Published in final edited form as:
Mol Diagn Ther. 2014 October ; 18(5): 587–593. doi:10.1007/s40291-014-0115-2.

Cytosine Deamination is a Major Cause of Baseline Noise in Next Generation Sequencing

Guoli Chen, MD/PhD[1], **Stacy Mosier**[1], **Christopher D. Gocke**[1,2], **Ming-Tseh Lin, MD/PhD**[1], and **James R. Eshleman, MD/PhD**[1,2]

[1]Department of Pathology, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA

[2]Department of Oncology, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA

Abstract

Background and Objectives—As next generation sequencing (NGS) becomes a major sequencing platform in clinical diagnostic laboratories, it is critical to identify artifacts that constitute baseline noise and may interfere with detection of low level gene mutations. This is especially critical for applications requiring ultrasensitive detection, such as molecular relapse of solid tumors and early detection of cancer. We recently observed approximately ~10-fold higher C:G > T:A mutations than other background noise in both wild type peripheral blood and formalin fixed paraffin embedded (FFPE) samples. We hypothesized that these might represent cytosine deamination events that have been seen using other platforms.

Methods—To test the hypothesis, we pretreated samples with uracil N-glycosylase (UNG). Additionally, we also tested whether some of the cytosine deamination might be laboratory artifact by simulating the heat associated with PCR thermocyling by subjecting samples to thermocycling in the absence of polymerase. To test the safety of universal UNG pretreatment, we tested known positive samples treated with UNG.

Results—UNG pretreatment significantly reduced these mutations, consistent with a biologic source for the cytosine deamination. The simulated thermocycling heated samples demonstrated significantly increased C:G > T:A mutations without affecting other baseline base substitutions. Samples with known mutations demonstrated no decrease in our ability to detect these after treatment with UNG.

Conclusions—Baseline noise during NGS is mostly due to cytosine deamination, the source of which is likely both biologic and an artifact of thermocycling, and it can be reduced by UNG pretreatment.

Corresponding author/reprint requests: James R. Eshleman, MD, PhD, The Sol Goldman Pancreatic Cancer Research Center, Suite 344, CRB-II, 1550 Orleans St., Baltimore, MD 21231, Tel: 410-955-3511; Fax: 410-614-0671; jeshlem@jhmi.edu.

Conflicts of Interest Guoli Chen, Stacy Mosier, Christopher D. Gocke, Ming-Tseh Lin and James R. Eshleman have no conflicts of interest that are directly relevant to the content of this study.

1 Introduction

Next Generation Sequencing (NGS) has been revolutionizing biomedical research and clinical practice for the past nine years since its introduction[1]. With substantial improvement in accuracy, read length, depth of coverage, and remarkable reduction in cost, NGS is becoming a major sequencing platform for clinical diagnostic laboratories. In the field of oncology, NGS has allowed detection of all potential driver gene mutations that cause a particular malignancy and thus clinically report the status of all genes known to predispose to a particular type of cancer[2–7]. To embrace the era of targeted therapy and personalized medicine, NGS will be an essential tool to evaluate all related genes that predict drug response and clinical outcome, and offer us comprehensive genetic profiling especially if the target-based treatment covers mutations in multiple gene/pathway family members[8, 9].

A promising clinical area in which NGS will likely play a role is minimal residual disease (MRD) testing of solid tumors. This will allow clinicians to initiate second or third line therapy more quickly based on residual tumor molecules, rather than waiting until the tumor progresses radiographically[10, 11]. On the other hand, a negative MRD test could be used to avoid the risk and expense associated with additional chemotherapy, when the patient is actually in remission.

Due to its extremely high depth of coverage, NGS will likely also be the tool of choice for early detection of cancer by measuring critical oncogene activating mutations[12], and structural abnormalities[13, 14]. While NGS holds tremendous promise for clinical use, it also faces challenges from technical complexity and result interpretation. Reproducible sequence artifacts constitute the baseline noise of sequencing and may interfere with detection of true gene mutations. This is a critical concern to address when extremely low levels of mutations need to be unambiguously detected for these and other applications.

Cytosine deamination has been demonstrated to contribute to background noise for DNA sequencing of ancient and formalin fixed paraffin embedded (FFPE)-treated DNA[15, 16]. Since it can manifest as either a base substitution of C to T (C>T) on the sense strand, or G>A mutations on the sense strand (arising from a C>T deamination on the antisense strand), these mutations are collectively designated C:G >T:A. Williams and colleagues observed formalin fixation-induced C:G>T:A transitions[17]. However, the same artifact was also observed in freshly prepared samples in a recent study validating detection of gene mutations within the Ion AmpliSeq Cancer Hotspot Panel using the Ion Torrent Personal Genome Machine™ in our laboratory[18]. We observed significantly higher C:G > T:A mutations than other background noise during NGS in both normal peripheral blood and FFPE samples by checking all common *KRAS* (codons 12 and 13), *BRAF* (V600E) and *EGFR* (T790M and L858R) gene point mutations. These C:G > T:A mutations must be either biologic (intrinsic to the sample prior to isolation) or an artifact of the molecular biology, including DNA isolation, PCR amplification and/or sequencing.

