Mosaic Loss of Chromosome Y in Blood Is Associated with Alzheimer Disease

Jan P. Dumanski, Jean-Charles Lambert, Chiara Rasi, Vilmantas Giedraitis, Hanna Davies, Benjamin Grenier-Boley, Cecilia M. Lindgren, Dominique Campion, Carole Dufouil, The European Alzheimer’s Disease Initiative Investigators, Florence Pasquier, Philippe Amouyel, Lars Lannfelt, Martin Ingelsson, Lena Kilander, Lars Lind, and Lars A. Forsberg

Men have a shorter life expectancy compared with women but the underlying factor(s) are not clear. Late-onset, sporadic Alzheimer disease (AD) is a common and lethal neurodegenerative disorder and many germline inherited variants have been found to influence the risk of developing AD. Our previous results show that a fundamentally different genetic variant, i.e., lifetime-acquired loss of chromosome Y (LOY) in blood cells, is associated with all-cause mortality and an increased risk of non-hematological tumors and that LOY could be induced by tobacco smoking. We tested here a hypothesis that men with LOY are more susceptible to AD and show that LOY is associated with AD in three independent studies of different types. In a case-control study, males with AD diagnosis had higher degree of LOY mosaicism (adjusted odds ratio = 2.80, p = 0.0184, AD events = 606). Furthermore, in two prospective studies, men with LOY at blood sampling had greater risk for incident AD diagnosis during follow-up time (hazard ratio [HR] = 6.80, 95% confidence interval [95% CI] = 2.16–21.43, AD events = 140, p = 0.0011). Thus, LOY in blood is associated with risks of both AD and cancer, suggesting a role of LOY in blood cells on disease processes in other tissues, possibly via defective immunosurveillance. As a male-specific risk factor, LOY might explain why males on average live shorter lives than females.

Introduction

Alzheimer disease (AD [MIM: 104300]) is the most common neurodegenerative disorder and constitutes a major public health problem worldwide. Its etiology is complex and several pathways probably contribute to the pathology. The delineation of each pathophysiological pathway and identification of the factors that modulate the clinical phenotypes are crucial for the development of effective treatments and improved definitions of at-risk groups. The identification of genes involved in early-onset monogenic forms of AD has significantly contributed to our knowledge of the disease mechanisms. The causal links between mutations, the functions of the mutated genes, and disease development prompted a hypothesis radically changing understanding of AD; i.e., the amyloid cascade hypothesis. However, most cases of AD present a sporadic and late-onset form of disease. Thus, we need to further characterize the etiology of AD and understanding the genetics of the disease appears to be one of the best ways forward, as has been the case for monogenic forms of AD. Indeed, it has been estimated that genetic risk factors account for up to 80% of the attributable risk for AD and one can thus argue that the majority of the AD pathophysiology is driven by or include genetic factors. In addition to APOE (MIM: 107741), which is the major genetic determinant of AD, the advent of genomic approaches (e.g., SNP-based genome-wide association studies [GWASs] and next-generation sequencing) has now led to the characterization of more than 25 genetic risk variants. These implicate pathways related to immune response, regulation of endocytosis, cholesterol transport, and protein ubiquitination. However, less than 50% of the genetic risk for AD has been characterized so far. It is also noteworthy that although SNPs have been extensively studied in AD, only a restricted number of reports have assessed the association of structural genetic variations with AD risk.

The field of human genetics is increasingly recognizing that genetic variation acquired during life (i.e., post-zygotic changes) is not sufficiently explored. This is valid for both the extent of detectable post-zygotic variants at various ages and in different tissues as well as their importance for variety of human phenotypes and diseases. The major insight over the past few years has been discoveries of frequent aberrant clonal expansions (ACEs) in peripheral blood cells from apparently disease-free subjects. It has now been thoroughly described that ACEs often carry various post-zygotic genetic alterations, the majority of which affect gene dosage of chromosomal segments and/or mutations in specific genes that are connected with cancer development. Thus, ACEs represent potential pre-cancerous changes and a clear association of ACEs with increasing age in humans has also been shown.
In agreement with the above, reports have shown strong associations of ACEs with cancer mortality and cancer diagnoses, both non-hematological and hematological. Furthermore, associations between ACEs and non-cancer related outcomes, such as type 2 diabetes and cardiovascular disease, have also been shown. This might suggest the importance of ACEs for conditions outside the cancer field and further emphasize possible effects of ACEs detectable in blood on disease processes in other tissues, i.e., outside the hematopoietic system.

Chromosome Y is recognized for its role in sex determination and normal sperm production, but it has long been considered as genetic wasteland and its characterization has lagged behind that of the rest of the genome. Furthermore, it has been known for more than 50 years, from the earliest cytogenetic analyses, that mosaic loss of chromosome Y (LOY) is frequent among aging men. However, the phenotypic consequences of LOY have been elusive and the prevailing consensus has been that it should be considered phenotypically neutral and related to normal aging. The recent molecular analyses suggest that mosaic LOY in normal blood samples can be detected in $\geq 10\%$ of blood cells in at least $15\%$ of males around and above 70 years of age. Thus, normal blood cells with nullisomy Y represent an ACE, which is detectable because cell clones without chromosome Y are enriched, possibly due to an increased proliferative potential. Moreover, other analyses of various cancers indicate that chromosome Y is lost in many types of tumors in frequencies ranging from $15\%$ to $80\%$ of cases. Thus, combined results on LOY in non-cancerous blood clones and in transformed tumor cells suggest that nullisomy Y is the most common human mutation, and it affects $\sim 1.6\%$ of the genome.

Recent analyses using two independent cohorts (ULSAM and PIVUS) showed an association between LOY in blood and risk for all-cause mortality as well as for cancer outside the hematopoietic system. Both of these cohorts are also included in the current study. The median survival time among men with LOY was 5.5 years shorter and about half as long when compared to those without LOY, which suggests that LOY in blood could become a predictive biomarker of male carcinogenesis. It was further demonstrated that smoking is associated with LOY in blood cells in three independent cohorts. The finding that smoking can induce LOY thus links a very common and preventable environmental risk factor with the most common human mutation. The above-mentioned results regarding LOY and cancer risk have recently been extended to several specific cancer diagnoses. It should also be mentioned that LOY has been suggested to be involved in the pathogenesis of rare autoimmune diseases in males, such as autoimmune thyroiditis and primary biliary cirrhosis, which points to a possible functional connection between LOY and dysfunction of the immune system. Here we tested the hypothesis that aging males with mosaic LOY are also more susceptible to develop other common diseases, concentrating on Alzheimer disease, using three independent studies of different types: two prospective investigations and a case-control study.

