Published in final edited form as:

Mutat Res. 2017 October ; 803-805: 22–25. doi:10.1016/j.mrfmmm.2017.07.003.

Single nucleotide variations in cultured cancer cells: effect of mismatch repair

Igor G. Panyutin*, Irina V. Panyutin, Ian Powell-Castilla, Laura Felix, and Ronald D. Neumann

Nuclear Medicine Division, Department of Radiology and Imaging Sciences, Clinical Center, National Institutes of Health, 9000 Rockville Pike, Bethesda, Maryland, 20892, USA

Abstract

We assessed single nucleotide variations (SNVs) between individual cells in two cancer cell lines; DU145, from brain metastasis of prostate tumor with deficient mismatch repair; and HT1080, a fibrosarcoma cell line. Clones of individual cells were isolated, and sequenced using Ion Ampliseq comprehensive cancer panel that covered the exomes of 409 oncogenes and tumor suppressor genes. Five clones of DU145 and four clones of HT1080 cells were analyzed. We found from 7 to 12 unique SNVs between DU145 clones, while HT1080 clones showed no more than one unique SNV. We then sub-cloned individual cells from some of these isolated clones of DU145 and HT1080 cells. The sub-clones were expanded from a single cell to approximately one million cells after about 20 cell divisions. The sub-clones of DU145 cells had from one to four new unique SNVs within the sequenced regions. No unique SNVs were found between sub-clones of HT1080 cells. Our data demonstrate that the extent of genetic variation at the single nucleotide level in cultured cancer cells is significantly affected by the status of the DNA mismatch repair system.

Graphical abstract

Keywords

Cancer Cells; Single Nucleotide Variations; Mismatch Repair

*Corresponding author: Igor G. Panyutin, ipanyuting@mail.nih.gov, 10 Center Dr. Room 1C492, Bethesda, MD 20854, USA.

Conflicts of Interest: The authors declare no conflict of interest.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

1. Introduction

Cultured cancer cell lines provide valuable information about molecular mechanisms of cancer that is often hard to obtain by *in vivo* studies. Studies of cultured cancer cells helped to establish chromosomal abnormalities, high mutation rates, and genetic instability in the cancer cell populations (reviewed in [1]). With advent of the next generation sequencing (NGS), it became possible to extend studies on genetic variability in cancer cells to the level of single nucleotide variations (SNVs) [2–5]. Thus, it has been established that *in vivo* cancer tumors consist of multiple genetically distinct cell populations; however, such information for the cultured cancer cells is incomplete.

In this pilot study, we assessed single nucleotide variations in two cancer cell lines, DU145 and HT1080. The prostate cancer-derived DU145 cell line with deficient mismatch repair (MMR) exhibited the highest mutation rate among all of the MMR-deficient cell lines [6]. In contrast, fibrosarcoma HT1080 is a near diploid cell line with wild type p53, and a relatively stable genome among the cancer cell lines [7].

DU145 cells showed heterogeneity in expression of cancer [8] and epigenetic [9] markers. They also exhibited chromosomal instability and karyotype heterogeneity, which were demonstrated by spectral karyotype analysis [10]. It was found that defects in MMR resulted in a significant increase in the rate of spontaneous mutations [6]. Recently, it was shown that MMR is responsible for variation in the somatic mutation rate across the human genome [11].

Herein, we extended genetic variability studies in cancer cell lines to SNVs analysis using NGS. To assess SNV between individual cells in the original cultures we isolated five clones of DU145 and four clones of HT1080 cells and analyzed them with targeted NGS. The targeted comprehensive cancer panel (CCP) covered exomes of 409 oncogenes and tumor suppressor genes. Even within this limited portion of the genome, we found a significant number of SNVs between individual clones in DU145 cells. The sub-clones derived from these clones also have detectable number of SNVs. In contrast, HT1080 cells showed an insignificant number of SNVs within the target regions.

