

Rejuvenation Research

Rejuvenation Research: <http://mc.manuscriptcentral.com/rejuvenationresearch>

Association of FOXO3A Locus with Extreme Longevity in the Southern Italian Centenarian Study

Journal:	<i>Rejuvenation Research</i>
Manuscript ID:	REJ-2008-0827.R1
Manuscript Type:	Research Articles
Date Submitted by the Author:	29-Jan-2009
Complete List of Authors:	viviani anselmi, chiara; IRCCS Policlinico MultiMedica, Genetics malovini, alberto; University of Pavia, Computer Engineering and Systems Science roncarati, roberta; IRCCS Policlinico MultiMedica, Genetics novelli, valeria; IRCCS Policlinico MultiMedica, Genetics villa, Francesco; IRCCS Policlinico MultiMedica, Genetics condorelli, Gianluigi; IRCCS Policlinico MultiMedica, Genetics bellazzi, riccardo; University of Pavia, Computer Engineering and Systems Science Puca, Annibale; IRCCS Policlinico MultiMedica, Genetics
Keyword:	Aging, Genomics



1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Title: Association of FOXO3A Locus with Extreme Longevity in the Southern Italian Centenarian Study

Authors: Chiara Viviani Anselmi, BSc^{a*}, Alberto Malovini, BSc^{ab*}, Roberta Roncarati, BSc^a, Valeria Novelli, BSc^a, Francesco Villa, BSc^a, Gianluigi Condorelli, MD, PhD^a, Riccardo Bellazzi, PhD^b, Annibale Alessandro Puca, MD^a

Affiliations: Unit of Genetics and Cardiovascular Research Institute, Istituto Ricovero Cura Carattere Scientifico "Multimedica", 20099 Sesto S.Giovanni, Italy; ^b Department of Computer Engineering and Systems Science, University of Pavia, via Ferrata 1, I-27100, Pavia, Italy.

Word count: [3859](#)

Deleted: ¶

Deleted: 3477

Formatted: Highlight

Address for corresponding author: Doctor Annibale Alessandro Puca, Unit of Genetics and Cardiovascular Research Institute, IRCCS Policlinico MultiMedica, Via Milanese 300, 20099 Sesto S.Giovanni, Italy; Tel: +39 0255406508; Fax: +39 02 55406570; E-mail: puca@longevita.org

Keywords: longevity; association; gender-stratification; FOXO3A.

Running title: FOXO3A and Extreme Longevity

FOOTNOTES:

these authors contributed equally to the work

R.R. and V.N. were supported by a fellowship of the Doctorate School of Molecular Medicine, Università degli

Studi di Milano.

Deleted: *

Formatted: Font: Arial Narrow, 12 pt

Formatted: Font: Arial Narrow, 12 pt, English (U.K.)

Formatted: Font: Arial Narrow, 12 pt

Formatted: Font: Arial Narrow, 12 pt, English (U.K.)

Formatted: Font: Arial Narrow, 12 pt

Formatted: Font: Arial Narrow, 12 pt, English (U.K.)

ABSTRACT

1
2
3
4
5
6 A number of potential candidate genes in a variety of biological pathways have been associated with longevity
7
8 in model organisms. Many of these genes have human homologues and thus have potential to provide
9
10 insights into human longevity. Recently, several studies suggested that FOXO3A functions as a key bridge of
11
12 various signaling pathways that influence aging and longevity. Interestingly, Willcox BJ and colleagues
13
14 identified several variants which displayed significant genotype-gender interaction **in male** human longevity. In
15
16 particular, a nested case-control study was performed in an ethnic Japanese population in Hawaii and five
17
18 candidate longevity genes were chosen based on links to the insulin-IGF1-signaling pathway. In the Willcox
19
20 study, the investigated genetic variations (rs2802292, rs2764264, and rs13217795) within the *FOXO3A* gene
21
22 were significantly associated with longevity in male centenarians.