Materials and Methods

Studied Cohorts

The EADI1 case-control study (European Alzheimer’s Disease Initiative stage 1) was collected specifically for research on AD and first described in Lambert et al. Controls for this investigation were selected from the 3C Study, which is a population-based study of the relationship between vascular factors and dementia. It has been carried out in three French cities: Bordeaux, Montpellier, and Dijon. A sample of non-institutionalized, 65 years old or older subjects was randomly selected from the electoral rolls of each city. The second cohort was Uppsala Longitudinal Study of Adult Men (ULSAM), which started in 1970–1974, when all 50-year-old men living in Uppsala County, Sweden, were invited to a health survey, initially focusing at identifying risk factors for cardiovascular disease. Out of a total of 2,841 men born in 1920–1924, 2,322 (82%) agreed to participate in the study. Since then, the cohort has been reinvestigated several times at ages 60, 70 (when the first blood samples for DNA extraction were collected), 77, 82, 88, 91, and 93 years. The third cohort was PIVUS (Prospective Investigation of the Vasculature in Uppsala Seniors), which started in 2001 with the primary aim of investigating the predictive power of various measurements of endothelial function and arterial compliance. Eligible participants were all aged 70 and were living in the community of Uppsala, Sweden. The subjects were randomly chosen from the community register, and 1,016 men and women participated. Two reinvestigations of the cohort were undertaken, starting in the spring of 2006 and in the spring of 2011 at the ages of 75 and 80 years, respectively. The number of subjects included in statistical analyses from all three cohorts and the confounding factors at baseline are shown in Table S1. All procedures were in accordance with the standards of the responsible local research ethics committees on human experimentation and proper informed consents were obtained.

Assessment of Dementia Status

We included participants only from clinical analyses performed with the highest possible standards and confidence of AD diagnoses. In the EADI1 cohort, AD cases were ascertained by neurologists from Bordeaux, Dijon, Lille, Montpellier, Paris, and Rouen. In the ULSAM and PIVUS cohorts, the AD diagnoses were assigned by experienced geriatricians at the Memory Clinic, Uppsala University Hospital. In all cohorts, AD was diagnosed according to the National Institute of Neurological and Communicative Diseases and Stroke and the AD and Related Disorders Association criteria.

Estimation of LOY from SNP-Array Data

All participants were genotyped using different versions of Illumina SNP arrays and we applied a strict quality criterion as described in the Results. A continuous variable, the median of the log R ratio (i.e., mLRRY), was used to estimate the degree of LOY for each subject and calculated as the median value of the SNP-array probes positioned within the male-specific region of
Results

We present results of association between mosaic LOY in blood and risk of AD diagnosis from analyses performed in three independent studies. First, we separately describe results from analyses of four confounding factors such as quality of genotyping experiments, sampling age of participants, APOE genotype, and smoking, which are particularly important for our study.

Assessment of Genotyping Quality and Validation of LOY via WGS

A continuous estimate of the level of LOY in each participant (i.e., mLRRY) was calculated from the SNP probes within the MSY, as described in Materials and Methods previously. The quality of individual SNP-array experiments could be a confounding factor when estimating the level of LOY and therefore needs careful consideration. To minimize the potential bias from bad-quality data, we applied a stringent criterion for all experiments, as recommended by the SNP-array manufacturer (Figure S2). In brief, we calculated this quality metric from each experiment as the standard deviation of the log R ratio values of the SNP probes located on chromosome 1. A high value reflects poor genotyping quality whereas experiments with value below 0.28 are considered high quality (see link in Web Resources to Illumina Tech Note). In total, 3,218 experiments passed this quality control and were included in further analyses (Figure S3). Furthermore, validation of the SNP-array-based inferences of LOY was achieved by analysis of next-generation whole-genome sequencing data (WGS). This was performed by estimating ploidy values from WGS data using the FREEC software. Using this approach in 100 ULSAM participants, all observations of LOY as well as other autosomal mosaic copy-number variants detected by SNP array could be validated using the low-coverage (~5x) WGS data, as described previously. A similar strategy was used to validate LOY observations in SNP-array data in the EADI1 study using WGS data with higher coverage (~30x) from 183 participants among the 1,611 passing the genotyping quality assessment (Figure S3). As illustrated in Figure 1, the two independent technologies applied in EADI1 produced fully concordant estimates of individuals’ LOY status.

Association between LOY in Blood Cells and Age

Of the 3,218 participants included in our analyses, 546 (17.0%) had a detectable level of LOY mosaicism, i.e., an mLRRY value lower than the 99% confidence limit of the experimental variation. The frequency of LOY was similar in the ULSAM, PIVUS, and EADI1 cohorts with 17.5%, 21.1%, and 15.4% of participants, respectively. A linear regression model performed on this dataset showed that LOY was more common in older participants ($F_{(1,3216)} = 26.79$, $p < 0.0001$) (Figure 2A), and grouping participants based on age (Figures 2E and 2F).
Figure 1. Reproducible Estimations of LOY via Two Independent Technologies

Validation of LOY mosaicism detected by SNP-array data by whole-genome next-generation sequencing (WGS) in 183 EADI1 participants. This was performed by estimating ploidy values from about 30\(\times\) NGS data using the FREEC software.\textsuperscript{46}

further illustrates that the fraction of men affected with LOY is increasing with participant age. Hence, both the frequency of LOY and the incidence of AD are positively correlated with age. This covariation was handled by fitting the effect of age at baseline as a continuous explanatory variable in the statistical models evaluating the association between LOY in blood cells and AD. We also evaluated the influence of age in the investigation of LOY on AD risk by analyses in two subsets of men within narrow age ranges (see below).

