ORIGINAL ARTICLES

An algorithm to evaluate the efficacy of detecting somatic mutations

Thurai Moorthy*

MultiGEN Diagnostics, United States

ABSTRACT

Detection of somatic mutations from late stage solid tumors is a critical part of cancer treatment. Although tumor content is used as a convenient parameter to measure efficacy of detection, it fails to include two basic factors: the lower limit of detection (LLOD), and the ratio of the mutant and wild type allele frequencies. Recently, the detection of somatic mutations has expanded to liquid biopsy, early stages of cancer and population screening, which all generally carry lower copy numbers of somatic mutations compared to late stage tumors. With the growing importance of these mutations for targeted chemotherapy and other clinical applications, there is a need re-evaluate the efficacy of detection of somatic mutations. Hence, a new algorithm, Detection Index (DI), is proposed to standardize the efficacy of all molecular methods and is applicable to all types of clinical samples. DI is based on two basic determinants: lower limit of detection of the mutant allele, and the ratio of the copies of the mutant allele to that of the wild-type. The benefits of DI include (a) standardization of methods detecting somatic mutations so that laboratory reports will have a uniform interpretation related to clinical picture, and (b) the flexibility to use appropriate amounts of DNA and assay conditions to achieve desired DI.

Key Words: Detection Index, Tumor content, Mutant allele, Somatic mutations

1. INTRODUCTION

There are a number of cancer-specific somatic mutations, some of which are targets for specific chemotherapies. Others are used for guiding selection of drugs and prognosis. Somatic mutations are tested for in a variety of heterogeneous cell sample types, including solid tumors, liquid biopsy, mammalian cell culture, stool, urine, saliva, cerebrospinal fluids, lung lavage, gastric, and tissue and/or organ transplants.[1–9] Further, the genetic variants could be single nucleotide polymorphism, multiple mutations, deletion, addition of nucleotides, epigenetic changes (e.g. Methylation), gene expression, splice variants and copy number variations of all the above genetic variants.[10–13]

Targeted chemotherapy is a standard treatment in late stage cancer and is reliant on the detection of specific somatic mutations.[14] Presently, tumor content from a formalin fixed paraffin embedded tissue (FFPE) is widely used as a measure of efficacy of detection for two practical reasons: (a) tumor content is a part of the routine histological analysis, and (b) the DNA for detection is extracted from FFPE slides.[15] Thus, tumor content became a choice of convenience to score the efficacy of detection. With continued uncontested usage, the concept of tumor content gained acceptance in scientific publications as "sensitivity" defined as the percentage ratio of cancer cells to normal cells. Since adequate human DNA can be extracted from blood samples and the fact that the allele frequency of the mutant and wild type allele is

---

*Correspondence: Thurai Moorthy; Email: moorthy@multigen-diagnostics.com; Address: MultiGEN Diagnostics, United States.

equal, neither the Lower limit of detection (LLOD) nor allele competition are then considered in evaluating efficacy of detecting germline mutations. At the same time, the efficacy of molecular methods detecting pathogens is based on LLOD. In contrast, detection of somatic mutations has two challenges; detecting at lower copy numbers, and detecting in the presence of a large number of wild type alleles. Use of tumor content accounts for part of the latter. Although tumor content can be a convenient parameter, it falls short of an accurate measure of detection for two reasons:

(1) LLOD. Detection of the intended analyte is determined by the ability of the assay to detect a minimum amount of the analyte in a patient sample matrix that carries both mutant and wild type alleles, referred to as the lower limit of detection.[17] Somatic mutations, especially the ones that are associated with targeted chemotherapy, have been well studied; the molecular pathways established and supported by numerous clinical studies. Hence, the presence of actionable somatic mutations is considered a "pathological" factor. Further, detection of actionable somatic mutations is a qualitative test; it should be expected to detect at as minimum a number of copies as possible, which guarantees LLOD as a legitimate base for evaluating efficacy of detection.