2. Materials and Methods

DU-145 and HT-1080 cell lines were obtained from ATCC and were grown to 50% confluency according to ATCC recommendations. Cells were collected by Trypsin/EDTA and seeded in 10 cm cell culture dishes at approximately 50 cells per dish. After 14 days, visible distinct colonies generated from single cells were collected using sterile cloning cylinders (Bel-Art–SP Scienceware, South Wayne, NJ) and the manufacturer's protocol. Collected colonies were transferred to separate plates for further growth. When the total number of cells in cultures reached approximately 1×10^6 , they were collected and DNA was purified for sequencing.

Genomic DNA was isolated using a PureLink® Genomic DNA Mini Kit (Thermo Fisher Scientific, Waltham, MA, USA) and quantified by a NanoVue Plus spectrophotometer (GE Healthcare Life Sciences, Piscataway, NJ, USA). A total of 50 ng of DNA of each sample

was subjected to PCR amplification of target regions using 16000 primer pairs in four pools according to Ion AmpliSeq™ Comprehensive Cancer Panel that covers exons of 409 genes (Thermo Fisher Scientific). Barcode adaptors from the Ion Xpress™ Barcode Adapters Kit (Life Technologies, Grand Island, NY, USA) were ligated to the amplicons to allow for sample multiplexing. Sequencing libraries were prepared using the Ion AmpliSeq™ Library Kit 2.0 (Thermo Fisher Scientific) as instructed by the manufacturer. All prepared libraries were quantified using RT PCR based Ion Library TaqMan Quantitation kit (Thermo Fisher Scientific).

Each uniquely barcoded library was diluted in nuclease-free water to a stock concentration of 100 pM. The samples were prepared for sequencing using Ion Chef system and Ion PGM Hi-Q Chef Kit, and were loaded onto Ion 318™ chips (Thermo Fisher Scientific). Semiconductor-based NGS was performed on an Ion PGM™ system (Thermo Fisher Scientific) using Ion PGM™ Hi-Q™ Sequencing Kit. The successful sequencing of a sample set was determined based on the manufacturer's recommendations of expected throughput and total number of reads with a quality score of AQ20 (one misaligned base per 100 bases) for each chip.

Base calling and alignment of sample sequence to the reference genome (hg19) were performed by the Torrent Suite™ software version 5.0.2 (Thermo Fisher Scientific). All sequencing reads were automatically barcode-sorted with low-quality reads removed. The resulting BAM files were uploaded to Ion Reporter software on Thermo Fisher Cloud server for variant analysis. Variants were annotated using AmpliSeq CCP single sample-PGM 5.0 Ion Reporter application. Variant differences between clones were identified by using a custom workflow based on Ion Reported AmpliSeq CCP tumor-normal application. The variants were further filtered using the following filters: homopolymer length ≤ 7 , variant type = SNV, allele read-count ≥ 20 , alternate allele count ≥ 20 . SNV frequency was measured as a ratio of the reads with the alternative allele to the total number of reads covering the genomic position of the SNV.

In parallel, NextGENe software package version 2.3.4 (SoftGenetics, State College, PA, USA) was also used for analysis of SNV differences between clones. Clones were compared pairwise as tumor-normal pairs using somatic mutation comparison option with the following parameters: maximum contamination = 0.5%; somatic allele count = 20; relative directional balance (T/N) = 3.0; somatic allele frequency ratio (T/N) = 1.5.

Additional analysis included visual inspection of the read alignment and the presence of nucleotide variants on the Integrative Genomics Viewer (IGV, Broad Institute, Boston, MA) to confirm the variant calls by checking for possible sequencing errors. Gene Ontology analysis was done using online PANTHER Classification System (release 11.1; panterdb.org).

3. Results

The original cultures of DU145 and HT1080 cells were grown to 50% confluency. The cells were collected by Trypsin/EDTA and seeded into 10 cm petri dishes at approximately 50

cells per dish. Colonies of individual clones were isolated as described in the methods section. We selected five DU145 clones (D-1 - D-5) and four HT1080 clones (H-1–H-4) for targeted sequencing. When compared with the reference human genome (hg19), DU145 clones contained 1120 common SNVs within the targeted region with frequencies greater than 10%, including 126 SNVs listed in the Catalogue of Somatic Mutations in Cancer (COSMIC) database. Interestingly, our targeted sequencing detected mutations in both *MLH1* and *PMS2* genes that were identified in ref. [6] as the reason for the dysfunctional MMR in this cell line. The corresponding numbers for HT1080 cells were slightly lower; their clones contained 945 common SNVs with frequencies more than 10%; of that, 93 SNVs were listed in the COSMIC database.

