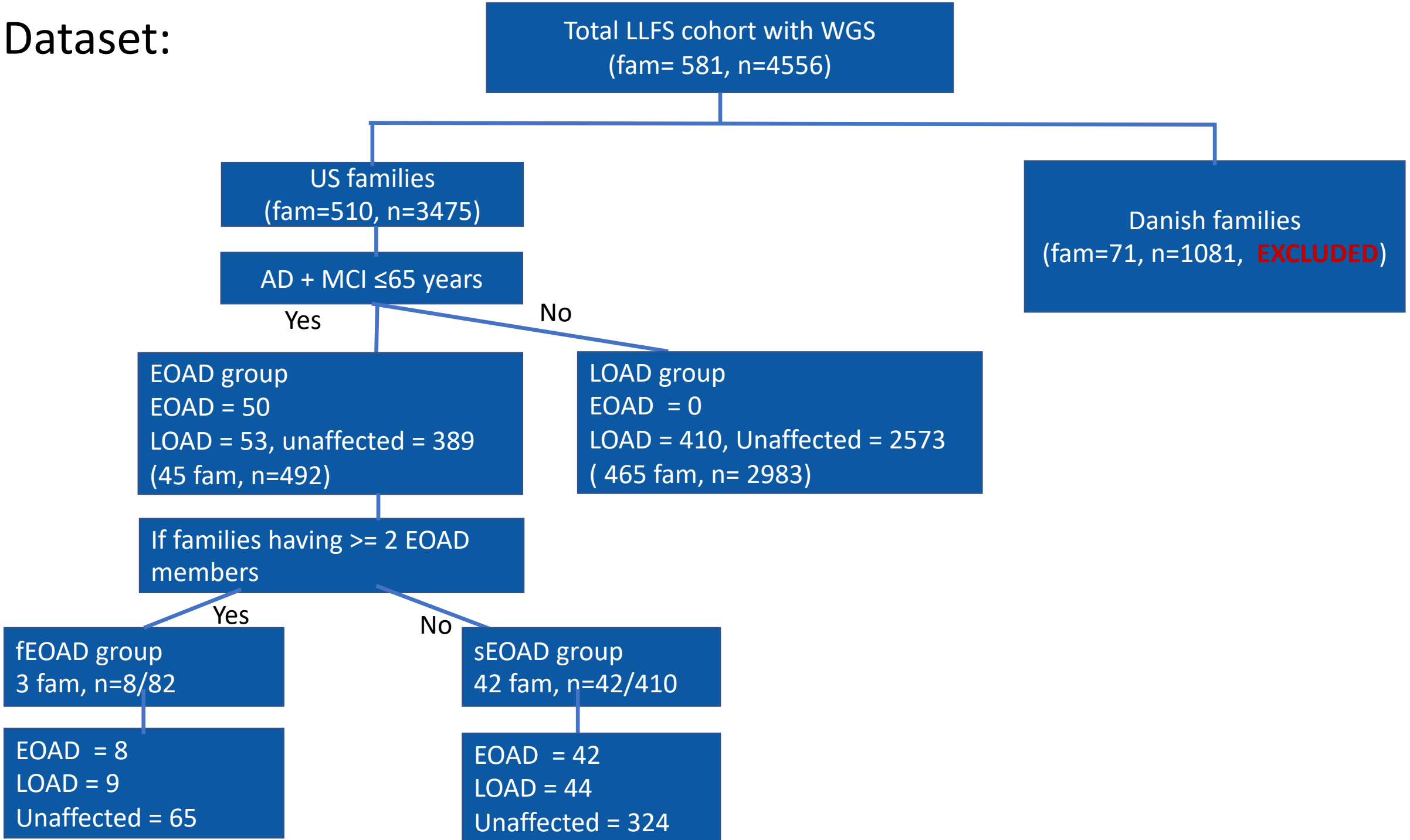
Variants Associated with Early Onset Alzheimer Disease in LLFS: A 2-Stage Study

Subinoy Biswas

Why study the genetics of Early Onset Alzheimer Disease (EOAD) in Long Life Family Study?

