ACS central science

Smoking-Related DNA Alkylation Events Are Mapped at Single-Nucleotide Resolution

Cite This: *ACS Cent. Sci.* 2023, 9, 346−348

Alvin T. Huang and Weixin Tang*

Genome-wide profiling of DNA alkylation events recapitulates mutational signatures found in smoking-associated lung cancers.

Exposure to environmental stressors can induce DNA damage, resulting in alterations to the genetic code. The accumulation of certain mutations, particularly in genes relevant to cell growth and division, can promote the development of various diseases including cancer.[1] Some of the most well-studied carcinogens originate from cigarettes.[2] One such chemical is benzo[a]-pyrene (BaP), a polycyclic aromatic compound that is metabolized by cytochrome enzymes to form the incredibly reactive BaP-diol-epoxide (BPDE).[3,4] BPDE irreversibly reacts with guanine to form a variety of ring-opened products, the most common being N^2-trans-(+)-anti-BPDE-deoxyguanosine (N^2-BPDE-dG) (Figure 1a).[5] The adduct can be tracelessly removed through nucleotide excision repair, but those evading repair often lead to C:G to A:T mutations.[6] Although the mechanism of damage is well understood, the exact locations of N^2-BPDE-dG have not been mapped in a cellular context. In this issue of *ACS Central Science*, Sturla and co-workers report the first single-nucleotide resolution map of N^2-BPDE-dG in the human genome which they use to elucidate the relationship between adduct formation and mutational signatures, especially those found in smoking-related lung cancers.[7]

The authors devised separate workflows to quantify the prevalence of N^2-BPDE-dG and characterize its distribution. Quantification was done using liquid chromatography tandem mass spectrometry on cellular and naked genomic DNA of human bronchial epithelial cells (BEAS-2B) exposed to BPDE (Figure 1b). The N^2-BPDE-dG adduct was identified in both samples, with naked DNA showing a higher abundance. Overall adduct levels increased with higher concentrations of BPDE. The behavioral discrepancy between naked and cellular DNA presumably involves cellular machinery such as histone, transcription, and repair proteins.

To reveal locations of N^2-BPDE-dG, the authors developed BPDE-dG-Damage-seq (Figure 1c). Briefly, genomic DNA is fragmented followed by immunoprecipitation. The modified DNA is subsequently replicated with DNA polymerase that stalls at the location of the adduct, leading to shorter fragments that end immediately before N^2-BPDE-dG. These fragments are then barcoded, sequenced, and aligned to the human genome to determine the location of damage. The high frequency of G at the called damage site across various BPDE concentrations, along with the lack of nucleotide enrichment at adjacent sites, supports the fidelity of BPDE-dG-Damage-seq.

The high frequency of G at the called damage site across various BPDE concentrations, along with the lack of nucleotide enrichment at adjacent sites, supports the fidelity of BPDE-dG-Damage-seq.

In line with the mass spectrometry data, the level of damage in cellular DNA was almost always lower than that of naked DNA. The impact of cellular processes on damage accumulation was inferred from the relative abundance of