Formatted: Highlight

Deleted: isolated Okinawa

23
24 We validated the association of FOXO3A polymorphisms with extreme longevity in males from the Southern
25
26 Italian Centenarian Study. Particularly, rs2802288, a proxy of rs2802292, showed the best allelic association
27
28 (MAF=0.49, p=0.003, OR=1.51, 95% CI: 1.15-1.98). Furthermore we undertook a meta-analysis in order to
29
30 explore the significance of rs2802292 association with longevity by combining the association results of the
31
32 current study and the findings coming from the Willcox BJ *et al.* investigation.

Deleted: of

33
34 Our data point to a key role of FOXO3A in human longevity and confirm the feasibility of the identification of
35
36 such genes with centenarian-controls studies. Moreover, we hypothesize the susceptibility to the longevity
37
38 phenotype may well be the result of complex interactions involving genes, environmental factors but also
39
40 gender.
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

INTRODUCTION

1
2
3
4
5
6
7
8 Aging is a multifactorial and complex process regulated by stochastic interactions (random damage to vital
9
10 molecules), extrinsic interventions (such as diet and caloric restriction) and intrinsic/genetic alterations. A
11 number of centenarian (people aged 100 years old or older) studies exploit the strong selection of favourable
12 genotypes in exceptionally aged individuals to study candidate genes and to perform genome wide analyses.
13
14 Replications of observed associations of genotypes with longevity, coupled with functional studies to define
15 mechanisms whereby specific genotypes influence lifespan associated phenotypes, are essential to delineate
16 true positive genetic findings. Such findings could potentially lead to preventive and therapeutic interventions
17 for several age-related diseases that cause significant morbidity and mortality among older people.
18
19 Numerous genes have been identified that are either positively or negatively selected in the centenarian
20 population as a consequence of a demographic selection ¹, but no consistent replications have been observed
21 in independent population, with the exception of *Apolipoprotein E* (APOE) ^{2,3}. In addition, animal studies have
22 provided insights into the types of genes that can be involved in the regulation of lifespan. The first longevity
23 mutant to be identified was the *C. Elegans* gene *PAX2 age-1* ⁴ that encodes phosphatidylinositol 3-kinase
24 (PI3K) ⁵, which has a key role in a signalling pathway that is homologous to the mammalian insulin– IGF1
25 (insulin-like growth factor 1) pathway. Genetic studies in *C. Elegans* have demonstrated that activation of the
26 PIK3/Akt pathway by IGF1 suppresses activity of DAF-16 forkhead transcription factor, the nematode ortholog
27 of mammalian FOXO proteins. FOXO transcription factors belong to the Forkhead family of proteins, a family
28 characterized by a conserved DNA binding domain (*Forkhead box*) for phosphorylation by the survival kinase
29 Akt ⁶. Particularly, in mammals the FOXO subgroup (FOXO1, FOXO3, FOXO4 e FOXO6) promote the
30 expression of numerous downstream genes that mediate stress resistance, innate immunity, metabolic
31 processes and toxin degradation ^{7,8}.

Deleted: s

Deleted: ortein

32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48 A moderate reduction in the intake of calories [also known as caloric restriction (CR)] is extremely effective in
49 delaying age-related decline and increasing longevity in organisms ranging from yeast to mammals ^{9,10}. In this

Deleted: nutrients

Deleted: aging

Deleted: At

regard, Bonkowski and colleagues observed a striking difference in the response of growth hormone (GH) receptor/GH-binding protein knockout (GHRKO) and normal mice to an identical regimen of CR ¹¹. When male and female longevity data were analyzed separately, it became evident that the effects of CR on maximal longevity were sexually dimorphic in GHRKO mice but not in normal mice.

To be noted, susceptibility to many common diseases, well as longevity, may well be the result of complex interactions involving genes, environmental factors and, intriguingly, gender.

In this scenario, *FOXO* is a key component of the insulin pathway, free radical production and human longevity.