- LLFS family members tend to live to old age without chronic diseases (Wojczynski, 2022)
- *Of ~510 US LLFS families, 3 have familial EOAD, and 42 had sporadic EOAD.* These 3 may represent a familial form of EOAD (fEOAD), and there may exist a variant(s) that cause EOAD.
- We reason that:
- *"The 3 families (fEOAD) are likely to carry a Mendelian risk variant(s) that cause EOAD, these LLFS families are selected because they live to extreme old ages without major chronic diseases. If we can identify the same causal variants from other LLFS families AND YET ARE NOT AFFECTED with EOAD, these individuals may carry protective variant(s).*
- *For this purpose, the best candidate families to examine are: (1) Sporadic EOAD families (since one person with EOAD may not carry the protective variant(s), while other family members with LOAD or not yet affected may carry protective variants that delayed the onset); and (2) LOAD (all family members carry the protective variant(s), which is less likely).*

Dataset:


Total LLFS cohort with WGS
(fam= 581, n=4556)
US families
(fam=510, n=3475)
Danish families
(fam=71, n=1081, EXCLUDED)
AD + MCI ≤65 years
Yes
No
EOAD group
EOAD = 50
LOAD = 53, unaffected = 389
(45 fam, n=492)
LOAD group
EOAD = 0
LOAD = 410, Unaffected = 2573
(465 fam, n= 2983)
If families having >= 2 EOAD members
Yes
No
fEOAD group
3 fam, n=8/82
sEOAD group
42 fam, n=42/410
EOAD = 8
LOAD = 9
Unaffected = 65
EOAD = 42
LOAD = 44
Unaffected = 324

Demographic and clinical characteristics of the US LLFS members

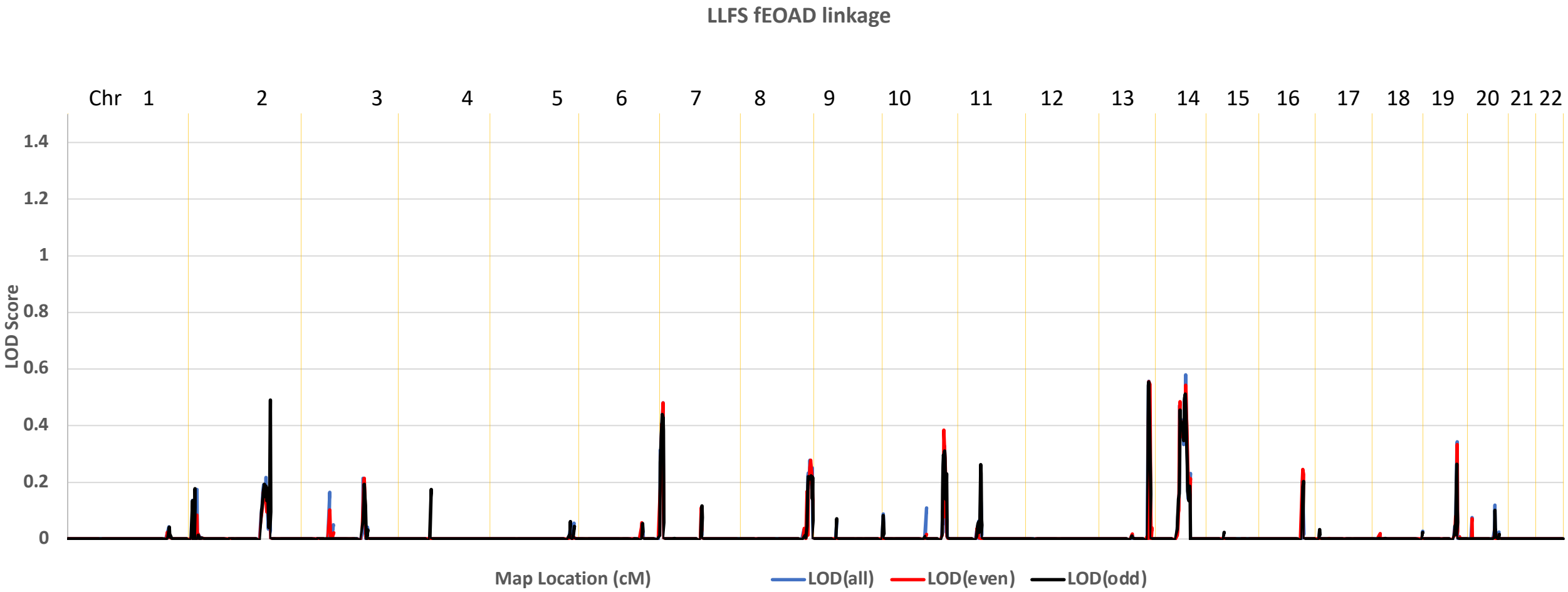
	Overall		fEOAD		sEOAD		LOAD	
Characteristics	n	Mean/%	n	Mean/%	n	Mean/%	n	Mean/%
Families	510		3		42		465	
Participants	3475		82		410		2983	
AD status								
EOAD Affected	50	1.43%	8	9.76%	42	10.24%	0	
EOAD Affected Age			8	57.25±6.20	42	60.11±4.76		
LOAD Affected	463	13.32%	9	10.98%	44	10.73%	410	13.74%
LOAD Affected Age				81.11±8.56		83.36±9.22	410	84.91±8.48
Unaffected	2962	85.23%	65	79.26%	324	79.02%	2573	86.25%
Unaffected Age		**73.65±14.33**		**71.64±13.31**		**73.52±13.69**		**73.72±14.44**
% Females	1916	55.13	46	56.1	225	54.9	1645	55.14
*Age		74.92±14.24		71.28±13.46		73.20±13.72		75.26±14.3