Published: March 7, 2023

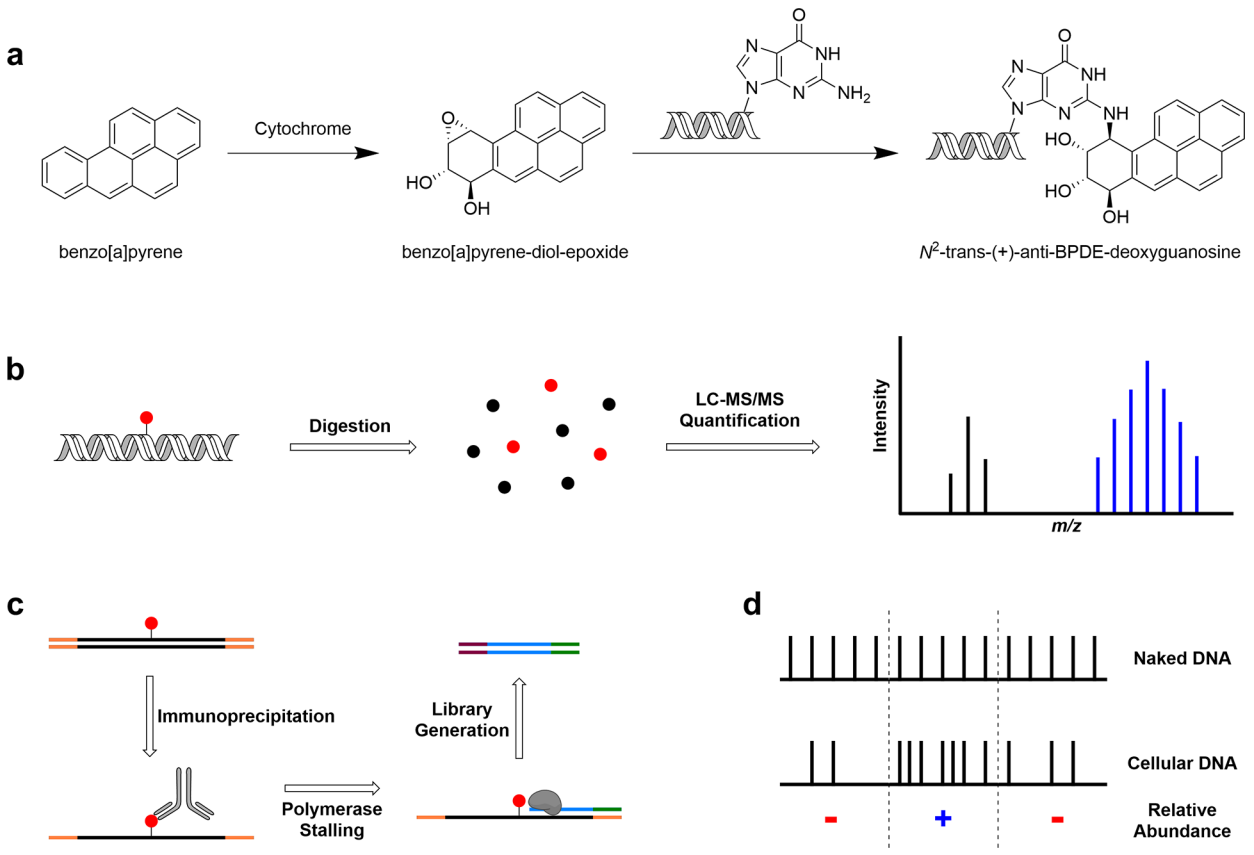

Figure 1. (a) BaP is metabolized by cytochrome enzymes to form the highly reactive BaP-diol-epoxide. A nucleophilic attack by the deoxyguanosine amine opens the epoxide ring forming N^2-trans-(+)-anti-BPDE-deoxyguanosine in DNA. (b) To quantify the amount of adduct, genomic DNA is extracted, digested, and analyzed by mass spectrometry. (c) To analyze the adduct distribution, immunoprecipitation is performed on fragmented genomic DNA to separate the modified DNA. A DNA polymerase stalling assay is then performed to create shortened fragments that end immediately in front of the adduct. These short amplicons are barcoded, sequenced, and aligned to the genome to locate modified sites. (d) Sequencing data from naked and cellular DNA are overlaid, and the difference in abundance of adduct is noted. Regions modified to lower/higher levels in cellular DNA compared to naked DNA are marked with a red "−" and a blue "+", respectively.

N^2-BPDE-dG in cellular and naked DNA (Figure 1d). Overlaying with a chromatin accessibility map revealed that open regions had fewer adducts compared to closed regions. This observation was largely consistent with the methylation state of the DNA. On the sequence level, highly transcribed genes carried less damage; the transcribed strand had less damage compared to the nontranscribed strand. These observations point to differences in how DNA is repaired: open chromatin is more accessible to the repair machinery compared to closed chromatin; the template strand may be preferentially repaired by transcription-coupled nucleotide excision repair.[8]

With this information in hand, the authors compared their genomic map with mutations identified in smoking-related lung cancers. The relative frequencies of N^2-BPDE-dG-bearing trinucleotide sequences observed by the authors were remarkably similar to those of mutational signatures found in cancer patients.[2] N^2-BPDE-dG was detected consistently in lung cancer hot-spot genes, including the chromatin modifying gene SMARCA4 and the tumor suppressor genes CDKN2A, KEAP1, and TP53. In particular, BPDE damage found in CDKN2A matches the most frequently mutated site of this gene in lung adenocarcinoma, which is also one of the top ranked genome-wide mutations observed in lung carcinomas. Almost all lung-carcinoma patients with mutations at this site are smokers, indicating a strong correlation between BPDE-mediated formation of DNA lesions and the progression of lung cancer.

Almost all lung-carcinoma patients with mutations at this site are smokers, indicating a strong correlation between BPDE-mediated formation of DNA lesions and the progression of lung cancer.