(2) Tumor content assumes that the cancer cells carry two copies of mutant alleles and wild type normal cells carry two copies of wild type alleles. However, incidence of somatic mutations among the transformed cells could be either heterozygous (single copy) or homozygous (two copies).[16] Although there could be such an allele variation in FFPE extracts, analytical validation for regulatory requirements uses heterozygous status of the mutant allele. Hence, in evaluating allele competition, it is more appropriate to use the heterozygous status of the mutant allele with three times copies of the wild-type allele (one from the cancer cell and two from the wild type) rather than tumor content which implies two copies of mutant alleles and two copies of wild type allele. Further, different detection platforms, including Sanger sequencing, have their own analytical validation, with specific cut-offs of tumor content and DNA input for their respective assay output.[18–21] Such procedural variations among the assays could lead to non-uniformity of laboratory results, which could in turn make clinical correlation more difficult for desired treatment outcome.

The importance of the ratio of the number of copies of mutant alleles to that of the wild type is explained as follows; Molecular methods use primers for amplification and probes

for identification, where target specific primers and probes are supposed to bind to their respective templates. However, such binding is also affected by ionic concentrations, temperature etc.[22, 23] Use of total nucleic acid extracts from clinical samples such as tumor biopsy could create reaction conditions that allow for the possibility of cross binding, where the primers specific to the mutated allele could bind to the wild-type allele and vice versa. Therefore, the efficacy of detection of the target mutant alleles depends on the relative competitiveness of the mutant allele and wild-type allele to their respective specific primers/probes. The tumor content is a measure of the mere relative number of the transformed cells to that of the non-transformed cells and is a false measure of the competitiveness of mutant and wild-type probes and/or primers to respective alleles. The true measure of the competitiveness will be the ratio of the number of copies of mutant allele to that of wild type in the sample that competes for its specific primers and/or probes.

Since tumor content does not reflect the molecular competitiveness in the reaction, there is a need for a new way to evaluate efficacy that is based on true determinants of detection. Further, detection of somatic mutations is expanding to clinical sample types such as liquid biopsy, population screening and early stage cancers, which could carry low numbers of copies of mutations.[24, 25] Hence, there is a necessity for more sensitive tests to address new emerging clinical needs. To have both the present and the more sensitive assays on the same playing field, there is a need for a new evaluation parameter. This manuscript outlines a new algorithm, Detection Index (DI), that reflects the true measure of efficacy of detection.

2. MATERIALS AND METHODS

It is summarized that two basic factors that determine the efficacy of detection of the mutant allele are:

(1) The number of copies of mutant alleles in the reaction. The fewer number of copies of mutant allele needed for detection, the more efficient the assay.
(2) The number of copies of wild-type alleles in the reaction. The assay that detects the mutant allele in the presence of higher number of copies of wild-allele, the more efficient the assay.

Polymerase chain reaction (PCR) is an integral part of companion diagnosis and is performed in an optimum reaction volume to avoid nonspecific amplification, while at the same time to be effective in amplifying the intended target DNA. However, for practical reasons, present companion diagnosis is performed in small volumes (5 μl - 20 μl), which could compromise the overall assay.[15] Since reaction volume is

variable, for evaluation purposes, it is more appropriate to have (1) and (2) per μl.

Now that the two determinants have been characterized, one could combine them into an algorithm, DI. The fewer the number of copies of the mutant allele necessary for detection, the more efficient the assay. Hence the efficacy of detection is inversely proportional to the number of copies of mutant allele/μl. Further, the assay is more efficient if the ratio of mutant: wildtype alleles is lower; hence DI is inversely proportional to the ratio of the mutant: wildtype alleles.

3. RESULTS AND DISCUSSION

On this basis, a DI algorithm is derived as follows:

$$\text{Copies of Mutant Allele}/\mu l = \frac{\text{Amount of DNA} \times 290 \times \% \text{tumor content}}{\text{Reaction volume}(\mu l) \times 100}$$

$$\text{Copies of Wildtype allele}/\mu l = \frac{(\text{Amount of DNA} \times 290 \times \% \text{tumor content}) + [(\text{Amount of DNA} \times 290 \times 100 - \% \text{Tumor content}) \times 2]}{\text{Reaction volume}(\mu l) \times 100}$$

Note: 1 ng of human DNA contains 290 haploid copies; Amount of DNA in ng.

