LETTER

doi:10.1038/nature24041

Dynamic landscape and regulation of RNA editing in mammals

Meng How Tan[1,2,3]*, Qin Li[1]*, Raghuvaran Shanmugam[2,3], Robert Piskol[1], Jennefer Kohler[1], Amy N. Young[1], Kaiwen Ivy Liu[3], Rui Zhang[1], Gokul Ramaswami[1], Kentaro Ariyoshi[4], Ankita Gupte[5], Liam P. Keegan[6,7], Cyril X. George[8], Avinash Ramu[9,10], Ni Huang[9,10], Elizabeth A. Pollina[1], Dena S. Leeman[1], Alessandra Rustighi[11], Y. P. Sharon Goh[12], GTEx Consortium†, Ajay Chawla[12], Giannino Del Sal[11], Gary Peltz[13], Anne Brunet[1], Donald F. Conrad[9,10], Charles E. Samuel[8], Mary A. O'Connell[6,7], Carl R. Walkley[5,14], Kazuko Nishikura[4] & Jin Billy Li[1]

Adenosine-to-inosine (A-to-I) RNA editing is a conserved post-transcriptional mechanism mediated by ADAR enzymes that diversifies the transcriptome by altering selected nucleotides in RNA molecules[1]. Although many editing sites have recently been discovered[2–7], the extent to which most sites are edited and how the editing is regulated in different biological contexts are not fully understood[8–10]. Here we report dynamic spatiotemporal patterns and new regulators of RNA editing, discovered through an extensive profiling of A-to-I RNA editing in 8,551 human samples (representing 53 body sites from 552 individuals) from the Genotype-Tissue Expression (GTEx) project and in hundreds of other primate and mouse samples. We show that editing levels in non-repetitive coding regions vary more between tissues than editing levels in repetitive regions. Globally, ADAR1 is the primary editor of repetitive sites and ADAR2 is the primary editor of non-repetitive coding sites, whereas the catalytically inactive ADAR3 predominantly acts as an inhibitor of editing. Cross-species analysis of RNA editing in several tissues revealed that species, rather than tissue type, is the primary determinant of editing levels, suggesting stronger *cis*-directed regulation of RNA editing for most sites, although the small set of conserved coding sites is under stronger *trans*-regulation. In addition, we curated an extensive set of ADAR1 and ADAR2 targets and showed that many editing sites display distinct tissue-specific regulation by the ADAR enzymes *in vivo*. Further analysis of the GTEx data revealed several potential regulators of editing, such as AIMP2, which reduces editing in muscles by enhancing the degradation of the ADAR proteins. Collectively, our work provides insights into the complex *cis*- and *trans*-regulation of A-to-I editing.

The prevalence and importance of A-to-I RNA editing have been illuminated in recent years largely owing to the rapid adoption of high-throughput sequencing technologies[11,12]. Separate laboratories have examined the RNA editome across many tissues or developmental stages in human and other mammals[13–17]. However, the published studies are limited in the number of samples and tissues examined and do not systematically compare the editing landscape across species or thoroughly dissect the regulation of editing. In this work, we performed multidimensional analyses of thousands of new and publicly available sequencing libraries to address major gaps in our fundamental knowledge of A-to-I editing.

To construct a mammalian reference atlas of A-to-I editing, we first compiled a comprehensive list of editing sites in human and mouse (Supplementary Note 1) and then examined the RNA editome across tissues using 8,551 RNA-sequencing (RNA-seq) samples derived from 552 donors in the GTEx project (Supplementary Information 1). Notably, the editing profiles across different tissues were highly correlated (Fig. 1a) and the overall editing activities were also generally similar, except for skeletal muscle, in which editing was significantly lower than in other tissues ($P < 2.2 \times 10^{-16}$, Wilcoxon rank sum test; Fig. 1b). Nevertheless, principal component analysis (PCA) showed that the brain regions could still be resolved from non-brain tissues (Extended Data Fig. 1a). Within the brain, the cerebellum was clearly segregated from other brain parts (Extended Data Fig. 1b), possibly owing to higher expression of *ADAR2* (also known as *ADARB1*) (Extended Data Fig. 1c). When we examined non-repetitive sites in coding regions only, the editing levels became more distinct among the various tissues (Fig. 1a). The different brain regions clustered together, as did heart and skeletal muscle. Unexpectedly, the artery was the most highly edited tissue type (Fig. 1c). The importance of RNA editing in vascular disease was demonstrated in a recent study[18]. We further validated the results obtained from the GTEx data by applying a targeted sequencing approach (microfluidics-based multiplex PCR and deep sequencing; mmPCR–seq)[19] (Supplementary Note 2) to examine 12,871 exonic sites in 672 loci (Supplementary File 2) on independent tissue samples from two individuals (Extended Data Fig. 2).

The extent to which variation in editing may be attributed to the expression of each ADAR enzyme is not well understood. From the GTEx data, we found that the expression of *ADAR1* (also known as *ADAR*) accounted for approximately 20% of the variation in overall editing of repetitive sites (Fig. 1d), which represented 97.7% of all known editing sites. By contrast, *ADAR2* expression explained 2.8% of the variation (Fig. 1d). However, when non-repetitive protein-coding sites were considered instead, *ADAR1* expression accounted for only 6% of the variation, whereas *ADAR2* expression accounted for 25% (Fig. 1e). The expression of *ADAR3* (also known as *ADARB2*), which localizes exclusively to the brain and has no enzymatic activity, was negatively correlated with editing levels in brain (Fig. 1f). When the negative influence of ADAR3 was taken into account, ADAR1 and ADAR2 were able to explain better the variation in editing (Fig. 1g), supporting the hypothesis that ADAR3 served predominantly as an

[1]Department of Genetics, Stanford University School of Medicine, Stanford, California 94305, USA. [2]School of Chemical and Biomedical Engineering, Nanyang Technological University, Singapore 637459, Singapore. [3]Genome Institute of Singapore, Agency for Science Technology and Research, Singapore 138672, Singapore. [4]The Wistar Institute, Philadelphia, Pennsylvania 19104, USA. [5]St Vincent's Institute of Medical Research, Fitzroy, Victoria 3065, Australia. [6]MRC Human Genetics Unit, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh EH4 2XU, UK. [7]Central European Institute of Technology, Masaryk University, Kamenice, Brno 625 00, Czech Republic. [8]Department of Molecular, Cellular and Developmental Biology, University of California-Santa Barbara, Santa Barbara, California 93106, USA. [9]Department of Genetics, Washington University School of Medicine, St Louis, Missouri 63108, USA. [10]Department of Pathology & Immunology, Washington University School of Medicine, St Louis, Missouri 63108, USA. [11]Department of Life Sciences, University of Trieste, 34127 Trieste, Italy and National Laboratory CIB (LNCIB), Area Science Park, 34149 Trieste, Italy. [12]Cardiovascular Research Institute, University of California-San Francisco, San Francisco, California 94158, USA. [13]Department of Anesthesia, Stanford University School of Medicine, Stanford, California 94305, USA. [14]Department of Medicine, St. Vincent's Hospital, University of Melbourne, Fitzroy, Victoria 3065, Australia.
*These authors contributed equally to this work.
†Lists of participants and their affiliations appear in the online version of the paper.