigenicity in prostate cancer. Proc Natl Acad Sci U S A **102** 719-724

38 Aikhionbare FO, Mehrabi S, Kumaresan K, Zavareh M, Olatinwo M, Odunsi K, Partridge E (2007) Mitochondrial DNA sequence variants in epithelial ovarian tumor subtypes and stages. J Carcinog **6** 1

39 Lorenc A, Bryk J, Golik P, Kupryjanczyk J, Ostrowski J, Pronicki M, Semczuk A, Szolkowska M, Bartnik E (2003) Homoplasmic MELAS A3243G mtDNA mutation in a colon cancer sample. Mitochondrion **3** 119-124

40 Vega A, Salas A, Gamborino E, Sobrido MJ, Macaulay V, Carracedo A (2004) mtDNA mutations in tumors of the central nervous system reflect the neutral evolution of mtDNA in populations. Oncogene **23** 1314-1320

41 Rothfuss O, Gasser T, Patenge N (2009) Analysis of differential DNA damage in the mitochondrial genome employing a semi-long run real-time PCR approach. Nucleic Acids Res doi:10 1093/nar/gkp1082

42 Gasparre G, Porcelli AM, Bonora E, Pennisi LF, Toller M, Iommarini L, Ghelli A, Moretti M, Betts CM, Martinelli GN, Ceroni AR, Curcio F, Carelli V, Rugolo M, Tallini G, Romeo G (2007) Disruptive mitochondrial DNA mutations in complex I subunits are markers of oncocytic phenotype in thyroid tumors. Proc Natl Acad Sci U S A **104** 9001-9006

43 Gasparre G, Hervouet E, de LE, Demont J, Pennisi LF, Colombel M, Mege-Lechevallier F, Scoazec JY, Bonora E, Smeets R, Smeitink J, Lazar V, Lespinasse J, Giraud S, Godinot C, Romeo G, Simonnet H (2008) Clonal expansion of mutated mitochondrial DNA is associated with tumor formation and complex I deficiency in the benign renal oncocytoma. Hum Mol Genet **17** 986-995

44 Zhidkov I, Livneh EA, Rubin E, Mishmar D (2009) mtDNA mutation pattern in tumors and human evolution are shaped by similar selective constraints. Genome Res **19** 576-580

Acknowledgements

AR and WP were supported by the FWF Austrian Science Fund Translational Research Programme (L397). We are grateful to o.Univ.-Prof. Dr. Raimund Margreiter (Department of Operative Medicine, Innsbruck Medical University) for supporting the study.

Abbreviations:

mtDNA - mitochondrial DNA

LCM - laser capture microdissection

ND - NADH dehydrogenase

CO – cytochrom oxidase

cytb – cytochrome b

OXPHOS – oxidative phosphorylation

Table I. List of observed CR sequence differences between cancerous and corresponding normal breast tissues of 15 patients.

patient No	haplogroup	normal tissue[a]	cancer tissue[a]	hits in the phylogeny[b]
5	H11a	16391G/A	16391G	A2g, I
9	H3	16390G	16390A/G	B4f, C5c, D5a1, D5c1, E, F1a3, L0a2b, L0d2a, L1c4, L2, M27b, M48, N1b, N9b1b, R7a1b2, R8b, R9b1, U3a
12	HV	215A/G	215A	A2p, C1c1, J2a1, L0k2, M11
13	M35b	16106G	16106G/A	T
		16304C	16304T/C	H5, F3, L3d3, L3h2, M17c, M21d, M22, M25, M27c, N9a2a, R0a2c, R5, R9, R31a, T2b, U5b3
14	J1c	152C/T	152T	common

a Minor contributions in point-heteroplasmic mixtures are indicated by reduced IUB base code letter size. Length variants and substitutions within the poly C stretches spanning positions 16183-16193 and 309-315 were disregarded.

b A polymorphic site survey performed using databases EMPOP (www.empop.org), mtDB (http://www.genpad.uu.se/mtDB), mtDNA tree Build 08 [9] (http://www.phylotree.org) mtDNA (http://ianlogan.co.uk/mtdna.htm) and MITOMAP (http://www.mitomap.org).