During routine NGS, we noted an increase in C:G>T:A transitions in normal samples and in this report, we identify the etiology of NGS baseline noise, explore its contributing factors and provide a partial solution. To confirm cytosine deamination as the source, we treated

DNA from peripheral blood with uracil N-glycosylase (UNG) and demonstrated that it significantly reduced the frequency of C:G >T:A mutations by checking all common *KRAS* (codons 12 and 13), *BRAF* (V600E), *EGFR* (T790M and L858R) gene point mutations. Consistent with a previous finding that prolonged heating may induce deamination of DNA[19], we found that the heat associated with thermocycling induced a significant increase in the C:G >T:A mutation level. We next showed that UNG pretreatment of positive control samples does not interfere with capacity of NGS to detect real mutations. Finally, we attempted to include a thermostable UNG in PCR reactions, but were unable to identify conditions that would allow both enzymes to work.

2 Materials and Methods

2.1 Materials

Under IRB approval, the specimens consisted of 4 peripheral blood specimens from normal donors, and 7 FFPE samples from patients carrying distinctive *EGFR, KRAS* and *BRAF* gene mutations. DNA was isolated as described previously[20, 21]. DNA concentration was determined by Qubit 2.0 Fluorometer (Life Technologies, Carlsbad, California). UNG enzyme (Life Technologies), 0.5μL (1 unit/μL) was added into each reaction (30ng DNA, 20 μL for total volumes) and incubated for 30 minutes at 50°C prior to thermocycling for library preparation.

2.2 Next generation sequencing (NGS) platform

The NGS was conducted using the Ion AmpliSeq Cancer Hotspot Panel (v2) for targeted multi-gene amplification as described in a previous study (18). Ion AmpliSeq Library Kit 2.0 was used for library preparation (PCR thermocycling 65°C for 30 minutes, 95°C for 2 minutes, 20 cycles of 95°C for 15 seconds and 60°C for 4 minutes with a hold at 10°C), Ion PGM Template OT2 200 Kit and Ion OneTouch ES Instrument for emulsion PCR and enrichment, Ion PGM Sequencing 200 Kit v2, Ion 318 Chips, and the Personal Genome Machine (PGM) sequencing platform for next generation sequencing (Life Technologies), as recommended by the manufacturers' protocols without modification. The DNA input for targeted multi-gene PCR was 30ng. Eight specimens were barcoded using Ion Xpress Barcode Adapters (Life Technologies), pooled and run on a single Ion 318 chip. For samples treated with extra heat, we cycled them at 99°C for 20 minutes, with 40 cycles of 99°C for 2 minutes and 60° for 4 minutes and a hold at 10°C prior to library preparation.

2.3 Data Analysis

Sequencing data were analyzed using Torrent Suite (Version 3.2.0) (Life Technologies). Frequency (%) of all of the common *KRAS* (codons 12 and 13), *BRAF* (V600E) and *EGFR* (T790M and L858R) gene point mutations and those at five randomly picked base positions with nucleotide G or C in amplified regions of chromosomes 7 and 12 were calculated. The C:G >T:A mutations include *KRAS* G12D (GGT>GAT), G12S (GGT>AGT), G13D (GGC>GAC), G13S (GGC>AGC), and *EGFR* T790M (ACG>ATG). The non-C:G >T:A mutations include *KRAS* G12A (GGT>GCT), G12C (GGT>TGT), G12R (GGT>CGT), G13A (GGC>GCC), G13C (GGC>TGC), G13R (GGC>CGC), and *BRAF* V600E (GTG>GAG).

3 Results

3.1 C:G >T:A mutation levels in peripheral blood are reduced by uracil DNA glycosylase

In a previous study, we found that C:G→T:A mutations were significantly higher than other mutations in both peripheral blood and FFPE specimens (18). As shown in Fig. 1a, the levels of C:G >T:A mutations including *KRAS* G12D, G12S, G13D, G13S, *EGFR* T790M and those at five randomly picked positions in PCR amplicons (dots) were significantly (about 8-fold) higher than other baseline noise (diamonds) (Fig. 1a and 1d, $p<0.01$). To test whether cytosine deamination had occurred prior to library construction, we treated normal peripheral blood specimens with uracil DNA glycosylase (UNG) before we conducted the initial AmpliSeq PCR. The glycosylase activity of UNG excises the uracil base from DNA, leaving the sugar-phosphate backbone intact, thereby functionally removing that strand from the PCR reaction as the polymerase cannot synthesize across the abasic site. UNG treatment reduced the C:G >T:A mutations at all of the above sites (dots) except for two of them (arrowheads, Fig. 1b and 1c), demonstrating that some of the C:G >T:A mutations arise from deamination of cytosine to uracil prior to PCR. The overall reduction in C:G >T:A mutations is approximately 30% and statistically significant (Fig. 1d, $p<0.05$), although the C:G >T:A mutations following UNG treatment were still significantly higher than the other mutations (Fig. 1d, $p<0.01$). In FFPE specimens, we also observed a 22% reduction of the mutation levels, but without statistical significance probably due to higher levels of variation in FFPE samples (data not shown).

