

Investigation of complement component *C4* copy number variation in human longevity

ST – Supplementary Tables

Supplementary Table S1a. Characteristics of samples

		Cases	Controls
C4A	Numbers before quality control	784	993
	Numbers after quality control	724	923
	Age		
	Median	101	53
	IQR	97-101	41-62
	Min – max	94-110	19-75
Sex (male:female)	363:361	491:432	
C4B	Smoking*		
	Smokers	21	195
	Non- smokers	640	306
	No information	63	422
	Numbers before quality control	784	993
	Numbers after quality control	728	920
C4C	Age		
	Median	101	53
	IQR	97-101	41-62
	Min – max	94-110	19-75
	Sex (male:female)	366:362	490:430
	Smoking*		
C4S	Smokers	22	194
	Non-smokers	643	302
	No information	63	424
	Numbers before quality control	784	993
	Numbers after quality control	736	929
	Age		
Median	101	53	
IQR	97-101	41-62	
Min – max	94-110	19-75	
Sex (male:female)	375:361	493:436	
C4S	Smoking*		
	Smokers	22	198
	Non-smokers	650	305
No information	64	426	

C4L	Numbers before quality control	784	993
	Numbers after quality control	649	899
	Age		
	Median	101	53
	IQR	98-101	41-62
	Min – max	94-110	19-75
Sex (male:female)	321:328	476:423	
C4L replication	Numbers before quality control	823	1202
	Numbers after quality control	714	890
	Age		
	Median	98	50
	IQR	96-99	40-65
	Min – max	91-108	18-79
Sex (male:female)	59:655	90:800	

IQR: interquartile range, *: “smokers” were defined as current smokers and recent ex-smokers (quit <3 years ago), “non-smokers” as never smokers and quit ≥3 years ago

For the C4L replication, smoking was not investigated.

Supplementary Table S1b. Quality control of *C4A*, *C4B*, *C4S* and *C4L*

	<i>C4A</i>		<i>C4B</i>		<i>C4S</i>		<i>C4L</i>		<i>C4L</i> replication	
	Cases	Controls	Cases	Controls	Cases	Controls	Cases	Controls	Cases	Controls
Before qc	784	993	784	993	784	993	784	993	823	1202
Replicate Number	770	955	771	953	766	947	770	956	822	1133
Confidence	745	937	757	945	755	937	663	903	716	890
Z score ≥ 2.65	741	934	746	929	752	934	663	903	715	890
$2.65 > z \text{ score} \geq 1.75$	738	927	742	924	750	933	663	903	714	890
$C4A+C4B \neq C4L+C4S$	724	923	728	920	736	929	649	899		

Before qc: numbers before quality control

Replicate number: removal of all samples with analyzed replicate number 0, 1 or 2

Confidence: removal of all samples with confidence < 0.95 and $|\text{CN}_{\text{predicted}} - \text{CN}_{\text{calculated}}| > 0.3$

Z score ≥ 2.65 : removal of all samples with z score ≥ 2.65

$2.65 > z \text{ score} \geq 1.75$: removal of all samples with $2.65 > z \text{ score} \geq 1.75$ and $|\text{CN}_{\text{predicted}} - \text{CN}_{\text{calculated}}| > 0.3$

$C4A+C4B \neq C4L+C4S$: removal of all samples with sum of *C4A* copies plus *C4B* copies not equal to the sum of *C4L* copies plus *C4S* copies (not available for *C4L* replication)

Supplementary Table S2. Frequencies of *C4A*, *C4S* and *C4L* copy numbers (absolute numbers are in parentheses)

	number of genes	0	1	2	3	4	5	6	10	p value
<i>C4A</i>	cases (LLI) n = 724	(5) 0.69%	(118) 16.30%	(426) 58.84%	(156) 21.55%	(19) 2.62%	(0) 0%	(0) 0%	(0) 0%	0.72
	controls n = 923	(10) 1.08%	(169) 18.31%	(517) 56.01%	(204) 22.10%	(22) 2.38%	(1) 0.11%	(0) 0%	(0) 0%	
<i>C4S</i>	cases (LLI) n = 736	(274) 37.23%	(342) 46.47%	(105) 14.27%	(13) 1.77%	(1) 0.14%	(1) 0.14%	(0) 0%	(0) 0%	0.51
	controls n = 929	(317) 34.12%	(437) 47.04%	(158) 17.01%	(13) 1.40%	(3) 0.32%	(1) 0.11%	(0) 0%	(0) 0%	
<i>C4L</i>	cases (LLI) n = 649	(1) 0.15%	(32) 4.93%	(172) 26.50%	(279) 42.99%	(154) 23.73%	(11) 1.69%	(0) 0%	(0) 0%	0.00065
	controls n = 899	(9) 1.00%	(73) 8.12%	(206) 22.91%	(328) 36.48%	(269) 29.92%	(13) 1.45%	(1) 0.11%	(0) 0%	
<i>C4L</i> replication	cases (LLI) n = 714	(10) 1.40%	(46) 6.44%	(176) 24.65%	(273) 38.24%	(184) 25.77%	(14) 1.96%	(10) 1.40%	(1) 0.14%	0.000021
	controls n = 890	(9) 1.01%	(80) 8.99%	(273) 30.67%	(342) 38.43%	(171) 19.21%	(15) 1.69%	(0) 0%	(0) 0%	