Table 1. Varying tumor content with 40 ng of DNA in 20 μl reaction volume

% Tumor content	Copies of Mutant allele/ul	Copies of Wildtype allele/ul	% Allele Ratio [#]	DI
1	5.8	1,154	0.5	3,448.28
2	11.6	1,148	1.0	862.07
5	29.0	1,131	2.5	137.93
10	58.0	1,102	5.0	34.48
20	116.0	1,044	10.0	8.62
30	174.0	986	15.0	3.83
40	232.0	928	20.0	2.16
50	290.0	870	25.0	1.38
60	348.0	812	30.0	0.96
70	406.0	754	35.0	0.7
80	464.0	696	40.0	0.54
90	522.0	638	45.0	0.43
100	580.0	580	50.0	0.34

[#] (Copies of mutant alleles / copies of total wildtype alleles) × 100

The algorithm is illustrated in detail with the following scenarios.

(1) As shown in Table 1, with a fixed amount of DNA (e.g. 40 ng) input, while there is a proportional correlation of tumor content and the percent allele ratio at low tumor content, both depart disproportionately at higher tumor content (see Figure 1a). A similar pattern is found when tumor content is compared to the mutant alleles (see Figure 1b). Hence, the

$$DI^{\text{Het}} = (1/MC) \times (1/\text{Ratio of MC: WC})$$

Somatic mutations could be either heterozygous or homozygous and for evaluation purposes we propose the heterozygous scenario. Hence it is assumed that in tumor biopsy samples, the wild type alleles will be three times that of the mutant alleles. MC is the number of copies of mutant allele/μl and WC is the number of copies of wild type allele/μl and the detection index is denoted as DI^{Het}.

Calculation of DI requires only two data points, the tumor content and the amount of DNA used per reaction. Both the number of copies of the mutant allele/μl and wild-type allele/μl can be calculated using:

tumor content is not the same as the ratio of mutant allele to that of the total wild-type alleles.

Detection of somatic mutations is not a standardized procedure; hence there are variations in all key steps including DNA extraction (sample preparation). These variations also include using varying numbers of FFPE sections per assay, and varying thickness (5 mm-10 mm) of the FFPE sections, thus resulting in varying amounts of DNA input per reaction. With the increasing amount of DNA per assay, the association between DNA input and ratio of the wild-type allele is disproportionate (see Figure 1c).

Table 2. Varying DNA with 5% tumor content in 20 μl reaction volume

DNA input	Copies of Mutant allele/ul	Copies of Wildtype allele/ul	% Allele Ratio (M:W) [#]	DI
10	7	282.75	2.5	551.72
20	15	565.50	2.5	275.86
30	22	848.25	2.5	183.91
40	29	1,131.00	2.5	137.93
50	36	1,413.75	2.5	110.35
60	44	1,696.50	2.5	91.95
75	54	2,120.63	2.5	73.56
90	65	2,544.75	2.5	61.3
100	73	2,827.50	2.5	55.17
250	181	7,068.75	2.5	22.07
500	363	14,137.50	2.5	11.03
750	544	21,206.25	2.5	7.36
1000	725	28,275.00	2.5	5.52