*Age at onset for AD individuals, otherwise latest visit age

Can we localize genes that are segregating with EOAD in families from a genome wide search?

Genome wide linkage analysis

Genome wide linkage signals of fEOAD

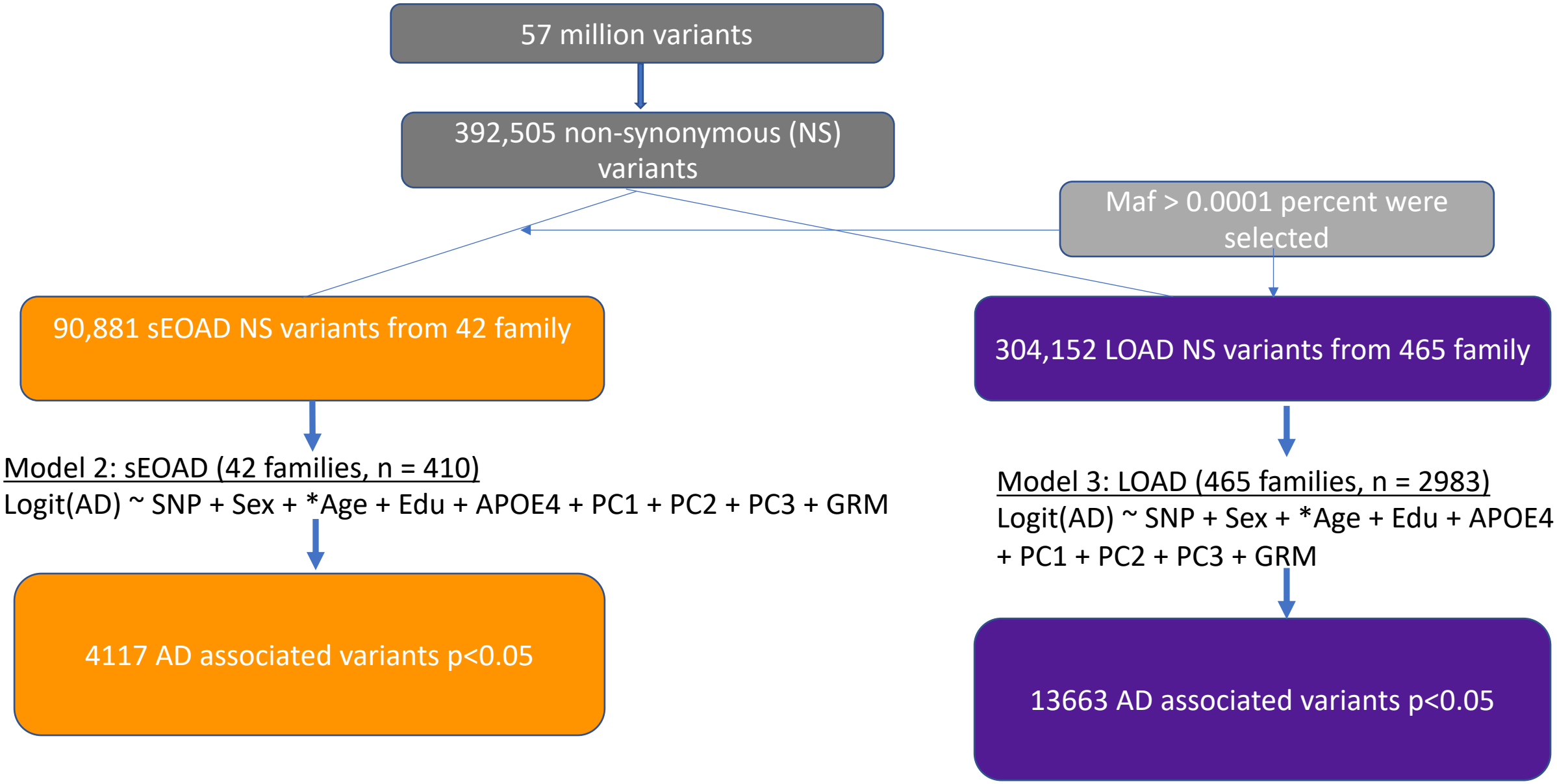


- **Trait(* Age follow)** ~ AD + Sex + Education + APOE2 + APOE4 + PC1 + PC2 + PC3 + Kinship coefficient
- *Age at onset for AD individuals, otherwise latest visit age
- SOLAR software was used

Are there non-synonymous potential causal variants that are associated with AD?