To identify SNVs that were not common in all the clones, their sequences were compared pairwise using tumor-normal algorithms of both Ion Reporter and NextGene software packages to identify SNVs. Based on these analyses for DU145 cells, a table was compiled that showed SNVs that were not present to all the clones but either unique for a particular clone or shared by several but not all the clones. (Supplementary Table S1).

The results are also presented as Venn diagrams in Figure 1. Five clones of DU-145 cells fall into two groups. One group consists of clones D-1, D-2, and D-3 that have 11 shared SNVs and 7, 9, and 7 unique SNVs, correspondingly. Another group consists of clones D-4 and D-5 with 23 shared SNVs, and 9 and 12 unique SNVs, respectively. There was only one shared SNV between these two groups that was common in clones D-2, D-3, D-4 and D-5. The clones of HT-1080 cell line had considerably less SNVs between them than the clones of DU-145 cell line (Figure 2). HT-1080 clones had at most one SNV; only clone H-4 had a unique SNV in the *IGF1R* (p.Gly1352Arg) gene; clones H-1 and H-2 shared an SNV in the *CDH20* (p.Pro328His; rs1943330) tumor suppressor gene that is absent from clones H-3 and H-4. Clones H-3 and H-4 had a common SNV in the *USP9X* (deletion with 100% frequency) gene that was not present in clones H-1 and H-2. Taken together, these data imply that the absence of functional mismatch repair system strongly affects SNV between individual cells in cancer cell lines.

With a few exceptions, almost all SNVs that were unique for a single clone have a frequency less than 2%, or were undetectable in the original DU-145 culture (Supplementary Table S1). In contrast, all SNVs that were shared in several clones were detected with frequencies greater than 10% in the original DU-145 cells, and therefore were “pre-existing”. This was confirmed with pairwise analysis of five DU-145 clones against DNA from the original DU-145 culture using the same comparison parameters.

The original cell cultures of DU-145 and HT-1080 cells that came from ATCC underwent multiple passages and most likely were not clonally purified. They also could have contained different cells from the original tumors. These factors could contribute to the level of SNV in the cultures. Therefore, we decided to test for the presence of *de novo* SNVs in our clonally purified daughter cell lines originated from single cells. The cells of the clones were subjected to a second round of clonal expansion, and several clones of individual cells were selected, which we called sub-clones. The SNVs between these sub-clones were analyzed by pairwise tumor-normal comparison algorithm as it was done for original clones, but the

“normal” now was the sequence of the original clone. The results for six sub-clones of the clone D1 and three sub-clones of the clone D2 of DU-145 cells are summarized in Tables 1 and 2 that show the gene in which SNVs were located and the corresponding frequencies of the SNVs. The sub-clones have from one to four SNVs that occurred during the expansion of the original clones from a single cell to approximately one million cells, i.e. in the course of ca. 20 cell divisions. SNVs in two genes, *PIK3R2* and *ARNT* are found in two sub-clones, D-1-5 and D-1-7, and D-1-3 and D-1-6, respectively. This may indicate that these SNVs were present in the parent D-1 clone with lower than the detection limit frequencies.

Six out of ten SNVs in the sub-clones of the clone D1 were located in the non-coding regions. Two out of four SNVs that resulted in changes to the amino acid sequences occurred in the genes are involved in regulation of cell growth; *ARNT*, Gene Ontology (GO) category “positive regulation of endothelial cell proliferation”, and *KMT2A*, GO category “negative regulation of cell proliferation”. In the sub-clones of D-2 clone, two SNVs were found in the non-coding regions and two SNVs changed the amino acid sequences of the transcription factor encoded by the *FLII* gene, and the tyrosine protein kinase ROS, GO category “cell proliferation”. The SNV in the *ROS1* gene is found in two sub-clones (Table 2).