3.2 Heat associated with thermocycling induces deamination

Noticing a previous observation that prolonged heat could induce cytosine deamination of DNA (19), we hypothesized that the denaturation phase of thermocycling during library prep and emulsion PCR might cumulatively cause such an effect. To test this, we thermocycled DNA from peripheral blood specimens (without performing PCR) prior to library preparation. As shown in Fig. 2, this treatment induced additional deamination effects at all susceptible positions (dots) except two of them (arrowheads, Fig. 2b and 2c). Compared to the statistically significant increase in the overall C:G >T:A mutations caused by additional thermocycling (Fig. 2d, $p<0.05$), the level of other baseline mutations (diamonds) at non-cytosine positions was not obviously affected by the treatment (Fig. 2d, $p>0.05$).

3.3 UNG does not interfere with detection of real mutations

Since UNG reduces deamination-related mutations, it is a potential tool to lower NGS background noise levels. With this consideration, we wanted to exclude the possibility that UNG may reduce the ability of NGS to detect bona fide gene mutations. We selected 7 positive control FFPE samples carrying a wide range of mutant allele frequency (Table 1) of distinct *EGFR*, *KRAS* and *BRAF* mutations. We then treated these DNA samples with UNG prior to library preparation. Compared to results from untreated specimens, the percent of the mutation detected after UNG treatment was consistent with the previously determined mutation frequency. Thus, UNG treatment does not interfere with the capability of NGS to detect the clinically important mutations (Fig. 3, Table 1), consistent with the findings of others[16].

4 Discussion

In this study, we demonstrate that the baseline noise of NGS is mainly attributed to cytosine deamination and that the source is both biologic and thermocycling induced. UNG eliminates the uracil that results from deamination and thus is a tool to reduce biologic background noise in NGS. Another possible method to eliminate PCR induced mutations should be to avoid heating by employing isothermal amplification technologies. Pretreating samples with UNG does not inhibit the ability to detect known positive control mutations. Our conclusions are all based on the Ion Torrent platform, however our findings are consistent with similar work on the MiSeq system analyzing FFPE samples[22].

NGS is a powerful tool to discover novel disease-related genetic variations, to clinically diagnose and predict disease based on comprehensive genetic profiling and to reveal therapeutic targets. Extremely high depth of coverage allows NGS to be highly sensitive, accordingly suitable for discovery of rare genetic variants, including early detection of cancer and monitoring MRD in cancer patients. In conducting NGS, reproducible sequence artifacts may produce false positives or interfere with detection of true gene mutations. For early detection of cancer or MRD monitoring, where extremely low levels of mutations need to be unambiguously identified, non-specific background noise will need to be minimized, if not eliminated.

Cytosine deamination is actually one of the most prevalent point mutations spontaneously occurring in nature, thereby contributing to background noise for sequencing[15, 16]. The two major underlying mechanisms include the deamination of 5-methylcytosine resulting in thymine and ammonia. In DNA, this reaction can be corrected by the enzyme thymine-DNA glycosylase prior to passage of the replication fork, otherwise a cytosine to thymine base substitution is generated[23, 24]. The other mechanism of deamination involves the hydrolysis of cytosine into uracil. This deamination in DNA is corrected by the DNA glycosylase UNG that removes the uracil base to generate an abasic site, which is then repaired by adding back a cytosine opposite the guanine. However, if the pro-mutagenic G˙U mispair is not repaired prior to the next round of DNA replication, a U:A mutation is generated[23, 24], which results in a T:A during the next round of synthesis.

Deamination can be attributed to multiple factors including biologic (intrinsic to the sample prior to isolation) or an artifact of the molecular biology[23–25]. For studies of ancient DNA, it is a major source of sequencing artifact[26]. In addition to age, it was observed that formalin induced C:G>T:A transitions in Sanger sequencing[17]. However, we identified the same artifact in freshly prepared samples in a recent study validating detection of gene mutations using NGS[18]. This adds evidence that deamination may also result from polymerase induced errors along with the lack of DNA repair or, directly from the heat associated with thermocycling[19]. Nevertheless, it appears that these mutations are higher in the FFPE samples than in peripheral blood[18] indicating that the process of specimen fixation also contributes to the noise level.