p value: p value for the comparison of allele frequencies between long-lived cases and younger controls (Fisher's exact test)

Supplementary Table S3. Results of the comparison of Q0 carrier state for the *C4B* gene with age groups as in Kramer et al. 1989 and Arason et al. 2003

a. Age distribution of cases and controls according to Kramer et al. 1989

	controls ≤53 years n = 463	cases (LLI) 61 - 110 years n = 970	p value
<i>C4B</i> *Q0 carriers	n = 118 25.49%	n = 265 27.32%	0.48

b. Age distribution of cases and controls according to Arason et al. 2003

	controls ≤54 years n = 487	cases (LLI) 55 - 110 years n = 1161	p value
<i>C4B</i> *Q0 carriers	n = 125 25.67%	n = 317 27.30%	0.50

p value: p value for the comparison of *C4B**Q0 frequencies between long-lived cases and younger controls (Fisher's exact test)

Supplementary Table S4. Results of the comparison of Q0 carrier state for the *C4L* gene (carrier and non-carrier information)

	≤ 45 n = 294	46 - 60 n = 351	61 - 75 n = 254	94 - 100 n = 233	> 100 n = 416	p value
<i>C4L</i> * <i>Q0</i> carriers	n = 29 9.86%	n = 32 9.12%	n = 21 8.27%	n = 11 4.72%	n = 22 5.29%	p_contr_trend = 0.52 p_case = 0.85 p_trend = 0.004
	all controls n = 82 9.12%			all cases (LLI) n = 33 5.08%		p_case_contr = 0.003
<i>C4L</i> * <i>Q0</i> non- carriers	n = 265 90.14%	n = 319 90.88%	n = 233 91.73%	n = 222 95.28%	n = 394 94.71%	
	all controls n = 817 90.88%			cases (LLI) n = 616 95.92%		

p_case_contr: p value for the comparison of *C4L***Q0* frequencies between long-lived cases and younger controls (Fisher's exact test; primary endpoint)

p_contr_trend: p value for trend test of *C4L***Q0* frequencies in the three control subgroups (Armitage trend test; secondary endpoint)

p_case: p value for the comparison of *C4L***Q0* frequencies in the two case subgroups (Fisher's exact test; secondary endpoint)

p_trend: p value for trend test of *C4L***Q0* frequencies in all five age groups (Armitage trend test; secondary endpoint)

Supplementary Table S5. Results Replication study in additional German sample: Results of the comparison of Q0 carrier state for the *C4L* gene (carrier and non-carrier information)

	≤ 45 n = 344	46 - 60 n = 248	61 - 79 n = 298	91 - 108* n = 714		p value
<i>C4L</i> * <i>Q0</i> carriers	n = 41 11.92%	n = 25 10.08%	n = 23 7.72%			p_contr_trend = 0.08 p_trend = 0.03
	all controls n = 89 10.00%			cases (LLI) n = 56 7.84%		p_case_contr = 0.14
<i>C4L</i> * <i>Q0</i> non- carriers	n = 303 88.08%	n = 223 89.92%	n = 275 92.28%	n = 577 92.62%	n = 81 89.01%	
	all controls n = 801 90.00%			cases (LLI) n = 658 92.16%		

* due to the small number of centenarians in the replication sample we did not subdivide the case sample into a nonagenarian and centenarian subgroup

For abbreviations see legend to Supplementary Table 4.

Supplementary Table S6. Replication study in additional German sample: Results of the comparison of Q0 carrier state for the *C4L* gene with age groups as in the initial study

	≤ 45 n = 344	46 - 60 n = 248	61 - 75 n = 254	94 - 108* n = 707	p value
<i>C4L</i> *Q0 carriers	n = 41 11.92%	n = 25 10.08%	n = 19 7.48%		p_contr_trend = 0.08 p_trend=0.03
	all controls n = 85 10.05%			cases (LLI) n = 56 7.92%	p_case_contr = 0.16

For abbreviations see legend to Supplementary Table 4.