4. Discussion

In this study, we showed that there was a significant number of SNVs in DU145 cancer cells even in the limited portion of the genome covered by the targeted sequencing. The region covered by CCP comprises only 0.1% of the genome. Assuming a uniform rate of mutations [11], we can estimate that there are about 10^4 unique SNVs per genome between individual cells. We found that in the original culture of DU145 cells there were at least two groups of clones with just one common SNV (Figure 1). This implicates that within the original cell population there were at least two “founding” clones. Of course, a detailed analysis of clonal evolution in the cultures of DU145 cells would require sequencing significantly more than the five clones that were sequenced in this study.

In contrast, HT1080 cells with a more stable genome had only one or two SNVs between cells detected within the CCP target region, which are too few for a meaningful clonal analysis and estimations of SNVs per genome. We believe that more extensive whole exome sequencing would be required to assess clonal evolution in the cultures of HT1080 cells.

We searched the list of SNVs (Supplementary Table 1) for genes that could affect cell growth, and identified the genes in the list that match Gene Ontology (GO) category “regulation of cell proliferation”. The results are presented in Table 3. Interestingly, six out of nine SNVs that resulted in amino acid change were unique to their clones, and not present in original DU145 culture. One of these mutations, in Thr336 of the *STK11* gene has been shown to completely inactivate this tumor suppressor gene, and thus give the mutant clone growth advantage [12,13].

Six out of 15 SNVs in the sub-clones that were located in the protein coding regions. Four out of these six SNVs resulted to changes to the amino acid sequences of proteins that are

involved in regulation of cell growth; *ARNT*, *KMT2A FLII* and *ROS1*. The high proportion of the SNVs that affect cell growth indicates that the populations of DU145 cells evolve rapidly. Cells with SNVs that confer growth advantage within a few generations can take over the population.

Our results indicate that cancer cell lines could potentially model clonal evolution in tumors under various genetic and environmental backgrounds that are easy to manipulate *in vitro*. They provide additional evidence of fast clonal evolution on the SNVs in the cultures of cancer cells that may result in significant genetic variations between these cultures in different laboratories.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

1. Stratton MR, Campbell PJ, Futreal PA. The cancer genome. *Nature*. 2009; 458:719–724. [PubMed: 19360079]
2. Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio SA, Behjati S, Biankin AV, Bignell GR, Bolli N, Borg A, Borresen-Dale AL, et al. Signatures of mutational processes in human cancer. *Nature*. 2013; 500:415–421. [PubMed: 23945592]
3. Eirew P, Steif A, Khattra J, Ha G, Yap D, Farahani H, Gelmon K, Chia S, Mar C, Wan A, et al. Dynamics of genomic clones in breast cancer patient xenografts at single-cell resolution. *Nature*. 2015; 518:422–426. [PubMed: 25470049]
4. Jesinghaus M, Pfarr N, Kloor M, Endris V, Tavernar L, Muckenhuber A, von Knebel Doeberitz M, Penzel R, Weichert W, Stenzinger A. Genetic heterogeneity in synchronous colorectal cancers impacts genotyping approaches and therapeutic strategies. *Genes Chromosomes Cancer*. 2016; 55:268–277. [PubMed: 26650777]
5. Martincorena I, Campbell PJ. Somatic mutation in cancer and normal cells. *Science*. 2015; 349:1483–1489. [PubMed: 26404825]
6. Glaab WE, Tindall KR. Mutation rate at the hprt locus in human cancer cell lines with specific mismatch repair-gene defects. *Carcinogenesis*. 1997; 18:1–8. [PubMed: 9054582]
7. Rasheed S, Nelson-Rees WA, Toth EM, Arnstein P, Gardner MB. Characterization of a newly derived human sarcoma cell line (ht-1080). *Cancer*. 1974; 33:1027–1033. [PubMed: 4132053]
8. Liu X, Chen X, Rycak K, Chao HP, Deng Q, Jeter C, Liu C, Honorio S, Li H, Davis T, et al. Systematic dissection of phenotypic, functional, and tumorigenic heterogeneity of human prostate cancer cells.
9. Putnik M, Wojdacz TK, Pournara A, Vahter M, Wallberg AE. Ms-hrm assay identifies high levels of epigenetic heterogeneity in human immortalized cell lines.
10. Beheshti B, Park PC, Sweet JM, Trachtenberg J, Jewett MA, Squire JA. Evidence of chromosomal instability in prostate cancer determined by spectral karyotyping (sky) and interphase fish analysis. *Neoplasia*. 2001; 3:62–69. [PubMed: 11326317]
11. Supek F, Lehner B. Differential DNA mismatch repair underlies mutation rate variation across the human genome. *Nature*. 2015; 521:81–84. [PubMed: 25707793]
12. Boudeau J, Kieloch A, Alessi DR, Stella A, Guanti G, Resta N. Functional analysis of lkb1/stk11 mutants and two aberrant isoforms found in peutz-jeghers syndrome patients. *Hum Mutat*. 2003; 21:172. [PubMed: 12552571]
13. Bai Y, Zhou T, Fu H, Sun H, Huang B. 14-3-3 interacts with lkb1 via recognizing phosphorylated threonine 336 residue and suppresses lkb1 kinase function. *FEBS Lett*. 2012; 586:1111–1119. [PubMed: 22575644]