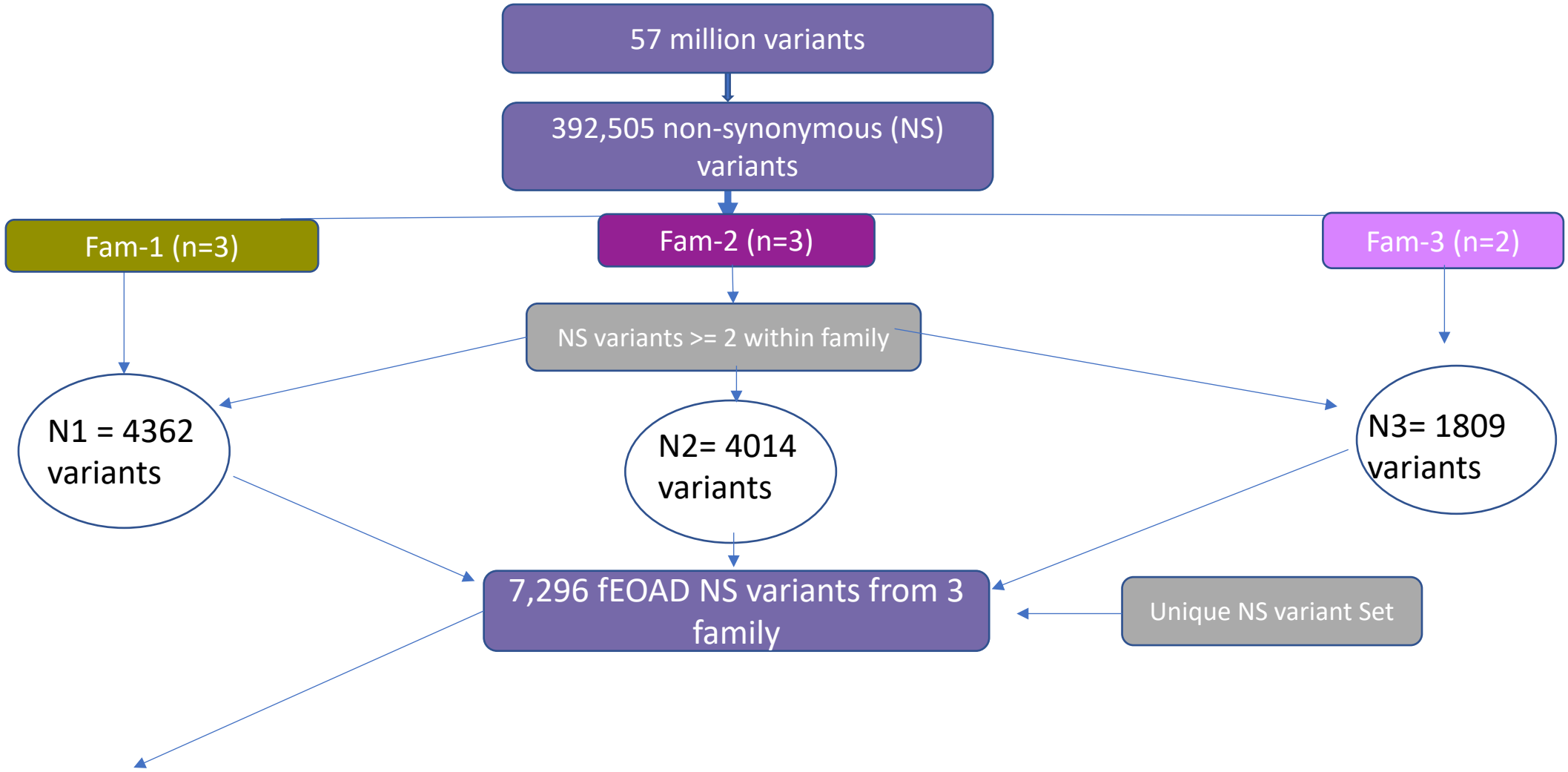
Non-synonymous variant analysis

Workflow of nonsynonymous variants study in sporadic early onset and late onset AD families



*Age at onset for AD individuals, otherwise latest visit age

Workflow of early onset nonsynonymous variants study



Model 1: fEOAD (3 Families, n =82)

Logit(AD) ~ SNP + Sex + *Age + Edu + APOE4 + PC1 + PC2 + PC3 + GRM

*Age at onset for AD individuals, otherwise latest visit age

80 AD associated potential causal variants p<0.01 and MAF of affected > unaffected

Table: fEOAD and sEOAD common significant variants by association analysis

CHR	SNP	BP	Gene	A1(alt)	A2(ref)	N	MAF	PVAL	OR	F_A	F_U	Link_LOD
1				G	C	76	0.132	0.0233	7.22	0.177	0.109	1.10
1				T	G	77	0.338	0.0177	3.59	0.559	0.262	1.03
1				G	T	76	0.086	0.0403	10.91	0.088	0.086	1.01
1				G	A	74	0.230	0.0069	5.44	0.313	0.198	1.20
1				G	C	77	0.227	0.0121	4.83	0.294	0.200	1.20
2				A	G	76	0.230	0.0099	4.17	0.353	0.180	1.13
2				A	G	77	0.266	0.0402	2.88	0.324	0.231	1.28
6				C	G	77	0.344	0.0427	2.84	0.559	0.315	1.20
7				T	C	71	0.218	0.0389	3.47	0.412	0.172	1.40
10				G	T	74	0.088	0.0169	8.65	0.177	0.066	1.01
17				C	T	69	0.341	0.0064	5.91	0.607	0.297	2.20
17				G	C	76	0.191	0.0406	3.59	0.324	0.148	0.86