In this study, we demonstrated that biological deamination contributes to the C:G >T:A mutations in background noise given that treating peripheral blood samples with UNG prior

to NGS led to a significant reduction of the C:G >T:A mutation level. While we favor a biologic source to explain the reduction by UNG treatment, we cannot eliminate the possibility that deamination is induced during DNA isolation. However, the reduction is only approximately 30% (Fig. 1d), suggesting that the other 70% of the mutations are already fixed (fully converted to T:A) prior to DNA isolation, or it is occurring during the process of PCR. In this regard, we found that the heat from denaturation phase of thermocycling induced a significant increase in this background noise, consistent with the prolonged heating used by Ehrlich et al[19]. To test whether it was solely an artifact of PCR, we attempted to add a thermostable UNG from an extreme thermophile organism *Archaeoglobus fulgidus*[27] during PCR. However, we were unable to identify conditions where this enzyme and the polymerase were both active (data not shown). An alternative approach may be to use an alternate thermostable UNG[28] or to replace traditional PCR with isothermal amplification.

5 Conclusion

In summary, a major cause of baseline noise in NGS is cytosine deamination. This appears to be pre-analytic (i.e. biologic in origin), but can also be induced by the heat associated with thermocycling. Routine use of UNG pretreatment and isothermal amplification are viable strategies to reduce the background noise level of NGS.

Acknowledgments

We acknowledge Dr. James Stivers (Johns Hopkins University School of Medicine) for helpful discussions. Grant Support: Women's Board of The Johns Hopkins Hospital (MTL), National Institutes of Health R21HG005745 (CDG), R21CA164592 (JRE), Pancreatic Cancer Action Network Innovation Award (JRE).

Literature Cited

1. Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, et al. Genome sequencing in microfabricated high-density picolitre reactors. Nature. 2005; 437:376–380. [PubMed: 16056220]
2. Jones S, Zhang X, Parsons DW, Lin JC, Leary RJ, Angenendt P, et al. Core signaling pathways in human pancreatic cancers revealed by global genomic analyses. Science. 2008; 321:1801–1806. [PubMed: 18772397]
3. Biankin AV, Waddell N, Kassahn KS, Gingras MC, Muthuswamy LB, Johns AL, et al. Pancreatic cancer genomes reveal aberrations in axon guidance pathway genes. Nature. 2012; 491:399–405. [PubMed: 23103869]
4. Wu J, Matthaei H, Maitra A, Dal Molin M, Wood LD, Eshleman JR, et al. Recurrent GNAS mutations define an unexpected pathway for pancreatic cyst development. Sci Transl Med. 2011; 3:92ra66.
5. Jiao Y, Shi C, Edil BH, de Wilde RF, Klimstra DS, Maitra A, et al. DAXX/ATRX, MEN1, and mTOR pathway genes are frequently altered in pancreatic neuroendocrine tumors. Science. 2011; 331:1199–1203. [PubMed: 21252315]
6. Dal Molin M, Hong SM, Hebbar S, Sharma R, Scrimieri F, de Wilde RF, et al. Loss of expression of the SWI/SNF chromatin remodeling subunit BRG1/SMARCA4 is frequently observed in intraductal papillary mucinous neoplasms of the pancreas. Hum Pathol. 2012; 43:585–591. [PubMed: 21940037]
7. Pritchard CC, Smith C, Salipante SJ, Lee MK, Thornton AM, Nord AS, et al. ColoSeq provides comprehensive lynch and polyposis syndrome mutational analysis using massively parallel sequencing. J Mol Diagn. 2012; 14:357–366. [PubMed: 22658618]