Table II. List of observed coding region sequence differences between cancerous and corresponding normal breast tissues of 15 patients

patient No	haplogroup	normal tissue[a]	cancer tissue[a]	gene	syn/non syn	hits in the literature[b]
1	J1c2c	12875T	12875C	ND5	Ile➔Thr	-
2	H	7379G	7379A/G	CO I	syn	M34a (Chandrasekar 09)
3	R0a1a	5703G	5703G/A	tRNA Asn		-
6	H5	9966A/G	9966A	CO III	Val➔Ile	G1c (Derenko 07), I1 (Derenko 07, Finnilä 01), L1c3a (Behar 08), M1a1c (Olivieri 06), M30c1a (Behar 08a), M31a2 (Barik 08)
7	H1	15341T	15341T/C	cyt b	Phe➔Leu	A5a (Tanaka 04), T1a1 (Coble 04)
8	K1a1b1	2998T	2998C/T	16s rRNA		lung cancer (Lorenc 03)
		2145G	2145A/G	16s rRNA		L3e3b (Kivisild 06), H1b (Herrnstadt 02)
11	H2a	12131T	12131C	ND 4	Ser➔Pro	-
12	HV	12803G	12803A/G	ND 5	Ser➔Asp	-
13	M35b2	1632T	1632C/T	tRNA Val		-
		5102A/G	5102A	ND 2	syn	R5 (Palanichamy 04), L3e2a (Behar 08)
		5390A/G	5390A	ND 2	syn	U2e (Palanichamy 04, Achilli 05)
14	J1c	1132T	1132C/T	12s rRNA		-
		1578A/G	1578A	12sr RNA		

a Minor contributions in point-heteroplasmic mixtures are indicated by reduced IUB base code letter size. Length variants and substitutions within the poly C stretches spanning positions 16183-16193 and 309-315 were disregarded.

b A polymorphic site survey performed using databases EMPOP (www.empop.org), mtDB (http://www.genpad.uu.se/mtDB), mtDNA tree Build 08 [9] (http://www.phylotree.org) mtDNA (http://ianlogan.co.uk/mtdna.htm) and MITOMAP (http://www.mitomap.org). References are available as supplementary file.

Table III. Frequency of variants found by cloning of PCR products obtained from normal distant breast tissues of patients no 1, 8, 11 and 13

Patient No	no. of clones/variants	relative frequency of variant (95% CI)
1	53/0	0 (0.0000-0.0582)
8	65/0	0 (0.000-0.048)
11	48/0	0 (0.000-0.0639)
13	80/0	0 (0.000-0.0393)

Fig. 1. Sequencing electropherograms showing different mutational conditions in the mitochondrial genomes of breast cancer cells as compared to the corresponding normal cells. Electropherograms of reverse strands are not shown. 1a/b: Full transitions on positions 12875 and 12131 from T in the normal to C in the corresponding cancerous mitochondrial genome from patients no 1 and 11. 2a-f: Heteroplasmic transitions in the tumor mt genome with the haplogroup - specific variant represented as the dominant type (position 7379: patient no. 2; 2145: patient no. 8; 2998: patient no. 8; 12803: patient no. 12; 1632: patient no. 13; 1132: patient no. 14). 3a/b: Heteroplasmic transitions in the tumor tissue where the haplogroup-specific variant constituted the minor contribution (position 5703: patient no. 3; 15341: patient no. 7). 4a-d: Heteroplasmic states in normal tissue that is not detected in mitochondrial tumor DNA (position 9966: patient no. 6; position 5390: patient no 13; 5102: patient no. 13; 1578: patient no. 14).

Fig. 2. Evolutionary conservation analysis. Up to 7 vertebrate sequences comprising positions 1632, 2998, 12803, 15341, 12875, 2145 as well as 12131 were aligned. GenBank accession numbers are given in brackets.

Fig. 3. Depiction of the germ line and somatic mutations of the observed haplotypes in a condensed phylogenetic tree. Mutations are denoted as differences to the rCRS according to van Oven version 08 [9]. Transversions are specified in capital letters, del indicates deletion, and + indicates insertion. Underlined mutations are recurrent within the phylogenetic tree. Sample numbers are marked in red bullets. Heteroplasmic positions are notated according to the IUPAC code. Sequences were deposited in GenBank (http://www.ncbi.nlm.nih.gov/Genbank/) with accession numbers GU592019-GU592048.