[#] Copies of mutant alleles / copies of non-mutant alleles

Table 3. Detection Index of Companion Diagnostics Tests

Method	Regulatory	Platform	Mutation	DNA (ng)	% Tumor content	Mutant allele/µl	Wild allele/µl	% Allele Ratio (M:W)	DI
Foundation One CDX (a)	FDA	NGS	EGFR -L858R	50	2.1	15.2	1435	1.05	62.55
Foundation One CDX (a)	FDA	NGS	EGFR -Del	50	5.1	37.0	1413	2.55	10.61
Foundation One CDX (a)	FDA	NGS	EGFR -T790M	50	2.5	18.1	1432	1.25	44.14
Foundation One CDX (a)	FDA	NGS	Braf V600E	50	2	14.5	1436	1.00	68.97
Foundation One CDX (a)	FDA	NGS	Kras G12/G13	50	2.3	16.7	1433	1.15	52.15
Oncomine™ Dx Target Test (a)	FDA	NGS	EGFR -L858R	10	8	11.6	278	4.00	21.55
Oncomine™ Dx Target Test (a)	FDA	NGS	EGFR -Del	10	6	8.7	281	3.00	38.31
Oncomine™ Dx Target Test (a)	FDA	NGS	Braf V600E	10	12	17.4	273	6.00	9.58
Paraxis™	FDA	NGS	Extended Ras panel	50	5	36.3	1414	2.50	11.03
Cobas (a)	FDA	Real time PCR	Braf V600E	125	5	90.6	2828	3.11	3.55
Cobas (a)	FDA	Real time PCR	Kras (G12D,G13D, G12V)	50	5	29.0	1131	2.50	13.79
Cobas (a)	FDA	Real time PCR	EGFR (L858R, Del747-A750, T790M)	50	5	36.3	1414	2.50	11.03
THxID (a)	FDA	Real time PCR	Braf V600E	20	80	232.0	348	40.00	0.11
Pyrosequencing	LDT	Sequencing	EGFR T790M	30	50	217.5	653	25.00	0.18
ddPCR	LDT	PCR	Braf p.V600E	30	50	217.5	653	25.00	0.18
SNapShot	LDT	Primer extension	Braf p.V600E	30	50	217.5	653	25.00	0.18
Tumorplex-MutiGEN	LDT	Sequencing (ASMS)	Braf p.V600E	2.5	1	0.15	29	0.50	13793.10

(a) FDA- https://www.fda.gov/MedicalDevices/ProductsandMedicalProcedures/InVitroDiagnostics/ucm301431.htm

(2) With fixed DNA input per assay (40 ng) and with varying tumor content, the number of copies of mutant allele/µl varies from 5.8 copies/µl to 580 copies/µl (see Table 1). On the other hand, with fixed tumor content (5%) per assay and varying amount of DNA input, the number of copies of mutant allele/µl varies from 7 copies/µl to 725 copies/µl (see Table 2). Hence, an assay capable of detecting positives (29 copies of mutant allele/µl) with 40 ng of DNA at 5% tumor content will report a sample that carries 50 mutant copies/µl as positive, whereas the assay that is capable of detecting (348 copies/µl) with 40 ng of DNA at 60% tumor content will report the same sample as negative. Similarly, an assay capable of detecting somatic mutations with 5% tumor content and 50 ng of DNA could only detect in a sample

that has more than 36 mutant allele/µl. Therefore, this same assay would report a sample with 5% tumor content using 10 ng of DNA as negative. Hence, having tumor content as a measure of accuracy could lead to false negatives, resulting in patients being prevented from receiving beneficial targeted chemotherapy. These examples show that there are discrepancies among methods determining the efficacy of detection based on the tumor content or the amount of DNA input alone.

(3) Two sets of DI are shown (see Tables 1 and 2), one with 40 ng of DNA input with varying tumor content where the Detection Index decreases with increase in tumor content, and the other with varying DNA input at 5% tumor content where the DI decreases with an increase in DNA input. With

this interactive table for a desired detection index, one could determine either the amount of DNA needed if tumor content is known or vice versa.