8. van der Heijden MS, Brody JR, Dezentje DA, Gallmeier E, Cunningham SC, Swartz MJ, et al. In vivo therapeutic responses contingent on Fanconi anemia/BRCA2 status of the tumor. Clin Cancer Res. 2005; 11:7508–7515. [PubMed: 16243825]
9. Turner NC, Lord CJ, Iorns E, Brough R, Swift S, Elliott R, et al. A synthetic lethal siRNA screen identifying genes mediating sensitivity to a PARP inhibitor. EMBO J. 2008; 27:1368–1377. [PubMed: 18388863]
10. Diehl F, Schmidt K, Choti MA, Romans K, Goodman S, Li M, et al. Circulating mutant DNA to assess tumor dynamics. Nat Med. 2008; 14:985–990. [PubMed: 18670422]
11. Heitzer E, Auer M, Gasch C, Pichler M, Ulz P, Hoffmann EM, et al. Complex tumor genomes inferred from single circulating tumor cells by array-CGH and next-generation sequencing. Cancer Res. 2013; 73:2965–2975. [PubMed: 23471846]
12. Kanda M, Knight S, Topazian M, Syngal S, Farrell J, Lee J, et al. Mutant GNAS detected in duodenal collections of secretin-stimulated pancreatic juice indicates the presence or emergence of pancreatic cysts. Gut. 2013; 62:1024–1033. [PubMed: 22859495]
13. Leary RJ, Kinde I, Diehl F, Schmidt K, Clouser C, Duncan C, et al. Development of personalized tumor biomarkers using massively parallel sequencing. Sci Transl Med. 2010; 2:20ra14.
14. Leary RJ, Sausen M, Kinde I, Papadopoulos N, Carpten JD, Craig D, et al. Detection of chromosomal alterations in the circulation of cancer patients with whole-genome sequencing. Sci Transl Med. 2012; 4:162ra54.
15. Hofreiter M, Jaenicke V, Serre D, von Haeseler A, Paabo S. DNA sequences from multiple amplifications reveal artifacts induced by cytosine deamination in ancient DNA. Nucleic Acids Res. 2001; 29:4793–4799. [PubMed: 11726688]
16. Do H, Dobrovic A. Dramatic reduction of sequence artefacts from DNA isolated from formalin-fixed cancer biopsies by treatment with uracil-DNA glycosylase. Oncotarget. 2012; 3:546–558. [PubMed: 22643842]
17. Williams C, Ponten F, Moberg C, Soderkvist P, Uhlen M, Ponten J, et al. A high frequency of sequence alterations is due to formalin fixation of archival specimens. Am J Pathol. 1999; 155:1467–1471. [PubMed: 10550302]
18. Lin MT, Mosier S, Cope L, Thiess M, Beierl K, Chen G, et al. Clinical Validation of Kras, Braf, and EGFR Mutation Detection using Next Generation Sequencing. Accepted, pending revision.
19. Ehrlich M, Norris KF, Wang RY, Kuo KC, Gehrke CW. DNA cytosine methylation and heat-induced deamination. Biosci Rep. 1986; 6:387–393. [PubMed: 3527293]
20. Tsiatis AC, Norris-Kirby A, Rich RG, Hafez MJ, Gocke CD, Eshleman JR, et al. Comparison of Sanger sequencing, pyrosequencing, and melting curve analysis for the detection of KRAS mutations: diagnostic and clinical implications. J Mol Diagn. 2010; 12:425–432. [PubMed: 20431034]
21. Lin MT, Tseng LH, Rich RG, Hafez MJ, Harada S, Murphy KM, et al. Delta-PCR, A Simple Method to Detect Translocations and Insertion/Deletion Mutations. J Mol Diagn. 2011; 13:85–92. [PubMed: 21227398]
22. Do H, Wong SQ, Li J, Dobrovic A. Reducing sequence artifacts in amplicon-based massively parallel sequencing of formalin-fixed paraffin-embedded DNA by enzymatic depletion of uracil-containing templates. Clin Chem. 2013; 59:1376–1383. [PubMed: 23649127]
23. Yonekura S, Nakamura N, Yonei S, Zhang-Akiyama QM. Generation, biological consequences and repair mechanisms of cytosine deamination in DNA. J Radiat Res. 2009; 50:19–26. [PubMed: 18987436]
24. Duncan BK, Miller JH. Mutagenic deamination of cytosine residues in DNA. Nature. 1980; 287:560–561. [PubMed: 6999365]
25. Bjelland S, Seeberg E. Mutagenicity, toxicity and repair of DNA base damage induced by oxidation. Mutat Res. 2003; 531:37–80. [PubMed: 14637246]
26. Briggs AW, Stenzel U, Meyer M, Krause J, Kircher M, Paabo S. Removal of deaminated cytosines and detection of in vivo methylation in ancient DNA. Nucleic Acids Res. 2010; 38:e87. [PubMed: 20028723]
27. Sandigursky M, Franklin WA. Uracil-DNA glycosylase in the extreme thermophile Archaeoglobus fulgidus. J Biol Chem. 2000; 275:19146–19149. [PubMed: 10777501]

28. Sartori AA, Fitz-Gibbon S, Yang H, Miller JH, Jiricny J. A novel uracil-DNA glycosylase with broad substrate specificity and an unusual active site. EMBO J. 2002; 21:3182–3191. [PubMed: 12065430]

Key Points

- The baseline noise in normal peripheral blood and FFPE samples detected by next-generation sequencing is dominated by CG>TA mutations, a signature mutation of cytosine deamination.
- Consistent with this, treating samples with an enzyme designed to remove uracil, reduced these mutations, suggesting that one source is biologic. We also demonstrated that the heat of thermocycling (in the absence of polymerase) can increase these mutations.
- We conclude that the major source of baseline noise in next generation sequencing is both biologic and lab-induced cytosine deamination.