Figure 1. Venn diagram shows the numbers of SNVs in DU145 (A) and HT1080 (B) clones that were unique for a particular clone or shared by several but not all the clones.

Table 1

Frequencies (in percent) of SNVs in the genes of DU-145 clone D-1 sub-clones.

Gene	D-1-1	D-1-3	D-1-5	D-1-6	D-1-7	D-1-10
<i>IGF2R</i>	50%					
<i>PIK3R2</i>			50%		50%	
<i>ARNT</i>		30%		20%		
<i>STK36</i>		50%				
<i>ATM</i>		30%				
<i>KMT2A</i>				30%		
<i>MDM4</i>						30%
<i>FANCD2</i>						40%
<i>PIK3CB</i>						30%
<i>MYH11</i>						50%

Table 2

Frequencies (in percent) of SNVs in the genes of DU-145 clone D-2 sub-clones

Gene	D-2-2	D-2-6	D-2-8
<i>ETV1</i>	10%		
<i>FLI1</i>	10%		
<i>IRF4</i>		30%	
<i>ROS1</i>		50%	50%
<i>RNF213</i>		100%	

Table 3

List of genes with SNVs in DU-145 clones included in GO category “regulation of cell proliferation”. Highlighted are unique SNVs in the clones.

Gene	Mutation	Function
ARID2	p.(=)	AT-rich interactive domain-containing protein 2
JAK2	p.(=)	Tyrosine-protein kinase
ABL2	p.Met772Ile	Abelson tyrosine-protein kinase 2
IGF1R	p.(=)	Insulin-like growth factor 1 receptor
FANCA	p.(=)	Fanconi anemia group A protein
PRKAR1A	NCR	cAMP-dependent protein kinase type I-alpha regulatory subunit
PKHD1	p.Ala877Thr	Fibrocystin;PKHD1
NOTCH2	p.Pro365Leu	Neurogenic locus notch homolog protein 2
FN1	p.Thr890Ala	Fibronectin
PBRM1	p.Glu970Asp	Protein polybromo-1
NOTCH1	NCR	Neurogenic locus notch homolog protein 1
AKT1	p.(=)	RAC-alpha serine/threonine-protein kinase
TNFAIP3	NCR	Tumor necrosis factor alpha-induced protein 3
CTNNA1	NCR	Catenin alpha-1
ABL1	p.Ala1061Ser	Tyrosine-protein kinase
STK11	p.Thr336Asn	Serine/threonine-protein kinase
TSC1	p.Arg953Lys	Hamartin
TSC1	p.Leu11Ile	Hamartin
EGFR	p.(=)	Epidermal growth factor receptor