(4) Examples of DI values are shown in Table 3. These include (a) DI values of FDA approved companion diagnostics tests, and (b) DI values of several laboratory developed tests (LDTs). Although the amount of DNA and tumor content vary among LDTs, DI values were calculated using commonly used amounts (30 ng of DNA with 50% tumor

content) (see Table 3). Realtime PCR uses labeled probes for identification. To distinguish the target signal from the background noise, there is a need to have high mutant allele/μl. Hence, they tend to have very low DI. Since NGS is designed to detect many markers per reaction, the DNA input is relatively high, increasing the copies of mutant allele/μl, and lowing the DI. The Braf p. V600E/K Tumorplex assay uses 1ng of DNA with 1% tumor content where the mutant allele copies are low, that in turn increases the DI to 13793.10.
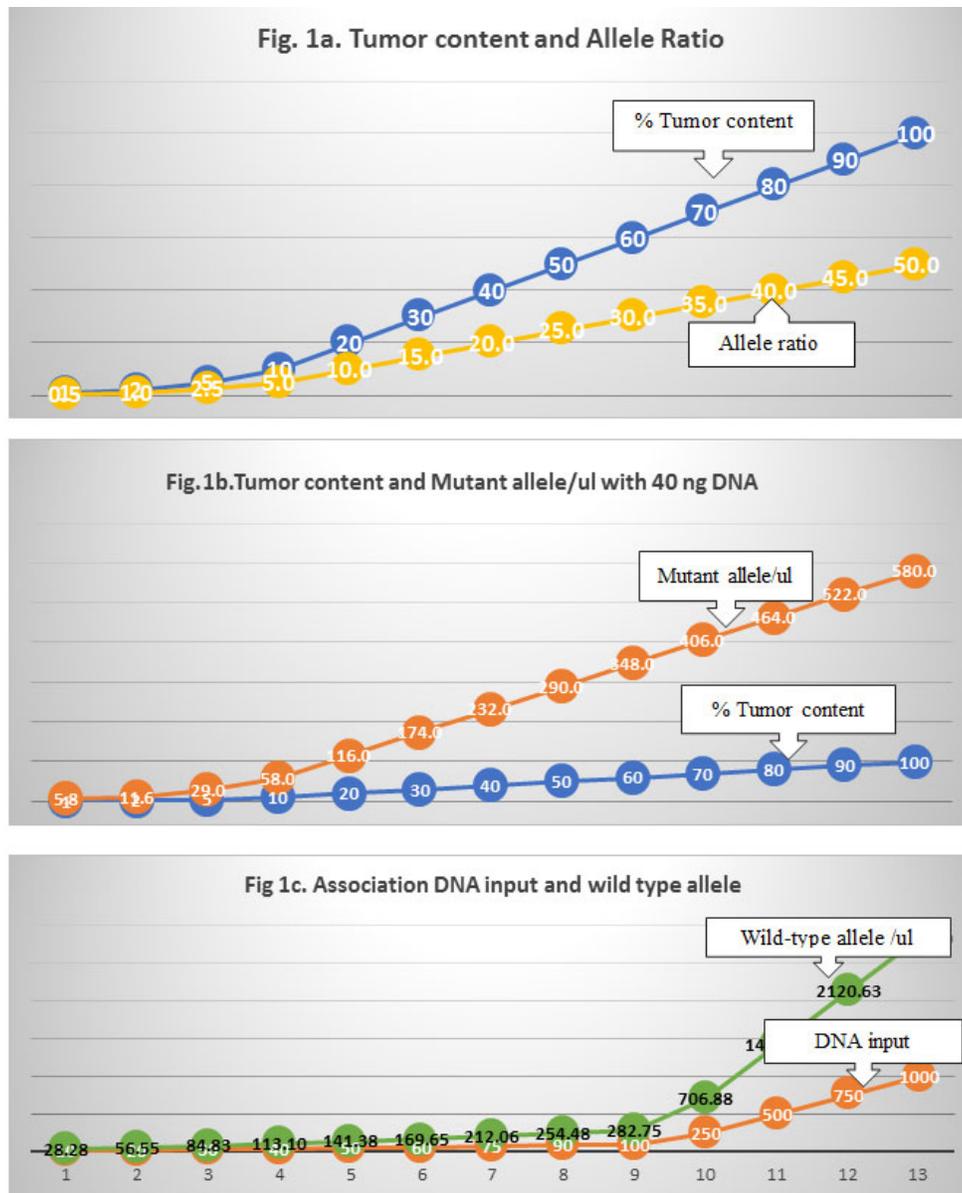
Figure 1. Correlation between different factors of the FFPE sample. 1a. The allele ratio does not proportionately increase with tumor content. 1b. Mutant allele does not proportionately increase with tumor content. 1c. Although at low concentration wild type allele correlates with DNA input, at higher DNA concentration they do not