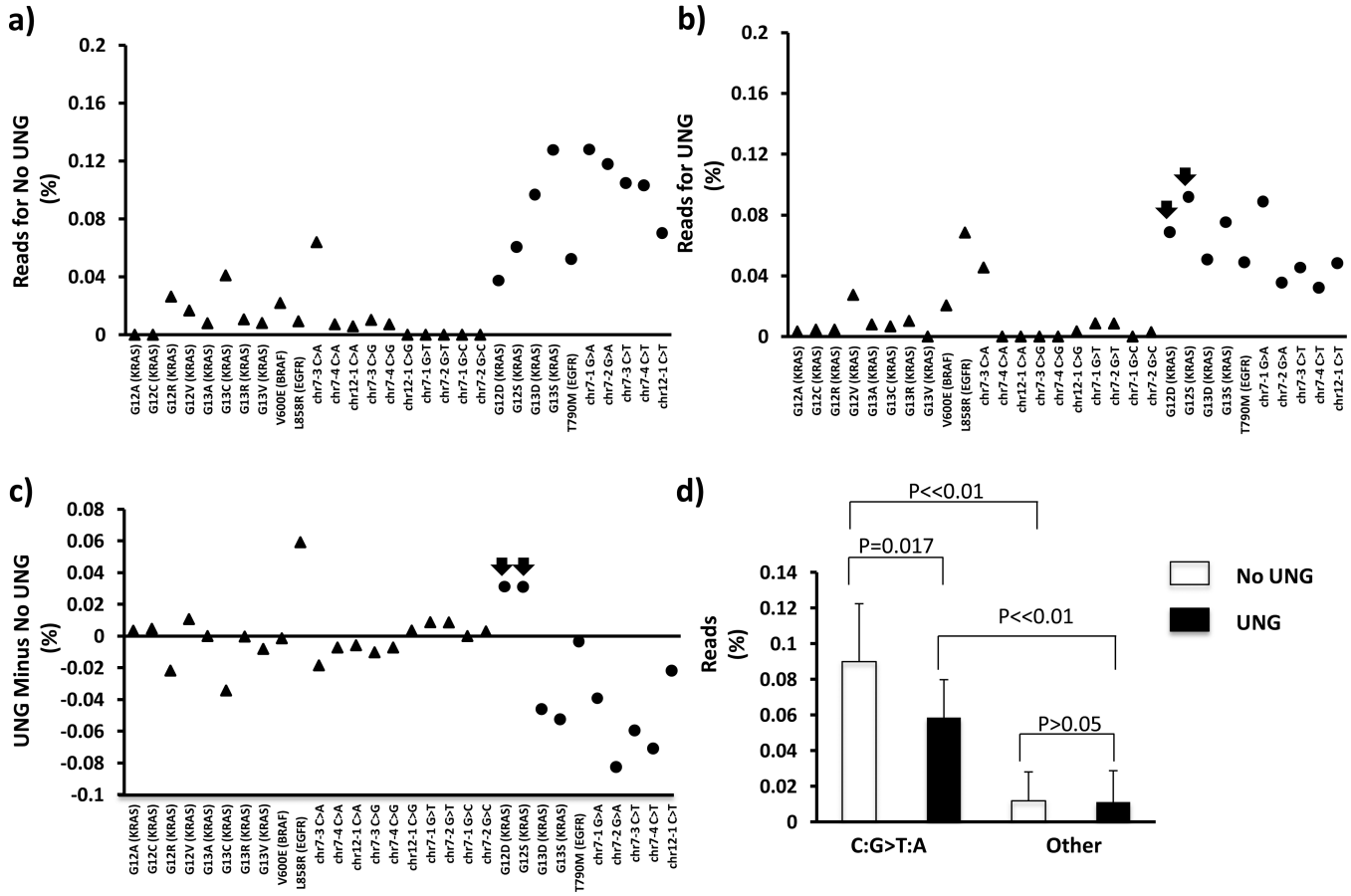

Fig. 1. UNG reduces the C:G >T:A mutations

Background noise of NGS for each of the common mutations within *KRAS* (codons 12 and 13), *BRAF* (V600E) and *EGFR* (T790M and L858R) genes, and five additional C:G sites in amplified regions of chromosomes 7 and 12 without (A) or with (B) UNG treatment prior to PCR amplification was evaluated on normal peripheral blood samples from four healthy subjects. The C:G→T:A mutations are labeled by filled circles (right), whereas other potential mutations are designated with triangles (left). The difference between two groups is demonstrated by subtraction of "No UNG" from "UNG" values (C). Note that most C:G->T:A mutations are reduced with UNG treatment except for two sites (arrows). The reason for the lack of effect on the latter two sites is unknown. The C:G>T:A mutations and other baseline noise (diamonds) are averaged, respectively (D). The C:G>T:A mutation level is significantly higher than that of other baseline noise (D) ($p<0.01$, student t-test); UNG treatment prior to PCR significantly decreases the C:G>T:A mutation level ($p<0.05$, paired t-test), without affecting other mutations ($p>0.05$, paired t-test). Following UNG treatment, the C:G>T:A mutation level is still significantly higher than that of other baseline noise ($p<0.01$, student t-test). The error bar represents standard deviation.