LOY and APOE Genotype

The Apolipoprotein E (\textit{APOE} [MIM: 107741]) genotype and in particular homozygosity for the APOE epsilon 4 risk allele is a well-established risk factor for AD.\textsuperscript{4} An analysis of the level of LOY observed among individuals with different APOE genotypes (i.e., A22, A23, A24, A33, A34, and A44) in the three cohorts showed no significant differences (ANOVA; \(F_{(3,3017)} = 0.4760, p = 0.7950\)) (Figure S4A). We also tested association between LOY and APOE genotype after grouping subjects homozygous for the APOE epsilon 4 risk allele or not and found no significant differences (Kolmogorov-Smirnov test; \(D = 0.0393, p = 0.7836\)) (Figure S4B). These results suggest that degree of LOY mosaicism within the studied men is independent from their APOE genotype. However, because APOE genotype is a strong risk factor for AD that explains a substantial portion of the variation in AD, a binary variable reflecting the APOE epsilon 4 allele homozygosity state (1/0) was fitted as in subsequent statistical models evaluating association between LOY and AD.

LOY and Smoking

We have previously showed that LOY in blood cells is associated with smoking.\textsuperscript{31} Moreover, some studies suggest that smoking might play a role in dementia and Alzheimer disease.\textsuperscript{38,49} The present dataset included 1,565 subjects passing genotyping quality control and with smoking data available (n = 1,097 and n = 468 in ULSAM and PIVUS, respectively). However, in the current dataset there was no significant association between smoking status at 70 years of age and AD diagnosis (logistic regression; \(p = 0.2570\)). The lack of association between AD diagnosis and smoking status was replicated also after adjusting for the confounding effects of APOE genotype and sampling age (logistic regression; \(p = 0.2459\)). Nevertheless, the potential effect from smoking on AD status was fitted as a binary variable (current smoker at age 70 or not) in the subsequent models testing association between LOY and risk for AD. Furthermore, we also tested association between LOY and risk for AD in a subset including non-current smokers only (see below).

Association between LOY and AD in Case-Control Analyses

Results from the EADI1 case-control study including 1,611 subjects with high-quality genotyping data showed that 606 AD-affected subjects had a significantly higher degree of LOY in blood cells (i.e., lower mLRRY) compared to 1,005 control subjects in an unadjusted test of association (Kolmogorov-Smirnov test: \(D = 0.07, p = 0.0198\)) (Figure 3A). Furthermore, this association was significant also in models adjusting for the effects from the confounding factors APOE genotype and age at sampling (Figure 3B and Table S2A). The confounders analyzed are summarized in Table S1. A logistic regression model further showed, after adjusting for APOE genotype and age, that EADI1 men with a higher fraction of cells without chromosome Y were more likely to be diagnosed with AD (OR = 2.80, 95% CI = 1.19–6.61, \(p = 0.0184\)) (Table S2B). We also performed case-control analyses after pooling the data from the EADI1, ULSAM, and PIVUS cohorts (n = 3,218) with results comparable to the above (unadjusted Kolmogorov-Smirnov test: \(D = 0.05, p = 0.0390\)) (Figure 3C) and in a model adjusting for APOE genotype and age (ANCOVA: \(F_{(1,3085)} = 7.44, p = 0.0064\)) (Figure 3D). Furthermore, we also analyzed the ULSAM and PIVUS data as case-controls, both separately (Figures 3E and 3F) and pooled together (Figures 3G and 3H). In summary, comparisons of participants from three independent studies show that men with AD diagnosis had a higher degree of LOY mosaicism in blood compared to controls.

Association between LOY and Risk for AD in Two Prospective Cohorts

The association between LOY in blood cells and risk for AD diagnosis was also examined using Cox proportional hazards regression model performed with the R package Survival (Web Resources). These analyses were performed using a pooled dataset encompassing results from two prospective cohorts (i.e., ULSAM+PIVUS) and no prevalent cases of AD were included, i.e., cases diagnosed prior
to study baseline \((n = 8)\). Among the 1,599 participants included in these analyses, 140 incident cases of AD were observed during follow-up time (123 in ULSAM and 17 in PIVUS). Study entry in the Cox regressions were the date of blood sampling and age was used as timeline. The median follow-up time was 8.7 years (range 0–20.2 years) and 7.8 years (range 0.3–9.8 years) for ULSAM and PIVUS studies, respectively. In the primary Cox regression model, the continuous mLRRY was used as predictor of AD risk. In the model, we adjusted for the effects from the following set of confounders: APOE epsilon 4 genotype, age at sampling, smoking, BMI, diabetes, LDL and HDL cholesterol, hypertension, exercise habits, education level, and alcohol consumption. We also tested whether subjects harboring blood cell clones with large-scale (>1 Mb) structural aberrations on the autosomal chromosomes was affecting probability of AD-free follow-up time. All of the baseline confounders are summarized in Table S1. The results from the adjusted Cox regression show a strong association between level of LOY mosaicism at the time of blood sampling and risk for incident AD diagnosis during the follow-up time (HR = 6.80, 95% CI = 2.16–21.43, \(p = 0.0011\)) (Table 1). It is also noteworthy that other previously known risk factors for AD, i.e., age at sampling and APOE genotype, showed strong associations with AD risk, independent of the risk from LOY (Table 1). A model including only the significant confounders in the model (i.e., APOE epsilon 4 genotype and age at sampling) showed a similar result (Table S3).

To visualize the above results, we scored participants using two different thresholds for level of LOY mosaicism, as explained in Figure S1 and published previously.\(^30\) The first was the lower 99% confidence limit of experimentally induced mLRRY variation and the second threshold was a stricter cut-off at mLRRY \(-0.4\). These thresholds represent LOY in \(~10\%\) and \(~35\%\) of leukocytes, respectively.\(^30\) We plotted the adjusted probabilities of an AD-free follow-up time for the groups of participants scored with LOY and

---

**Figure 2. Mosaic LOY in Blood Cells Increases with Participant Sampling Age**

(A) Illustration of the association between of LOY and age of sampling in all 3,218 subjects from three independent cohorts included in the analyses. Linear regression shows that LOY (i.e., mLRRY) was associated with sampling age (ANOVA; \(F_{(1,3216)} = 26.79, p < 0.0001\)).

(B–D) Corresponding plots of the data from ULSAM (B), PIVUS (C), and EADI1 (D) cohorts. The dotted horizontal lines show the cut-off used for LOY scoring at the 99% confidence interval in the three independent cohorts.