Model

Logit(AD) ~ SNP + Sex + *Age + Edu + APOE4 + PC1 + PC2 + PC3 + GRM

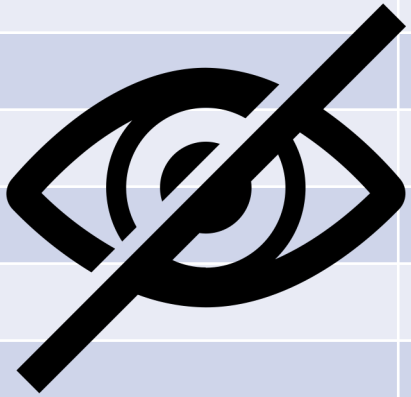
*Age at onset for AD individuals, otherwise latest visit age

Can we localize variants that are associated with AD by joint linkage analysis?

Pseudomarker analysis

Q: What are the top non-synonymous variants within *fEOAD group

				Dominant	Recessive
Chr	SNP	BP	Gene	Link(LOD)	Link(LOD)
2			**Culprit gene**	**2.00**	0.43
2			**Culprit gene**	**1.98**	0.43
2			**Culprit gene**	**1.66**	0.42
2			**Culprit gene**	**1.83**	0.35
2			**Culprit gene**	**1.95**	0.41
2			**Culprit gene**	**1.82**	0.34
2			**Culprit gene**	**2.00**	0.43
2			**Culprit gene**	**1.97**	0.42
7			--	0.66	1.68
11			--	1.71	0.22
14			--	1.61	0.76
17			--	1.62	0.86
17			--	2.06	0.85
19			--	1.65	0.69



*AD status were kept unknown for Age follow non-AD members >65 and <=70 years.

Pseudomarker linkage signals of culprit gene across fEOAD, sEOAD and LOAD

				*fEOAD_masked		sEOAD		LOAD	
Chr	SNP	BP	Gene	LOD DOM	LOD REC	LOD DOM	LOD REC	LOD DOM	LOD REC
2				**2.00**	0.43	0.34	0.27	0.07	0.21
2				**1.98**	0.43	0.50	0.70	0.05	0.16
2				**1.66**	0.42	0.76	0.35	0.01	0.03
2				**1.83**	0.35	0.87	0.46	0.00	0.01
2				**1.95**	0.41	0.57	0.54	0.20	0.40
2				**1.82**	0.34	0.36	0.28	0.07	0.23
2				**2.00**	0.43	0.25	0.34	0.06	0.22
2				**1.97**	0.42	1.13	0.83	0.00	0.04

*AD status were kept unknown for non-AD members >65 and <=70 years.