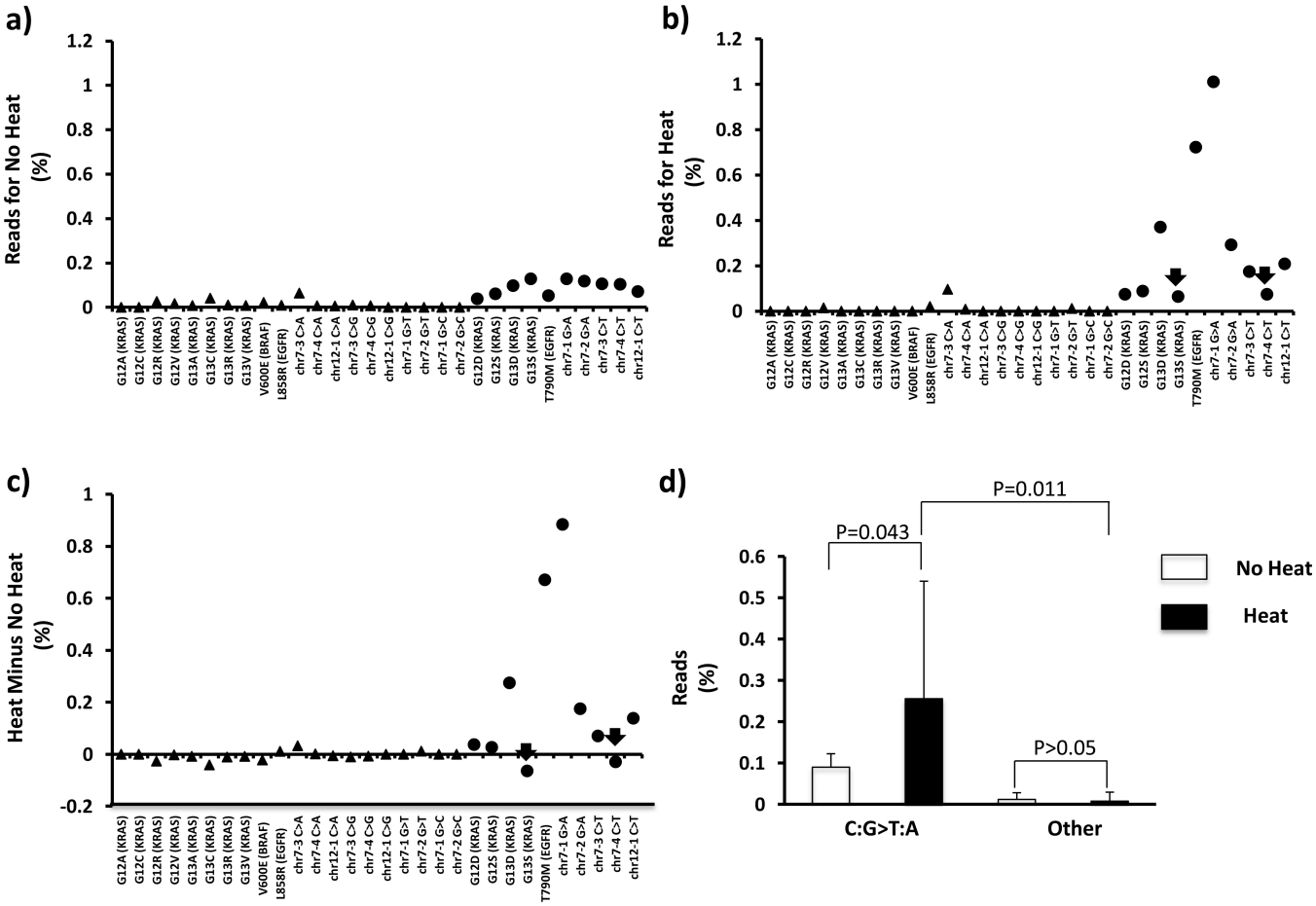

Fig. 2. Thermocyling induces deamination

Background noise of NGS for each of the common mutations within *KRAS* (codons 12 and 13), *BRAF* (V600E) and *EGFR* (T790M and L858R) genes, and five additional C:G sites in amplified regions of chromosomes 7 and 12 without (A) or with (B) prolonged heating treatment prior to library preparation was evaluated on peripheral blood samples from four health subjects. The difference between two groups is demonstrated by subtraction of "No Heat" from "Heat" values (C). While most values are above zero, there is clearly some variability among positions in the response to heat. The C:G→T:A mutations are labeled by filled circles (right) and induced with heat treatment except for two sites (arrows). Control mutations are represented by filled triangles (left). The C:G>T:A mutations and other baseline noise are averaged, respectively (D). Prolonged heat significantly increases the C:G>T:A mutation level ($p<0.05$, paired t-test), without affecting the frequency of other mutations ($p>0.05$, paired t-test). The error bar represents standard deviation.