In the prospective population based Leiden 85 plus Study (866 females and 379 males, aged 85 years or more), the effect of genetic variants in *FOXO1A* and *FOXO3A* were analyzed on metabolic profiles, fertility, fecundity, age-related diseases (i.e. diabetes, myocardial infarction, etc) and mortality ¹². In detail, *FOXO3A* haplotypes showed no association with longevity and other investigated variables. One possible drawback of this finding is total absence of gender stratification. In fact, the hormonal environment as well as tissue specific gene expression is known to differ significantly between genders in vertebrates. Hormones may differentially affect gene expression in somatic tissues, thus leading to the gender specific susceptibility to a complex phenotype, as longevity ¹¹. At this regard, Willcox BJ *et al.* ¹³ performed a nested case-control study of five candidate longevity genes (*ADIPOQ*, *FOXO1A*, *FOXO3A*, *SIRT1* and *COQ7*) with links to the insulin/IGF signaling pathway (IIS). This nested case-control study was conducted in an ethnic Japanese population from the Island of Oahu (Hawaii) with 213 male longevity “cases” and 402 male “average-lived” controls, and found three SNPs associated with the “longevity” phenotype in the locus of *FOXO3A* gene: rs2764264 (p=0.0002), rs13217795 (p=0.0006) and rs2802292 (p<0.0001) ¹³. The aim of our study was to support the potential role of *FOXO3A* as a key component that influence longevity with a case-control study. In this regard, we screened 480 long-living “cases” (281 males and 199 females) and 335 young controls (195 males, 140 females) from an isolated and homogeneous population of Southern Italy (Southern Italian Centenarian Study-SICS) with Illumina BeadChip 300K and we focused on *FOXO3A* gene. As previously described, we hypothesized that gender may influence the association between FOXO3A and longevity; thus we analyzed males and females

- Deleted: the
- Deleted: these
- Deleted: Probably, o
- Deleted: on isolated Okinawa population
- Deleted: Single Nucleotide Polymorphisms (
- Deleted:)
- Deleted: One possible drawback of this study is the lack of correction for population stratification that can bring to spurious associations (nominal type I errors). ¶
- Deleted: At
- Deleted: of
- Deleted: To avert any possible drawbacks of Willcox study,
- Deleted: However,
- Formatted: English (U.S.), Highlight
- Deleted: an additional factor to consider is the impact gender has on the longevity susceptibility. A
- Formatted: English (U.S.), Highlight
- Formatted: Highlight
- Formatted: English (U.S.), Highlight
- Formatted: Highlight
- Formatted: English (U.S.), Highlight
- Formatted: Highlight
- Formatted: English (U.S.), Highlight
- Formatted: Font: Italic, Highlight
- Formatted: English (U.S.), Highlight
- Deleted: the investigated phenotyp
- Formatted: Highlight
- Deleted: e
- Formatted: Highlight
- Formatted: English (U.S.), Highlight
- Formatted: Highlight
- Deleted: performed
- Deleted: the statistical analysis on
- Deleted: the only male samples
- Formatted: Highlight

Deleted: w

Formatted: Highlight

1
2 separately. We performed a powerful method for correcting for stratification of our population. Lastly, we
3
4 exploited a meta-analysis approach to pool Willcox's findings in order to find evidence for statistical
5
6 association for SNP rs2802292.
7
8
9

10 RESULTS

11 We recruited and phenotyped 281 male longevity "cases" (mean age 95 years) and 195 male "controls" (mean
12
13 age 33 years) as part of the SICS. Subsequently, we genotyped these subjects (Table 1) using the Illumina
14
15 300k SNPs mapping BeadChip.
16

Deleted: Single Nucleotide Polymorphisms
(

Deleted:)