(E) The increasing degree of LOY in four different age groups in the three included studies and p values adjusted for multiple testing using Tukey’s method is shown. The whiskers extend to illustrate the 1.5 inter-quantile range of the total variation in each age group.

(F) Summary of the observed frequencies of LOY in the different age groups plotted in (E).
men without detectable LOY, using these two thresholds. The comparison between the groups confirmed the results from analyses using the continuous mLRRY, that LOY in blood cells is a significant risk factor for AD diagnosis (Figures 4A and 4C, Tables S4 and S5). Interestingly, the risk of AD diagnosis during follow-up was higher in men harboring LOY in ~35% of cells compared to subject with LOY in ~10% of cells. Furthermore, risk for AD diagnosis during follow-up time among men with and without LOY in blood at sampling was evaluated separately in the ULSAM and PIVUS cohorts as shown in Figure S5. Among the 674 ULSAM men without any cancers, 90 were diagnosed with AD during follow-up and LOY in blood at sampling was a significant risk factor (HR = 1.90, 95% CI = 1.13–3.20, p = 0.0148). In the separate analysis of the smaller PIVUS cohort, the association between LOY and risk for AD was not significant but a similar trend could be observed.

Three independent papers recently demonstrated associations between LOY in blood and risk for various cancers^30,36,37 and in the present study we show that LOY in blood is also associated with risk for AD diagnosis. It is therefore reasonable to hypothesize that cancer and AD would act as competing risks when analyzing the effects of LOY. The suspicion of competing risks was supported by results from an analysis of 990 ULSAM and PIVUS men that were free from any cancer diagnosis and without AD diagnosis before the date of blood sampling (Figure S3). Specifically, we performed corresponding Cox regressions in this subset of 990 men, as described above for the 1,599 participants, and found an even stronger association between LOY and AD, first using the continuous mLRRY (HR = 28.41, 95% CI = 7.05–114.44, p < 0.0001) (Table S6) and then scoring participants with or without LOY using the same thresholds as above (Figures 4B and 4D, Tables S7 and S8). Furthermore, our recent report on an association between increased risk for non-hematological cancer from LOY in blood cells (HR = 3.76, 95% CI = 1.21–11.67, p = 0.022)^30 was done disregarding the competing risk of AD, and we therefore performed here a new analogous analysis after excluding participants with AD diagnosis. As anticipated, the association between LOY in blood cells and risk for non-hematological cancer mortality was strengthened in this refined analysis (HR = 5.58, 95% CI = 1.72–18.06, p = 0.0041) (Table S9).
As mentioned above, age and LOY are both associated with risk of AD and covariates in statistical analyses. Hence, the effect from age at sampling was fitted in Cox models testing association between LOY and AD (Table S10 and S11), supporting the first hypothesis. Thus, a reasonable interpretation of these results so far is that LOY has a direct influence on the pathogenesis of AD and cancer. The analysis of our data can also be discussed in the perspective of the differences between chronological and biological age, because LOY could be considered a marker for biological age, and as such a possible driver behind AD and cancer outcomes. Future investigations of LOY could therefore benefit from analyzing the biological rather than the chronological age. Our results are further consistent with LOY in peripheral blood representing the most frequently occurring ACE, affecting up to 20% of the oldest men in the studied population (Figure 2). In our analyses we also considered the ACEs containing post-zygotic structural aberrations on autosomes, but no significant association between risk for AD and ACEs carrying autosomal aberrations of >1 Mb in blood cells could be observed.

We have previously shown that cigarette smoking is likely to induce LOY in blood in a dose-dependent and transient manner. It has further been suggested that smoking is a risk factor for development of AD. Moreover, smoking is a well-known risk factor for cancer development, with lung cancer being the prime cause of cancer-associated death. Smoking is also a risk factor for tumors outside the respiratory tract and these are more common in males than females. Considering the above, we have thoroughly tested the possible effects of smoking in the statistical models in the current paper. In order to understand the relationship between smoking, LOY, and risk for AD, we included smoking as a co-covariate in our analyses. There was no significant association between smoking status at 70 years of age and AD diagnosis in our dataset. We further tested a model when current smokers were excluded from the analysis. The results showed that the subset of non-smokers still have an increased risk for AD when affected with LOY (Table S12). In summary, our results imply that smoking could induce LOY, which in turn increases risks for both cancer and AD.

### Discussion

Our results suggest that men with mosaic LOY in peripheral blood cells are at an increased risk of AD (Figures 3 and 4), in addition to the previously described risk of non-hematological cancer. A critical question is whether LOY per se is important in the pathogenesis of AD and cancer, or whether LOY is a phenotypically neutral product of aging that co-varies with other common aging-related phenotypes, for example cancer and AD? We carefully considered the latter hypothesis by fitting the age at sampling as a co-covariate in the statistical models evaluating LOY and AD. Moreover, we performed analyses of association between LOY and AD in two subsets of men within narrow age windows, which would tend to neutralize an effect from age at sampling. The association between LOY in blood and incident AD was robust in these analyses, whereas the effect of age was not significant (Tables S10 and S11), supporting the first hypothesis. Thus, a reasonable interpretation of these results so far is that LOY has a direct influence on the pathogenesis of AD and cancer. The analysis of our data can also be discussed in the perspective of the differences between chronological and biological age, because LOY could be considered a marker for biological age, and as such a possible driver behind AD and cancer outcomes. Future investigations of LOY could therefore benefit from analyzing the biological rather than the chronological age. Our results are further consistent with LOY in peripheral blood representing the most frequently occurring ACE, affecting up to 20% of the oldest men in the studied population (Figure 2). In our analyses we also considered the ACEs containing post-zygotic structural aberrations on autosomes, but no significant association between risk for AD and ACEs carrying autosomal aberrations of >1 Mb in blood cells could be observed.

We have previously shown that cigarette smoking is likely to induce LOY in blood in a dose-dependent and transient manner. It has further been suggested that smoking is a risk factor for development of AD. Moreover, smoking is a well-known risk factor for cancer development, with lung cancer being the prime cause of cancer-associated death. Smoking is also a risk factor for tumors outside the respiratory tract and these are more common in males than females. Considering the above, we have thoroughly tested the possible effects of smoking in the statistical models in the current paper. In order to understand the relationship between smoking, LOY, and risk for AD, we included smoking as a co-covariate in our analyses. There was no significant association between smoking status at 70 years of age and AD diagnosis in our dataset. We further tested a model when current smokers were excluded from the analysis. The results showed that the subset of non-smokers still have an increased risk for AD when affected with LOY (Table S12). In summary, our results imply that smoking could induce LOY, which in turn increases risks for both cancer and AD.