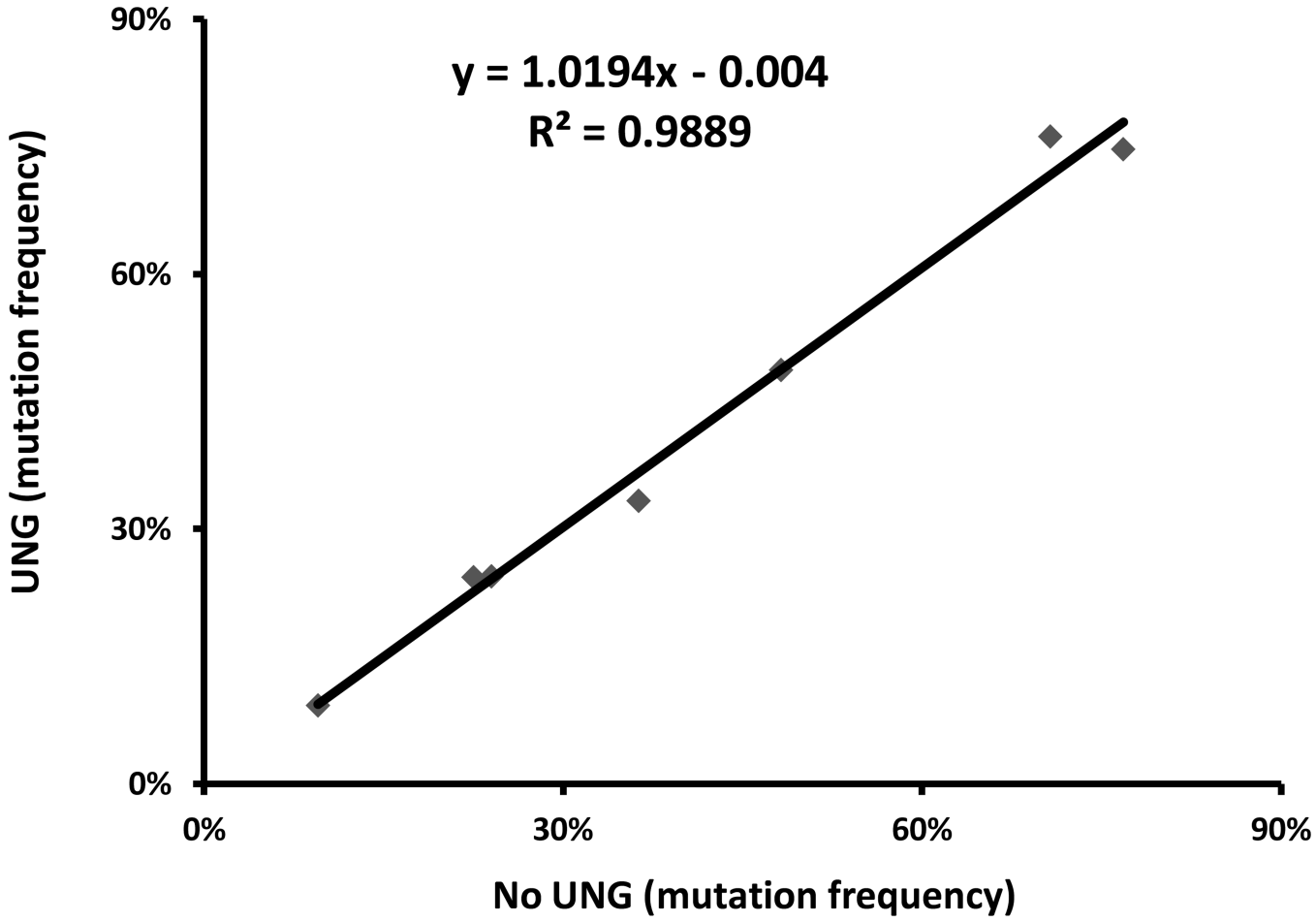

Fig. 3. UNG does not interfere with detection of real mutations by NGS

The mutation frequencies detected in UNG treated positive control samples (y-axis) are compared to those in untreated ones (x-axis).

Table 1

Mutation frequencies with and without UNG treatment

Sample	Gene	Mutation. cDNA	Mutation. Protein	Mutation Freq	Mutation Freq (UNG)
1	*KRAS*	c.35G>A	p.G12D	9.50%	9.23%
2	*KRAS*	c.35G>A	p.G12D	22.50%	24.31%
3	*KRAS*	c.35G>C	p.G12A	24%	24.37%
4	*BRAF*	c.1799T>A	p.V600E	36.30%	33.33%
5	*EGFR*	c.2573T>G	p.L858R	48.20%	48.71%
6	*EGFR*	c.2310_2311insAACCCCCAC	p.N771_H773dup	70.70%	76.16%
7	*EGFR*	c.2236_2249del15	p.E746_A750del	76.80%	74.72%

Mol Diagn Ther. Author manuscript; available in PMC 2015 October 01.