|  |  |
| --- | --- |
| I. Project Title and pCode | |
| 201710261504-Christiani: Epigenome-wide survival analysis based on integrating meQTL for non-small cell lung cancer202002111052-Christiani: The association of genetically predicted blood protein biomarkers and lung cancer survival202002111102-Christiani: Fine mapping of MHC region in lung cancer prognosis202002121034-Christiani: Genetic landscape analysis of short tandem repeats on the risk and survival of lung cancer | **Date:**  Apr 19, 2023 |
| **II. Project Leaders (List Names and Institutes)** | **Primary Contact**  **Information** |
| Project leaders:   * David C. Christiani, Harvard T.H Chan School of Public Health * Mulong Du, Harvard T.H Chan School of Public Health   Investigators:   * ILLCO contributing study/ data PI’s | **Name:**  David C. Christiani  **Email:**  dchris@hsph.harvard.edu  **Tel:**  6174323323 |
| III. Progress to date. *What has been done so far to move the project forward (data pooling, grant submission, etc), and when it was done* | |
| Up to Feb, 2023, we have collected several survival GWAS datasets of lung cancer patients, and performed genome-wide imputation and SNP-level/individual-level quality control for each cohort, including:   * ILCCO (BLCS part): 4503 patients; about 5 millions SNPs * TCGA: 901 patients; about 5 millions SNPs * PLCO: 1185 patients; about 5 millions SNPs * UK Biobank: 1700 patients; about 5 millions SNPs | |
| **IV. Project Status and Preliminary Results**. *S****ummarize the current status of the project. Report preliminary results if there is any*** | |
| 1. we have assessed the association between genome-wide SNPs and lung cancer survival in each cohort, and released the survival-based GWAS summary statistics of UK Biobank (SUMMER database; *Nucleic Acids Res. 2023 Jan 6;51(D1):D1160-D1167*). The above survival datasets can be applied in our projects for independent validation. 2. we have proposed a solid analysis framework about Mendelian randomization (*BMC Med. 2022 May 11;20(1):168*), which can be applied in the multi-omics projects, such as the integration of GWAS with CpG and protein biomarkers. 3. for some datasets (e.g., TCGA), we have completed the imputation of MHC region and short tandem repeats (STRs). | |
| **V. Justification of Extension** | |
| Considering the update and release of some datasets (e.g., UK Biobank), we need to add more analyses to improve and validate our results. | |
| **VI. Future Plans & Timeline** | |
| We need 6 months to update the results, and extral 6 months to write the manuscripts. These projects will be finished by end of Apr, 2024. | |