17 After a preliminary phase of quality control (see methods section), we looked for evidence of population
18
19 heterogeneity on a set of 258 males centenarians (cases) and 178 males controls by applying the Principal
20
21 Component Analysis (PCA) as implemented in EigenSoft 2.0 package ¹⁴. The estimation of the over-dispersion
22
23 of trend test statistics (λ) was 1.05, decreasing to 1.027 after correction for the first 4 significant principal
24
25 components: this indicates a small or null confounding effect of population structure on our association results.
26
27 (Fig.1)

28
29 In order to test for associations between SNPs and the longevity phenotype, we compared allele and genotype
30
31 frequencies between cases and controls by means of allelic and genotypic association tests (Table 2, Table
32
33 S1). Results from all analyses were very similar for significantly associated SNPs whether or not adjustments
34
35 for population structure were applied (Table S2). None of the analyzed SNPs showed deviations from the
36
37 Hardy Weinberg Equilibrium (p -HWD<0.05) or significant differences in missing data fractions between cases
38
39 and controls (p <0.05). All SNPs showed a genotyping call rate > 98.8%.

40 We also performed logistic regression analyses assuming an additive model and by comparing homozygous
41
42 minor *versus* homozygous major alleles on a set of SNPs located on the *FOXO3A* locus (Table 2). In order to
43
44 confirm the null-influence of population structure on our results we also performed association tests correcting
45
46 along the top 4 principal components by means of logistic regression assuming an additive effect of allele
47
48 dosage on the same SNP set.
49

Deleted: Results from all analyses were very similar for significantly associated SNPs whether or not adjustments for population structure were applied (Table S2). None of the analyzed SNPs showed deviations from the Hardy Weinberg Equilibrium (p -HWD<0.05) or significant differences in missing data fractions between cases and controls (p <0.05). All SNPs showed a genotyping rate > 98.8%.

We first focused on rs2802288, showing Linkage Disequilibrium (LD) with rs2802292 ($r^2=1$) in the HapMap CEU data panel ¹⁵. Particularly, allelic association tests highlighted significant association for rs2802288 in males (MAF=0.49, $p=0.003$, OR=1.51, 95% CI: 1.15-1.98). Moreover, this association was also confirmed by logistic regression under an additive model ($p=0.0026$, OR=1.53, 95% CI: 1.16-2.03) and by comparing homozygous minor *versus* homozygous major alleles ($p=0.0023$, OR=2.44, 95% CI: 1.37-4.34). (Table 2)

Successively, we imputed missing SNPs by means of Impute ¹⁶ software to better represent the 6q21 region. Rather, interestingly, the imputed rs2802292 (MAF=0.47) that was previously indicated by Willcox BJ *et al.* ¹³ as the best association, showed $r^2 = 0.79$ with rs2802288 (MAF=0.49), similarly to that observed in the HapMap CEU panel ¹⁵. In fact, our imputation analysis confirmed that rs2802292, unrepresented on the BeadChip, *was responsible for the strongest association effect in this region* (MAF=0.468, $p=0.0028$; OR: 1.53; CI 95%: 1.16-2.03, allelic association test). Additionally, we confirmed this significant association by logistic regression under additive model ($p=0.0022$; OR: 1.57; CI 95%: 1.17-2.10) and by comparing homozygous minor *versus* homozygous major alleles ($p= 0.0019$; OR: 2.58; CI 95%: 1.42-4.70). (Table 2)

To evaluate a possible joint effect of FOXO3A gene and the environmental exposure (such as smoke, etc) or FOXO3A and the non -genetic attributes (such as clinical-cognitive variables), we performed a covariates analysis. No significant association of genotypes with cognitive variables (Barthel score, cognitive score, anxiety, depression) or clinical phenotypes (diabetes mellitus, glaucoma, high blood pressure, kidney diseases, osteoporosis, smoking history, cancer) was detected. Then, we can hypothesize that FOXO3A have an effect on longevity independently of the tested variables. Subsequently, we generated a LD plot with the software HaploView v4.0 ¹⁷. LD blocks are delineated by black lines and defined using the Four Gamete Rule, a variant on the algorithm described in previously published study ¹⁸, as implemented in HaploView software ¹⁹.