The choice of targeted chemotherapy depends on the detection of specific somatic mutations; hence the detection of respective somatic mutations becomes very critical, carrying the burden of accuracy for patients fighting for recovery. The proposed Detection Index helps the laboratory to provide more accurate results that will help the clinicians to improve treatment. Just like in clinical chemistry (e.g. determination of blood sugar), use of Detection Index will enable companion diagnosis to define its limit of efficacy using two data points. Operating within such a limit will eliminate potential false negatives or false positives of the test. Since Detection Index is a true scientific evaluation of the tests, clinical data interpretation and improved treatment strategies can be adopted.

$$\text{Amount of DNA} = \frac{\text{Copies of Mutant Allele} \times 100}{290 \times \%\text{tumor contents}}$$

Hence, with the interactive DI table, the end user has the freedom over the design of the assay. For example, once the desired Detection Index is decided, one could determine the amount of DNA input based on the tumor content.

4. Conclusion

The increasing clinical importance of somatic mutations calls for second look at the efficacy of detection. Tumor content has been a convenient but crude evaluation criterion that fails to accommodate basic factors of detecting mutant alleles in a heterogeneous cell population. Hence a new criterion, DI, is proposed that is based on fundamental scientific determinants in detecting somatic mutations in a clinical heterogeneous cell population.

Conflicts of Interest Disclosure

The author declares that there is no competing interest.

References

[1] Johnson GE. Mammalian cell HPRT gene mutation assay: test methods. Methods Mol Biol. 2012; 817: 55-67. PMid:22147568. https://doi.org/10.1007/978-1-61779-421-6_4

[2] Imperiale TF, Ransohoff DF, Itzkowitz SH, et al. Multitarget Stool DNA Testing for Colorectal-Cancer. N. Engl J Med. 2014; 370: 1287-1297.

[3] Dahmcke CM, Steven KE, Larsen LK, et al. A Prospective Blinded Evaluation of Urine-DNA Testing for Detection of Urothelial Bladder Carcinoma in Patients with Gross Hematuria. Eur Urol. 2016 Dec; 70(6): 916-919. https://doi.org/10.1016/j.eururo.2016.06.035. Epub 2016 Jul 11.

[4] Cohen N, Gupta M, Doerwald-Munoz L, et al. Developing a new diagnostic algorithm for human papilloma virus associated oropharyngeal carcinoma: an investigation of HPV DNA assays. J. Otolaryngol. Head Neck Surg. 2017 Feb 13; 46(1): 11. https://doi.org/10.1186/s40463-017-0189

[5] Huang TY, Piunti A, Lulla RR, et al. Detection of Histone H3 mutations in cerebrospinal fluid-derived tumor DNA from children with diffuse midline glioma Acta Neuropathol. Commun. 2017 Apr 17; 5(1): 28. https://doi.org/10.1186/s40478-017-0436-6

[6] Park S, Hur JY, Lee KY, et al. Assessment of EGFR mutation status using cell-free DNA from bronchoalveolar lavage fluid. Clin Chem Lab Med. 2017 Aug 28; 55(10): 1489-1495. https://doi.org/10.1515/cclm-2016-0302

[7] Zuo Y, Lv Y, Qian X, et al. Inhibition of HHIP Promoter Methylation Suppresses Human Gastric Cancer Cell Proliferation and Migration. Cell Physiol Biochem. 2018; 45(5): 1840-1850. https://doi.org/10.1159/000487875. Epub. 2018 Feb 28.