### Table 1. Cox Hazards Regression Model Evaluating the Association between LOY in Blood Cells and Risk to Be Diagnosed with Incident Alzheimer Disease during Follow-up Time in the ULSAM and PIVUS Studies after Adjusting for Potential Confounders

|                       | HR    | 95% CI          | p Value |
|-----------------------|-------|-----------------|---------|
| APOE genotype         | 2.80  | 1.42–5.54       | 0.0030**|
| Age at sampling       | 1.24  | 1.15–1.33       | <0.0001***|
| Smoking               | 1.36  | 0.77–2.39        | 0.2870  |
| BMI                   | 0.95  | 0.89–1.01        | 0.1298  |
| Diabetes              | 0.96  | 0.53–2.04        | 0.9141  |
| LDL cholesterol       | 1.07  | 0.87–1.33        | 0.5025  |
| HDL cholesterol       | 1.05  | 0.54–1.66        | 0.8558  |
| Hypertension          | 0.81  | 0.54–1.22        | 0.3145  |
| Exercise habits        | 0.94  | 0.37–3.11        | 0.9022  |
| Education level       | 1.01  | 0.69–1.48        | 0.9711  |
| Alcohol               | 1.00  | 0.97–1.03        | 0.9910  |
| Autosomal aberrations (>1 Mb) | 1.95  | 0.89–4.28        | 0.0976  |
| LOY (continuous mLRRY)| 6.80  | 2.16–21.43       | 0.0011**|

Abbreviations are as follows: HR, hazard ratio; CI, confidence interval. ***p < 0.001, **p < 0.01.
Our findings may suggest a role of LOY observed in non-cancerous blood cells on disease processes that take place in other tissues. Thus, an intriguing question is: what is the mechanism involving LOY in blood in the development of two radically different conditions, one being a neurodegenerative process in the central nervous system (CNS) and another encompassing abnormal proliferation of cells leading to tumors in various organs? We hypothesized previously that defective functions of immunosurveillance of the immune cells in blood might be related the effect of LOY on increased risk of tumor development in other organs.\(^{30,31}\) A deficient immunosurveillance in the CNS, i.e., a process that normally should eliminate abnormal cells related to AD phenotype in the brain, has been proposed as a mechanism in disease development.\(^{51–53}\) Moreover, the AD literature provides independent lines of evidence pointing toward the important role of the immune system in disease development.\(^{10,54}\) For instance, certain bacterial and viral infections are associated with increased risk of AD,\(^{54,55}\) suggesting that stress or deficiencies in the normal functions of the immune system could be involved in pathogenesis of AD. Hence, generally disturbed immune system functions, as an effect of LOY, could link the increased risk of AD as well as the risk of various tumor types. Another and non-mutually exclusive hypothesis is that LOY in blood cells could be a mirror of parallel processes of chromosomal instability also present in other cell types, such as neurons. It has been shown that AD neurons re-enter the cell cycle, leading to increased risk of various mitotic errors, such as aneuploidy, that has been documented in both normal and AD-affected brain cells.\(^{56,57}\)

Regardless of the underlying mechanism(s) for the increased risks of AD and cancer in men with LOY in blood, our and others’ published results reinforce a role of factors on chromosome Y in various, still poorly explored biological processes, other than sex determination and sperm production.\(^{20,21,36–39,58}\) Furthermore, our results demonstrate the importance of ACEs harboring post-zygotic aberrations, i.e., lifetime-acquired genetic variants, on the risk of development of common disease. In developed countries, dementia represents the third most common cause of morbidity/mortality in humans\(^ {59}\) and about one in
Table 2. Summary of the Main Findings from Analyses of EADI1 Study and Combined Analyses of ULSAM and PIVUS Cohorts

| Association between LOY and AD without Excluding Men with Cancer | EADI1 | ULSAM + PIVUS |
|---------------------------------------------------------------|--------|--------------|
| Unadj. K-S test | D = 0.07 (p = 0.0198) | D = 0.13 (p = 0.0137) | Figure 3 |
| Adj. logistic regression | OR = 2.80 (p = 0.0184) | | Table S2B |
| Adj. ANCOVA | F = 5.63 (p = 0.0178) | | Table S2A |
| Adj. Cox (cont. mLRRY) | HR = 6.80 (p = 0.0011) | | Table 1 |
| Adj. Cox (LOY 1/0 99% CI) | HR = 4.16 (p = 0.0085) | | Table S3 |
| Adj. Cox (LOY 1/0 −0.4) | HR = 1.63 (p = 0.0260) | Figure 4A |
| Adj. Cox (LOY 1/0 −0.4) | HR = 2.90 (p = 0.0250) | Figure 4C |

| Association between LOY and AD among Men 70–75 Years Old | EADI1 | ULSAM + PIVUS |
|----------------------------------------------------------|--------|--------------|
| Adj. Cox (cont. mLRRY) | HR = 25.82 (p < 0.0001) | | Table S10 |

| Association between LOY and AD among Men 75–80 Years Old | EADI1 | ULSAM + PIVUS |
|----------------------------------------------------------|--------|--------------|
| Adj. Cox (cont. mLRRY) | HR = 9.92 (p = 0.0489) | | Table S11 |

| Association between LOY and AD in Non-smokers | EADI1 | ULSAM + PIVUS |
|------------------------------------------------|--------|--------------|
| Adj. Cox (cont. mLRRY) | HR = 6.04 (p = 0.0131) | | Table S12 |

| Association between LOY and Cancer without Excluding Men with AD | EADI1 | ULSAM + PIVUS |
|---------------------------------------------------------------|--------|--------------|
| Adj. Cox (cont. mLRRY) | HR = 3.76 (p = 0.022)† | | Table S9 |

| Association between LOY and Cancer after Excluding Men with AD | EADI1 | ULSAM + PIVUS |
|---------------------------------------------------------------|--------|--------------|
| Adj. Cox (cont. mLRRY) | HR = 5.58 (p = 0.0041) | | Table S9 |

The “cont. mLRRY” is the continuous mLRRY estimate (i.e., the median of the log R ratio values of SNP-array probes positioned within the male-specific region of chromosome Y) reflecting the degree of LOY mosaicism in each participant. We also scored participants as 1 or 0 based on their mLRRY value using two different thresholds, i.e., mLRRY < −0.4 and mLRRY < 99% CI, as further described in the text. Abbreviations are as follows: D, the Kolmogorov-Smirnov test statistic; HR, hazard ratio; OR, odds ratio.