Using this analysis, multimarker association tests showed a slightly significant association ($p= 0.03$) for a 24kb haplotype (Fig 2, Block 1; Table 3) comprising 3 SNPs carrying the minor allele (A) for rs2802288, and defining boundaries of a locus affected by a potential causative mutation for the extreme longevity phenotype. As can be observed in Figure 2, the LD-plot defines a locus that includes the 5' untranslated regions and the first

Deleted: Hence,
Deleted: w

Deleted: what

Formatted: Font: Italic

Formatted: Highlight

Formatted: Font: Italic, Highlight

Formatted: Highlight

Formatted: Font: Italic, Highlight

Formatted: Highlight

Formatted: Font: Italic, Highlight

Formatted: Highlight

Deleted: N

Formatted: Highlight

Deleted: both,

Deleted: and

Deleted: risk factor

Formatted: Highlight

Formatted: Font: Italic, Highlight

Formatted: Highlight

Deleted: To be noted, our samples have not a uniform distribution in some variables requested for this statistical analysis.

Formatted: Highlight

Formatted: Highlight

Formatted: Highlight

Deleted: ¶

Deleted: it

coding region of the investigated gene. Subsequently, in order to confirm the gender-specificity of the FOXO3A association, we also studied 199 female longevity “cases” (mean age 98 years) and 140 female “controls” (mean age 31 years) drawn from the same population (88% power to detect the association). No differences in terms of allele frequencies between centenarians and controls have been detected for rs2802288 in female individuals (176 centenarians and 113 controls that pass the quality control filtering criteria) (MAF=0.48, $p=0.3975$, OR= 0.86, 95% CI: 0.62-1.21) (Table 2). The absence of association in the female gender has been further confirmed by logistic regression assuming an additive model ($p=0.4034$, OR=0.87, 95% CI: 0.62-1.21) and by comparing homozygous minor *versus* homozygous major alleles ($p=0.4048$, OR=0.75, 95% CI: 0.39-1.47) (Table 2).

Deleted: recruited

Deleted:

Deleted: .

Formatted: Not Highlight

Deleted: Table S3

Formatted: Highlight

Deleted: Table S3

Formatted: Highlight

Furthermore, we undertook a meta-analysis in order to explore the significance of the rs2802292 association with longevity by combining the association results of the current study and the findings coming from the Willcox BJ *et al.* investigation. We combined the logarithm of the OR estimates and 95% CI for homozygous minor *vs* homozygous major alleles from the current analysis (OR=2.44, 95% CI: 1.37-4.34) and from Willcox BJ *et al.* study (OR=2.75, 95% CI: 1.51-5.02) to obtain a summary OR using fixed²⁰ and Der Simonian and Laird random effects models²¹. Fixed effects models assume that the effect of a risk allele has the same value in each dataset while random effects models assume that the risk allele effects for each study vary around some overall average effect²². The inconsistency metric I^2 ²³ was 0%, while the Q statistic²⁴ was not statistically significant ($p=0.88$) indicating a null between-study heterogeneity evidence. Since no evidence of between-study heterogeneity have been proven, fixed effects²⁰ and Der Simonian and Laird random effects²¹ coincide, generating identical summary point estimates (OR= 2.66, 95% CI: 1.74-4.07, $p=0.0001$) (Fig 3).

Deleted: As shown in Table 4 t

Deleted: Table 4,

DISCUSSION

Research into ageing has been rejuvenated by the recent discovery of SNPs in candidate genes that can extend lifespan of laboratory model organisms. Thus, these polymorphisms can also delay, ameliorate or even abolish the impact of many ageing-related diseases, including cardiovascular disease, neurodegeneration and

1
2 cancer. FOXO transcription factors have been implicated in regulating different cellular functions, such as
3
4 differentiation, metabolism, proliferation, and survival ^{7,25}.