[8] Maemondo M, Inoue A, Kobayashi K, et al. Gefitinib or chemotherapy for non-small-cell lung cancer with mutated EGFR. N Engl J Med. 2010; 362: 2380-2388. PMid:20573926. https://doi.org/10.1056/NEJMoa0909530

[9] Verghese PS, Schmeling DO, Filtz EA, et al. Transplantation of solid organ recipients shedding Epstein-Barr virus DNA pre-transplant: A prospective study. Clin Transplant. 2017 Nov; 31(11). https://doi.org/10.1111/ctr.13116. Epub 2017 Oct 8.

[10] Cibulskis K, Lawrence MS, Carter SL, et al. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. Nature Biotechnology. 2013; 31: 213-219. PMid:23396013. https://doi.org/10.1038/nbt.2514

[11] Pan Q, Shai O, Lee LJ, et al. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. Nat Genet. 2008; 40: 1413-1415. PMid:18978789. https://doi.org/10.1038/ng.259

[12] Anders HL, Lohuizen MV. Epigenetics and cancer. Gene Development. 2004; 18: 2315-2335. PMid:15466484. https://doi.org/10.1101/gad.1232504

[13] Smadbeck JB, Johnson SH, Smoley SA, et al. Copy Number Variant Analysis using Genome-Wide Mate-Pair Sequencing. Genes Chromosomes Cancer. 2018.

[14] Qu K, Pan Q, Zhang X, et al. Detection of BRAF V600 Mutations in Metastatic Melanoma Comparison of the Cobas 4800 and Sanger Sequencing Assays. J. Mol. Diag. 2013; 15: 790-795.

[15] Pikor LA, Enfield KS, Cameron H, et al. DNA extraction from paraffin embedded material for genetic and epigenetic analyses. J Vis Exp. 2011; (49): 2763. Published online 2011 Mar 26. https://doi.org/10.3791/2763

[16] Improta1 G, Pelosi G, Tamborini E, et al. Biological insights into BRAFV600 mutations in melanoma patient. OncoImmunology. 2013; 2(8): e25594.

[17] David A, Armbruster TP. Limit of Blank, Limit of Detection and Limit of Quantitation. Clin Biochem. Rev. 2008; 29: S49-52.18.

[18] Weyant GW, Wisotzkey JD, Benko FA, et al. BRAF mutation testing in solid tumors. A methodological comparison. J. Mol. Diag. 2014; 16(5).

[19] Vinayagamoorthy T, Zhang D, Ye F, et al. Can detection of Braf p. V600E mutation be improved? Comparison of allele specific multiplex sequencing to present tests. Journal of Solid Tumors. 2017; 7(2): 14-22. https://doi.org/10.5430/jst.v7n2p14

[20] Available from: https://www.accessdata.fda.gov/cdrh_docs/pdf12/P120014B.pdf

[21] Available from: `https://www.accessdata.fda.gov/cdrh_docs/pdf11/P110020B.pdf`

[22] Bessekri MW, Aggoune A, Lazreg S, et al. Comparative study on the effects of reduced PCR reaction volumes and increased cycle number, on the sensitivity and the stochastic threshold of the AmpFlSTR Identifiler Plus kit. Forensic Science International: Genetics Supplement Series. 2013; 4(1).

[23] Owczarzy R, Moreira BG, You Y, et al. Predicting Stability of DNA Duplexes in Solutions Containing Magnesium and Monovalent Cations. Biochemistry. 2008; 47: 5336-5353. PMid:18422348. `https://doi.org/10.1021/bi702363u`

[24] Haber DA, Velculesc VE. Blood-Based Analyses of Cancer: Circulating Tumor Cells and Circulating Tumor DNA. Cancer Discov. 2014; 4(6): 650-661. PMid:24801577. `https://doi.org/10.1158/2159-8290.CD-13-1014`

[25] Ellen H, Samantha P, Jochen BG, et al. The potential of liquid biopsies for the early detection of cancer. Precision Oncology. 2017; 1: 36. PMid:29872715. `https://doi.org/10.1038/s41698-017-0039-5`