*aUnadj. K-S test = Unadjusted Kolmogorov-Smirnov test.*

*bLogistic regression model using AD status (1/0) as dependent variable and adjusting for the confounders APOE genotype and age at sampling.*

*cANCOVA model testing the continuous mLRRY estimate as dependent variable and adjusting for the confounders APOE genotype and age at sampling.*

*dCox hazards regression models testing the effect from the level of LOY in blood and risk for AD diagnosis during follow-up time, after adjusting survival from the 12 confounders summarized in Table S1 (i.e., APOE epsilon 4 genotype, age at sampling, smoking, BMI, diabetes, LDL and HDL cholesterol, hypertension, exercise habits, education level, alcohol consumption, and autosomal aberrations >1 Mb).*

*eCox hazards regression model testing the effect from level of LOY in blood and risk for AD diagnosis during follow-up time, after adjusting survival only for the significant confounders (Table S3).*

*fThe association between LOY and risk for cancer without excluding men with AD has been published.*

The “cont. mLRRY” is the continuous mLRRY estimate (i.e., the median of the log R ratio values of SNP-array probes positioned within the male-specific region of chromosome Y) reflecting the degree of LOY mosaicism in each participant. We also scored participants as 1 or 0 based on their mLRRY value using two different thresholds, i.e., mLRRY < −0.4 and mLRRY < 99% CI, as further described in the text. Abbreviations are as follows: D, the Kolmogorov-Smirnov test statistic; HR, hazard ratio; OR, odds ratio.

*aUnadj. K-S test = Unadjusted Kolmogorov-Smirnov test.*

*bLogistic regression model using AD status (1/0) as dependent variable and adjusting for the confounders APOE genotype and age at sampling.*

*cANCOVA model testing the continuous mLRRY estimate as dependent variable and adjusting for the confounders APOE genotype and age at sampling.*

*dCox hazards regression models testing the effect from the level of LOY in blood and risk for AD diagnosis during follow-up time, after adjusting survival from the 12 confounders summarized in Table S1 (i.e., APOE epsilon 4 genotype, age at sampling, smoking, BMI, diabetes, LDL and HDL cholesterol, hypertension, exercise habits, education level, alcohol consumption, and autosomal aberrations >1 Mb).*

*eCox hazards regression model testing the effect from level of LOY in blood and risk for AD diagnosis during follow-up time, after adjusting survival only for the significant confounders (Table S3).*

*fThe association between LOY and risk for cancer without excluding men with AD has been published.*

The American Journal of Human Genetics 98, 1208–1219, June 2, 2016

three people will get a cancer diagnosis during life. We hypothesize that the measurement of LOY in blood cells of adult/aging men could become a new, early predictive biomarker for AD and cancer, thus helping to relieve some of the huge burden that these diseases pose on individuals and society. This would also be well in line with an anticipated shift into a more preventive and personalized medical care. Finally, it has been known for centuries that men have a shorter life expectancy compared to women, even in regions of the world with well-developed healthcare, but the underlying factor(s) behind this sex difference are not clear. Mosaic LOY in blood, being a male-specific risk factor for both AD and cancer, might at least partly explain why men on average live shorter than women.

Accession Numbers

The accession numbers for the genetic variants detected and analyzed in this study are dbVar: nstd92 and nstd127.

Supplemental Data

Supplemental Data include 5 figures, 12 tables, and a list of the members of The European Alzheimer’s Disease Initiative Investigators and can be found with this article online at http://dx.doi.org/10.1016/j.ajhg.2016.05.014.
Conflicts of Interest

J.P.D. and L.A.F. are co-founders and shareholders in Cray Innovation AB with patent application PCT/EP2015/052898 protecting the commercial applications arising from the work in this publication.

Acknowledgments

We thank John Armour, Bradley Hyman, Alfredo Ramirez, and Constantin Polychronachos for critical evaluation of the manuscript. The study was sponsored by funding from the Olle Enqvist Byggmästare Foundation and European Research Council (ERC) Starting Grant to L.A.F. and by the Swedish Cancer Society, the Swedish Research Council, the Swedish Heart-Lung Foundation, the Torsten Söderberg Foundation, and Sci-Life-Lab-Uppsala to J.P.D. Genotyping and next-generation sequencing was performed by the SNP&SEQ Technology Platform in Uppsala, Sweden, and supported by Wellcome Trust Grants WT098017, WT064890, WT090532, Uppsala University, Uppsala University Hospital, the Swedish Research Council, and the Swedish Heart-Lung Foundation. The SNP&SEQ Technology Platform is part of Science for Life Laboratory at Uppsala University and supported as a national infrastructure by the Swedish Research Council. C.M.L. is a Wellcome Trust Research Career Development Fellow (086596/Z/09/2). The work on EAD1 dataset was funded by the French National Fondation on Alzheimer’s disease and related disorders, The Fondation pour la recherche sur le cerveau (FRC), the Lille Métropole Communauté urbaine council, and the LABEX (laboratory of excellence program investment for the future) DISTALZ Metropole Communaute urbaine council, and the LABEX (laboratory of excellence program investment for the future) DISTALZ grant (Development of Innovative Strategies for a Transdisciplinary approach to ALZheimer’s disease).