5
6 At this respect, in the Leiden Prospective Study the overall and individual haplotype frequencies were not

7
8 different between the elderly and young control group for the FOXO3A locus ¹². However, possible limitations
9
10 of the study is in the absence of gender stratification. In fact, due to the profound effect of this pathway on the
11
12 reproductive/hormonal system, that differs between males and females, it is possible that variants that affect
13
14 genes of this pathway are gender specific enriched/depleted as the populations age. Hence, variants that
15
16 influence longevity in males may not affect female longevity and vice-versa.

17
18 Recently, Willcox BJ and colleagues demonstrated an association with exceptional longevity and *FOXO3A*

19
20 gene in an ethnic Japanese population in Hawaii, focusing on male individuals only ¹³. Although there was no

21
22 population stratification, effect observed in this study a possible limitation in most. Stratification is the result of
23
24 imbalance of ethnic background between cases and controls. It is usually modest when the recruitment is

25
26 consistent among cases and controls, and for populations that didn't suffer of recent immigration. In the

27
28 present study, we extrapolated the FOXO3A locus from our genome wide association data on a part of the

29
30 SICS. Particularly, we genotyped (with BeadChip 300k, Illumina) 281 SIC male nonagenarians and 195 SIC

31
32 male controls. It was encouraging to report the association between rs2802292 at FOXO3A locus and

33
34 longevity phenotype was convincingly and independently confirmed in our SICS homozygous minor *versus*

35
36 homozygous major alleles ($p=0.0019$; OR: 2.58; CI 95%: 1.42-4.70). Moreover, to avoid a possible drawback

37
38 of population structure we demonstrated a small or null confounding effect of population structure on our

39
40 association results, thus excluding type I errors (Fig. 1, Table S2). Our aim was to replicate rs2802292 in male

41
42 longevity, and to explore the *FOXO3A* locus in males and females to generate eventual new candidate SNPs

43
44 for longevity and to clarify the gender interaction on the FOXO3A locus and longevity. Replication efforts

45
46 require adequate sample size and our male population was sufficient to achieve >90% power to detect an

47
48 allele with 40% allele frequency and a factor of 1.5 effect, as previously reported in Willcox paper for

49
50 rs2802292 ²⁶.

Formatted: Font: Not Italic

Deleted: consists

Deleted:

Deleted: isolated

Deleted: from

Deleted: Okinawa

Deleted: A possible limitation of this study was the lack of controlling for stratification.

Deleted: ,

Deleted: as for the Okinawa population. ¶

Deleted: G

Deleted: W

Deleted: A

Deleted: .

Meta-analysis represents a well-established method to summarize results and draw conclusions from different studies for a set of common hypothesis, therefore we applied this approach in order to combine our association results for the SNP rs2802292 with the findings coming from Willcox BJ *et al.* study and to achieve more statistical power. The null-significance of the Q-statistics ²⁴ ($p=0.88$) and the inconsistency metric ²³ value ($I^2=0\%$) confirmed the absence of between-study heterogeneity (Table 4). In absence of heterogeneity between the two datasets we obtained identical estimates of the ORs and the 95% CIs for homozygous minor vs homozygous major alleles from both fixed ²⁰ and random effects ²¹ models (OR= 2.66, 95% CI: 1.74-4.07, $p=0.0001$) (Fig 3).

Deleted: Table 4,

To be noted, this is the first convincing replication of an association with longevity after the APOE association ³.

As stressed by National Cancer Institute-National Human Genome Research Institute (NCI-NHGRI) Working Group on Replication in Association Studies ²⁷, validation of the postulated gene-investigated phenotype association in a population different from that of the previous study is of great value.