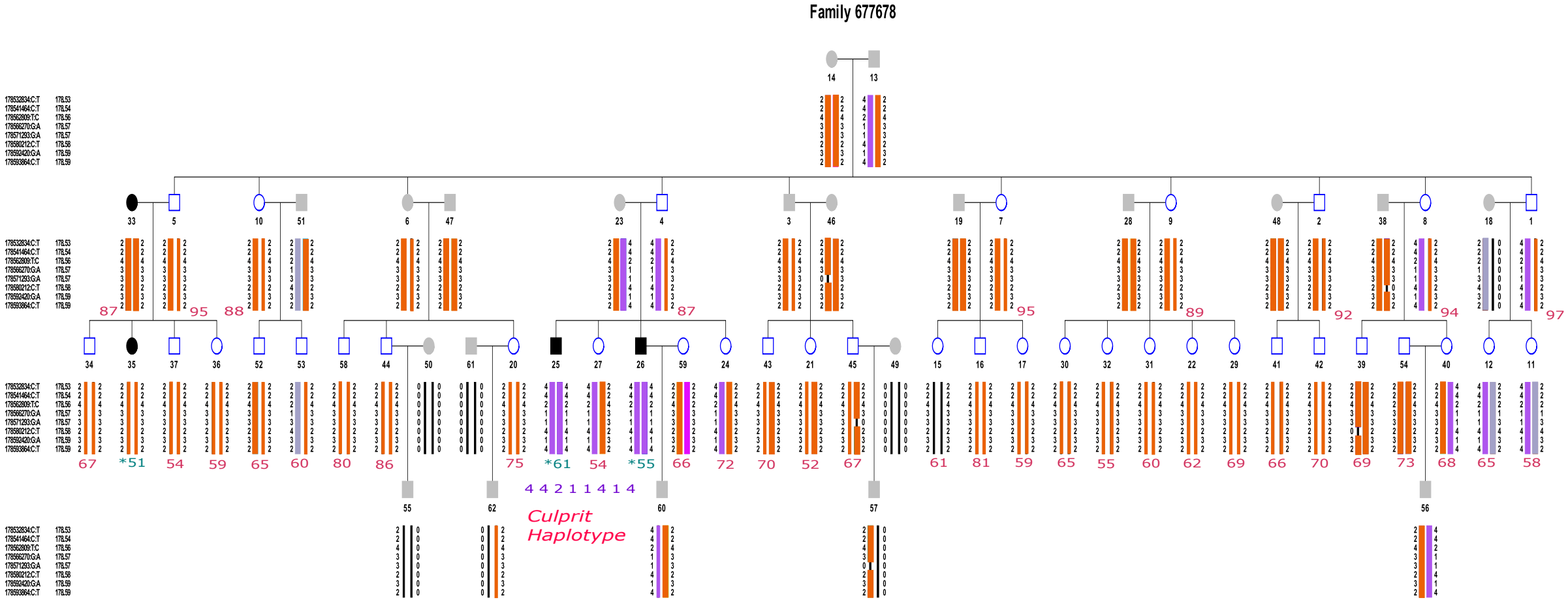
* Same results from fEOAD without masked group

Q: How culprit gene associatiated with AD in fEOAD, sEOAD and LOAD groups?

						fEOAD(3fam, AD=17/82)					sEOAD(42fam, AD=86/410)					LOAD(465fam, AD=410/2983)				
Chr	SNP	POS	Gene	A1(alt)	A2(ref)	N	F_A	F_U	PVAL	OR	N	F_A	F_U	PVAL	OR	N	F_A	F_U	PVAL	OR
2				T	C	76	0.294	0.125	0.0025	6.04	377	0.169	0.141	0.850	1.05	2705	0.197	0.154	0.0115	1.34
2				T	C	74	0.313	0.129	0.0014	6.89	371	0.161	0.145	0.870	0.96	2678	0.199	0.155	0.0109	1.34
2				G	A	75	0.353	0.183	0.0103	4.19	374	0.256	0.205	0.726	1.08	2673	0.281	0.238	0.0820	1.19
2				C	T	74	0.353	0.186	0.0112	4.11	378	0.259	0.215	0.715	1.09	2687	0.277	0.242	0.1952	1.13
2				A	G	77	0.294	0.146	0.0073	4.50	377	0.209	0.170	0.689	1.10	2683	0.215	0.186	0.1655	1.16
2				A	G	73	0.294	0.133	0.0033	5.92	377	0.163	0.142	0.965	0.99	2682	0.193	0.153	0.0278	1.29
2				T	C	76	0.294	0.148	0.0076	4.46	382	0.212	0.174	0.834	1.05	2712	0.223	0.193	0.1523	1.16
2				A	G	76	0.353	0.180	0.0099	4.17	378	0.247	0.201	0.787	1.06	2715	0.272	0.240	0.1815	1.14
2				A	G	77	0.294	0.123	0.0024	6.09	381	0.167	0.140	0.860	1.04	2712	0.194	0.154	0.0183	1.31
2				T	C	72	0.294	0.133	0.0035	5.85	366	0.175	0.144	0.824	1.06	2642	0.191	0.158	0.0269	1.29

Logistic regression model
AD status ~ NS_SNP + Age follow + Sex + Edu + APOE4 + PC1 .. PC3 + GRM