Received: February 19, 2016
Accepted: May 9, 2016
Published: May 23, 2016

Web Resources

A Package for Survival Analysis in S, http://www.mayo.edu/research/documents/tr53pdf/doc-10027379
Burrows-Wheeler Aligner, http://bio-bwa.sourceforge.net/
Control FREEC, http://bioinfo-out.curie.fr/projects/freec/dbVar, http://www.ncbi.nlm.nih.gov/dbvar/
GATK, https://www.broadinstitute.org/gatk/
IGAP Summary Statistics, http://web.pasteur-lille.fr/en/recherche/u744/igap/igap_download.php
ILLUMINA Tech Note, www.illumina.com/content/dam/illumina-marketing/documents/products/appnotes/appnote_cnvs_loh.pdf
OMIM, http://www.omim.org/
Picard, http://broadinstitute.github.io/picard/
PIVUS, http://www.medsci.uu.se/pivus/
ULSAM, http://www2.pubcare.uu.se/ULSAM/index.htm

References

1. Bettens, K., Sleegers, K., and Van Broeckhoven, C. (2013). Genetic insights in Alzheimer’s disease. Lancet Neurol. 12, 92–104.
2. Hardy, J., and Selkoe, D.J. (2002). The amyloid hypothesis of Alzheimer’s disease: progress and problems on the road to therapeutics. Science 297, 353–356.
3. Gatz, M., Reynolds, C.A., Fratiglioni, L., Johansson, B., Mortimer, J.A., Berg, S., Fiske, A., and Pedersen, N.L. (2006). Role of genes and environments for explaining Alzheimer disease. Arch. Gen. Psychiatry 63, 168–174.
4. Genin, E., Hannenquin, D., Wallon, D., Sleegers, K., Hiltunen, M., Combarros, O., Bullido, M.J., Engelborghs, S., De Deyn, P., Berr, C., et al. (2011). APOE and Alzheimer disease: a major gene with semi-dominant inheritance. Mol. Psychiatry 16, 903–907.
5. Lambert, J.C., Ibrahim-Verbaas, C.A., Harold, D., Naj, A.C., Sims, R., Bellenguez, C., DeStafano, A.L., Bis, J.C., Beecham, G.W., Grenier-Boley, B., et al.; European Alzheimer’s Disease Initiative (EADI); Genetic and Environmental Risk in Alzheimer’s Disease; Alzheimer’s Disease Genetic Consortium; Cohorts for Heart and Aging Research in Genomic Epidemiology (2013). Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer’s disease. Nat. Genet. 45, 1452–1458.
6. Escott-Price, V., Bellenguez, C., Wang, L.S., Choi, S.H., Harold, D., Jones, L., Holmes, P., Gerrish, A., Vedernikov, A., Richards, A., et al.; United Kingdom Brain Expression Consortium; Cardiovascular Health Study (CHS) (2014). Genome-wide analysis detects two new susceptibility genes for Alzheimer’s disease. PLoS ONE 9, e94661.
7. Guerreiro, R., Wojtas, A., Bras, J., Carraquillo, M., Rogaeva, E., Majounie, E., Cruchaga, C., Sassi, C., Kauwe, J.S., Younkin, S., et al.; Alzheimer Genetic Analysis Group (2013). TREM2 variants in Alzheimer’s disease. N. Engl. J. Med. 368, 117–127.
8. Lambert, J.C., Grenier-Boley, B., Harold, D., Zelenika, D., Chouraki, V., Kamatani, Y., Sleegers, K., Ikram, M.A., Hiltunen, M., Reitz, C., et al. (2013). Genome-wide haplotype association study identifies the FRMD4A gene as a risk locus for Alzheimer’s disease. Mol. Psychiatry 18, 461–470.
9. Benitez, B.A., Jin, S.C., Guerreiro, R., Graham, R., Lord, J., Harold, D., Sims, R., Lambert, J.C., Gibbs, J.R., Bras, J., et al.; 3C Study Group; EAD1 consortium; Alzheimer’s Disease Genetic Consortium (ADGC); Alzheimer’s Disease Neuroimaging Initiative (ADNI); GERAD Consortium (2014). Missense variant in TREM2 protects against Alzheimer’s disease. Neuron 83, 510.e19–510.e26.
10. International Genomics of Alzheimer’s Disease Consortium (IGAP) (2014). Convergent genetic and expression data implicate immunity in Alzheimer’s disease. Alzheimers Dement. 11, 658–671.
11. Chapman, J., Rees, E., Harold, D., Ivanov, D., Gerrish, A., Sims, R., Hollingworth, P., Stretton, A., Holmes, P., Owen, M.J., et al.; GERAD1 Consortium (2013). A genome-wide study shows a limited contribution of rare copy number variants to Alzheimer’s disease risk. Hum. Mol. Genet. 22, 816–824.
12. Forsberg, L.A., Absher, D., and Dumanski, J.P. (2013). Non-coding genome-wide analysis identifies the FRMD4A gene as a risk locus for Alzheimer’s disease. Neurobiol. Aging 35, 1510.e19–1510.e26.
13. Forsberg, L.A., Risi, C., Razaghian, H.R., Pakalapati, G., Waite, L., Thibault, K.S., Ronowicz, A., Wineinger, N.E., Tiwari, H.K., Boomsma, D., et al. (2012). Age-related somatic structural changes in the nuclear genome of human blood cells. Am. J. Hum. Genet. 90, 217–228.
14. Jacobs, K.B., Yeager, M., Zhou, W., Wacholder, S., Wang, Z., Rodriguez-Santiago, B., Hutchinson, A., Deng, X., Liu, C., Horner, M.J., et al. (2012). Detectable clonal mosaicism and its relationship to aging and cancer. Nat. Genet. 44, 651–658.
Bonnefond, A., Skrobek, B., Lobbens, S., Eury, E., Thuillier, D., Janvion, M., Delattre, O., Duport, A., and Gouya, M.T. (2013). The significance of isolated Y chromosome loss in bone marrow metaphase cells from males over age 50 years. Leuk. Res. 37, 105–109.

Bonnefond, A., Skrobek, B., Lobbens, S., Eury, E., Thuillier, D., Cauchy, S., Lantieri, O., Falk, D., Riboli, E., Marre, M., et al. (2013). Association between large detectable clonal mosaicism and type 2 diabetes with vascular complications. Nat. Genet. 45, 1040–1043.

dec Carvalho, C.M., and Santos, E.R. (2005). Human Y-chromosome variation and male dysfunction. J. Mol. Genet. Med. J. 63, 75.

Hughes, J.F., and Rozen, S. (2012). Genomics and genetics of human and primate Y chromosomes. Annu. Rev. Genomics Hum. Genet. 13, 83–108.