Collectively, Sturla and co-workers construct a genome-wide map for N^2-BPDE-dG that links carcinogen-DNA adducts to mutational signatures seen in medical cases. The study also provides a general method to genetically map

https://doi.org/10.1021/acscentsci.3c00246
ACS Cent. Sci. 2023, 9, 346−348

DNA alkylation events and predict cancer mutations. A key insight of the study is that genomic features shape the formation and retention of alkylation products. DNase data from BEAS-2B cells, instead of a generic DNase map of the human genome as used by the present study, would enable a more specific analysis.

The fact that open DNA carries fewer adducts raises an intriguing paradox: unlike nucleosome-packed DNA, open DNA is less protected from damage, but is also more actively repaired. Results of the current study suggest that it is the repair process, rather than reactivity, that dominates the accumulation of alkylation products in the genome. It would be interesting to determine if such a trend persists with shorter treatment time or in different cell types. Many of the observed phenomena can be mechanistically dissected through time-resolved profiling and by comparing damage profiles in cells of active and disabled repair machinery. Ultimately, these studies will provide important insights into the relationship between DNA alkylation, damage accumulation, and cancer progression.

Author Information

Corresponding Author

Weixin Tang − *Department of Chemistry, The University of Chicago, Chicago, Illinois 60637, United States; Institute for Biophysical Dynamics, The University of Chicago, Chicago, Illinois 60637, United States;* orcid.org/0000-0002-5739-5416; Email: weixintang@uchicago.edu

Author

Alvin T. Huang − *Department of Chemistry, The University of Chicago, Chicago, Illinois 60637, United States; Institute for Biophysical Dynamics, The University of Chicago, Chicago, Illinois 60637, United States*

Complete contact information is available at: https://pubs.acs.org/10.1021/acscentsci.3c00246

Notes

The authors declare no competing financial interest.

REFERENCES

(1) Tubbs, A.; Nussenzweig, A. Endogenous DNA Damage as a Source of Genomic Instability in Cancer. *Cell* **2017**, *168* (4), 644−656.

(2) Alexandrov, L. B.; Ju, Y. S.; Haase, K.; Van Loo, P.; Martincorena, I.; Nik-Zainal, S.; Totoki, Y.; Fujimoto, A.; Nakagawa, H.; Shibata, T.; Campbell, P. J.; Vineis, P.; Phillips, D. H.; Stratton, M. R. Mutational signatures associated with tobacco smoking in human cancer. *Science* **2016**, *354* (6312), 618−622.

(3) Sims, P.; Grover, P. L.; Swaisland, A.; Pal, K.; Hewer, A. Metabolic activation of benzo(a)pyrene proceeds by a diol-epoxide. *Nature* **1974**, *252* (5481), 326−328.

(4) Wood, A. W.; Levin, W.; Lu, A. Y.; Yagi, H.; Hernandez, O.; Jerina, D. M.; Conney, A. H. Metabolism of benzo(a)pyrene and benzo (a)pyrene derivatives to mutagenic products by highly purified hepatic microsomal enzymes. *J. Biol. Chem.* **1976**, *251* (16), 4882−489.

(5) Jerina, D. M.; Chadha, A.; Cheh, A. M.; Schurdak, M. E.; Wood, A. W.; Sayer, J. M. Covalent Bonding of Bay-Region Diol Epoxides to Nucleic Acids. *Adv. Exp. Med. Biol.* **1991**, *283*, 533−553.

(6) Seeberg, E.; Steinum, A. L.; Nordenskjöld, M.; Söderhäll, S.; Jernström, B. Strand-break formation in DNA modified by benzo-[a]pyrene diolepoxide Quantitative cleavage by *Escherichia coli uvr ABC* endonuclease. *Mutat. Res.* **1983**, *112* (3), 139−145.

(7) Jiang, Y.; Mingard, C.; Huber, S. M.; Takhaveev, V.; McKeague, M.; Kizaki, S.; Schneider, M.; Ziegler, N.; Huerlimann, V.; Hoeng, J.; Sierro, N.; Ivanov, N. V.; Sturla, S. Quantification and mapping of alkylation in the human genome reveal single nucleotide resolution precursors of mutational signatures. *ACS Cent. Sci.* **2023**, DOI: 10.1021/acscentsci.2c01100.

(8) Perlow, R. A.; Kolbanovskii, A.; Hingerty, B. E.; Geacintov, N. E.; Broyde, S.; Scicchitano, D. A. DNA Adducts from a Tumorigenic Metabolite of Benzo[a]pyrene Block Human RNA Polymerase II Elongation in a Sequence- and Stereochemistry-Dependent Manner. *J. Mol. Biol.* **2002**, *321* (1), 29−47.