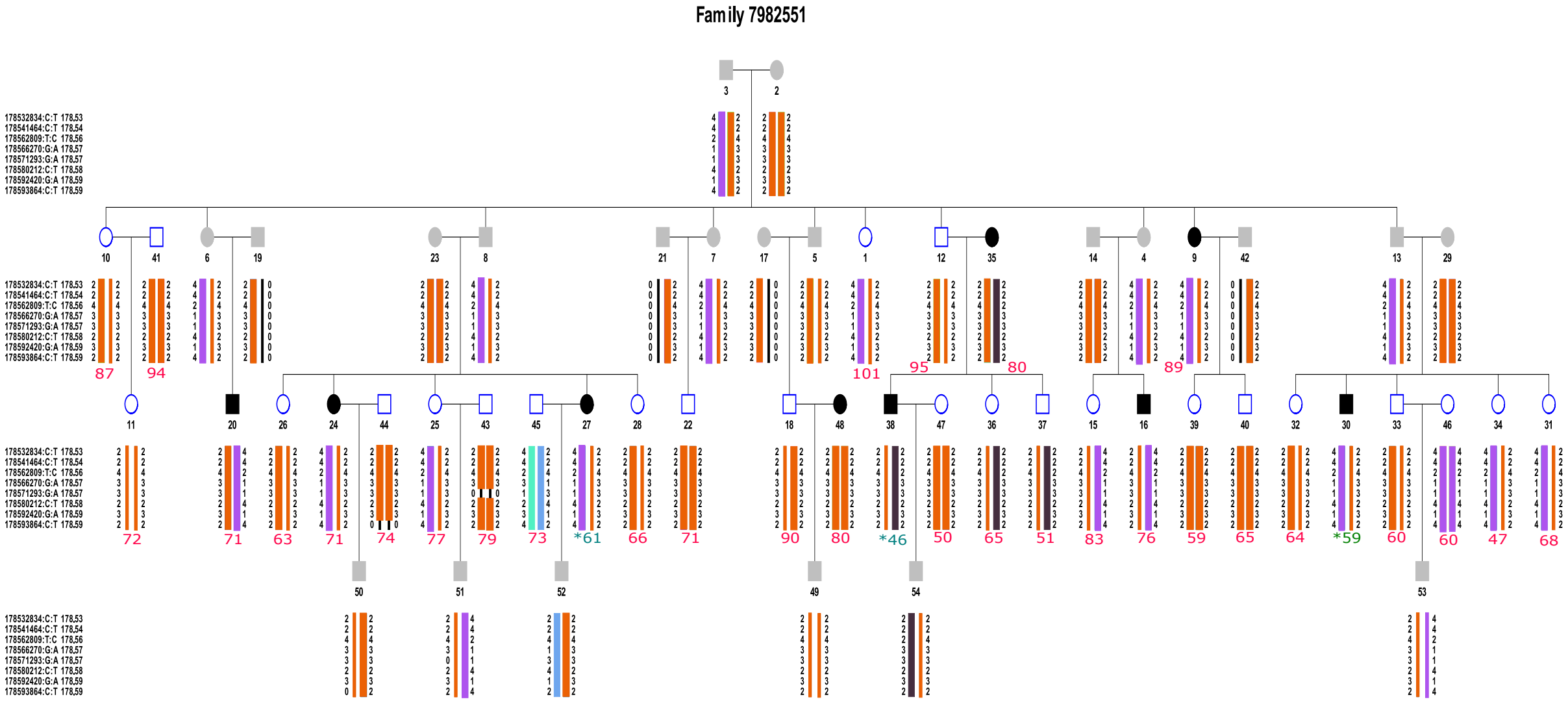
Haplotype of culprit gene in Family 30677678



The number below each individual is the Age. Either it's age-at-onset of AD or the age of last visit on non-AD.

* Marked individuals are early onset AD.

Haplotype of culprit gene in Family 37982551



Haplotype of culprit gene in Family 47735361