Jacobs, P.A., Brunton, M., Court Brown, W.M., Doll, R., and Goldstein, H. (1963). Change of human chromosome count distribution with age: evidence for a sex difference. Nature 197, 1080–1081.

Pierre, R.V., and Hoagland, H.C. (1972). Age-associated aneuploidy: loss of Y chromosome from human bone marrow cells with aging. Cancer 30, 889–894.

Nowinski, G.P., Van Dyke, D.L., Tillick, B.C., Jacobsen, G., Babu, V.R., Worsham, M.J., Wilson, G.N., and Weiss, L. (1990). The frequency of aneuploidy in cultured lymphocytes and type 2 diabetes with vascular complications. Nat. Genet. 45, 1040–1043.

Bonnet, A., Skrobek, B., Lobbens, S., Eury, E., Thuillier, D., Cauchy, S., Lantieri, O., Falk, D., Riboli, E., Marre, M., et al. (2013). Association between large detectable clonal mosaicism and type 2 diabetes with vascular complications. Nat. Genet. 45, 1040–1043.

dec Carvalho, C.M., and Santos, E.R. (2005). Human Y-chromosome variation and male dysfunction. J. Mol. Genet. Med. J. 63, 75.

Hughes, J.F., and Rozen, S. (2012). Genomics and genetics of human and primate Y chromosomes. Annu. Rev. Genomics Hum. Genet. 13, 83–108.

Jacobs, P.A., Brunton, M., Court Brown, W.M., Doll, R., and Goldstein, H. (1963). Change of human chromosome count distribution with age: evidence for a sex difference. Nature 197, 1080–1081.

Pierre, R.V., and Hoagland, H.C. (1972). Age-associated aneuploidy: loss of Y chromosome from human bone marrow cells with aging. Cancer 30, 889–894.

Nowinski, G.P., Van Dyke, D.L., Tillick, B.C., Jacobsen, G., Babu, V.R., Worsham, M.J., Wilson, G.N., and Weiss, L. (1990). The frequency of aneuploidy in cultured lymphocytes and type 2 diabetes with vascular complications. Nat. Genet. 45, 1040–1043.

Bonnet, A., Skrobek, B., Lobbens, S., Eury, E., Thuillier, D., Cauchy, S., Lantieri, O., Falk, D., Riboli, E., Marre, M., et al. (2013). Association between large detectable clonal mosaicism and type 2 diabetes with vascular complications. Nat. Genet. 45, 1040–1043.
45. Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25, 1754–1760.

46. Boeva, V., Popova, T., Bleakley, K., Chiche, P., Cappo, J., Schleiermacher, G., Janoueix-Lerosey, I., Delattre, O., and Barillot, E. (2012). Control-FREEC: a tool for assessing copy number and allelic content using next-generation sequencing data. Bioinformatics 28, 423–425.

47. R Development Core Team (2012). R: A language and environment for statistical computing (Vienna, Austria: R Foundation for Statistical Computing).

48. Cataldo, J.K., Prochaska, J.J., and Glantz, S.A. (2010). Cigarette smoking is a risk factor for Alzheimer's Disease: an analysis controlling for tobacco industry affiliation. J. Alzheimers Dis. 19, 465–480.

49. Norton, S., Matthews, F.E., Barnes, D.E., Yaffe, K., and Brayne, C. (2014). Potential for primary prevention of Alzheimer's disease: an analysis of population-based data. Lancet Neurol. 13, 788–794.

50. Jha, P., Ramasundarahettige, C., Landsman, V., Rostron, B., Thun, M., Anderson, R.N., McAfee, T., and Peto, R. (2013). 21st-century hazards of smoking and benefits of cessation in the United States. N. Engl. J. Med. 368, 341–350.

51. Simard, A.R., Soulet, D., Gowing, G., Julien, J.P., and Rivest, S. (2006). Bone marrow-derived microglia play a critical role in restricting senile plaque formation in Alzheimer's disease. Neuron 49, 489–502.

52. Schwartz, M., and Shechter, R. (2010). Protective autoimmunity functions by intracranial immunosurveillance to support the mind: The missing link between health and disease. Mol. Psychiatry 15, 342–354.

53. Ousman, S.S., and Kubes, P. (2012). Immune surveillance in the central nervous system. Nat. Neurosci. 15, 1096–1101.

54. Maheshwari, P., and Elick, G.D. (2015). Bacterial infection and Alzheimer's disease: a meta-analysis. J. Alzheimers Dis. 43, 957–966.

55. Lovheim, H., Gilthorpe, J., Johansson, A., Eriksson, S., Hal-mans, G., and Elgh, F. (2015). Herpes simplex infection and the risk of Alzheimer’s disease-A nested case-control study. Alzheimers Dement. 11, 587–592.

56. Arendt, T. (2012). Cell cycle activation and aneuploid neurons in Alzheimer's disease. Mol. Neurobiol. 46, 125–135.

57. McConnell, M.J., Lindberg, M.R., Brennand, K.J., Piper, J.C., Voet, T., Cowing-Zitron, C., Shumilina, S., Lasken, R.S., Vermeesch, J.R., Hall, I.M., and Gage, F.H. (2013). Mosaic copy number variation in human neurons. Science 342, 632–637.

58. Case, L.K., Wall, E.H., Dragon, J.A., Saligrama, N., Krementsov, D.N., Moussawi, M., Zachary, J.F., Huber, S.A., Blankenhorn, E.P., and Teuscher, C. (2013). The Y chromosome as a regulatory element shaping immune cell transcriptomes and susceptibility to autoimmune disease. Genome Res. 23, 1474–1485.

59. James, B.D., Leurgans, S.E., Hebert, L.E., Scherr, P.A., Yaffe, K., and Bennett, D.A. (2014). Contribution of Alzheimer disease to mortality in the United States. Neurology 82, 1045–1050.

60. Ahmad, A.S., Ormiston-Smith, N., and Sasieni, P.D. (2015). Trends in the lifetime risk of developing cancer in Great Britain: comparison of risk for those born from 1930 to 1960. Br. J. Cancer 112, 943–947.

61. Blatt Kalben, B. (2000). Why Men Die Younger. N. Am. Actuar. J. 4, 83–111.

62. Central Intelligence Agency (2013). The World Factbook 2013-14 (Washington, DC: Central Intelligence Agency).