Deleted: s

It is also established that mitochondrial Sirtuin 3 interacts with and under specific cellular conditions regulates the activity of FOXO3A. Additionally, SIRT3 regulates a series of essential intracellular processes that defend the cell against multiple types of cellular damage, including oxidative damage ²⁸. A rise in the level of reactive oxygen species (ROS) has two important effects: it can damage proteins, lipids, and DNA, leading to cell death or it can trigger the activation of specific physiologic signalling pathway ²⁹. Moreover, a GH-IGF1 pathway has been implicated in determination of longevity in a variety of species ^{10,30}.

Formatted: Highlight

Genetic variants that are either positively or negatively selected as population ages impact on survival (demographic selection). In this regard, FOXO3A gene SNP rs2802292 should have a protective role.

Formatted: Font: Italic, Highlight

Formatted: Highlight

Deleted: ies

These observations imply that the FOXO3A gene SNP rs2802292, or SNPs in LD, have a protective role by partially increasing the ability of other proteins to activate FOXO3A, or by increasing FOXO3A activity on downstream targets. Re-sequencing of the FOXO3A locus will clarify the nature of the genetic variation that is tracked by rs2802292. It is possible that the variation does not need to affect coding genes, since coding genes from less than one-third of the evolutionary conserved genome ²⁶. Furthermore, future population

Deleted: being

studies on diseases of aging will clarify if *FOXO3A* associate across diseases, as for the longevity gene *APOE* and Alzheimer's / cardiovascular diseases.

Deleted: C

To be noted, from the data that we obtained in our population, the association is strictly linked with male gender. Taking in consideration these new evidences, we would suggest a gender specific analysis for the longevity phenotype.

Formatted: Highlight

Formatted: Highlight

Formatted: Highlight

Formatted: Highlight

Formatted: Highlight

To conclude, the discovery and replication of a convincing association of *FOXO3A* locus contributes to the hypothesis that the impact of the IGF-I/insulin pathway on longevity is a property that has been evolutionarily conserved throughout the animal kingdom.

MATERIALS and METHODS

Subject features

This case/control study was carried on as a part of the Southern Italian Centenarian Study (SICS). The SIC study (2002) has assembled a large cohort of 600 DNA samples of nonagenarians/centenarians (age range 90-109 years) and 800 DNA samples of young controls (age range 18-48 years) from a isolated region of Southern Italy east of Naples with a high prevalence of long lived and healthy and characterized by a high level of endogamy ³¹. All participants for the current study was well-characterized, in fact detailed phenotypic information about self reported health history, cognitive and physical functions measured through Blessed memory tests and Barthel scores of activity of daily living, as well as demographic and exposure to common risk factors such as alcohol and smoking were collected. For the aim of the present investigation, 281 long-lived men (age range 90-108 years) and 195 controls (age range 18-48 years) were genotyped (Table 1).

Deleted: 2

The subjects were enrolled by an Italian subsidiary (Associazione Longevità) of Elixir Pharmaceuticals (former Centagenetix Inc). All subjects donated blood samples for DNA study and gave written informed consent to the study. The study was conducted in accordance with the ethical principles that have their origins in the Declaration of Helsinki.

Genotyping: Illumina BeadChip

1
2 Genotyping was carried out using the Infinium II Assay-HumanHap BeadChip 317K-duo (Illumina) using
3 standard protocol. In fact, Illumina's Infinium II Whole-Genome Genotyping Assay is designed to interrogate a
4 large number of single nucleotide polymorphisms (SNPs) at unlimited levels of loci multiplexing. HumanHap
5
6 317-Duo workflow can be divided in three main segments: (1) sample preparation, (2) sample fragmentation
7
8 and hybridization, (3) extension, staining, and scanning.
9

10 The BeadChps are imaged using a two-color confocal laser system with 0.8- μ m resolution. The bead
11 intensities are extracted, and genotypes are calculated using an Illumina-supplied cluster file, which is based
12 on a set of reference samples. The normalization algorithm adjusts for nominal offset, cross-talk, and intensity
13 variations observed in the two-color channels. The data for each BeadChip is self-normalized using information
14 contained within the array. This normalization algorithm removes outliers, adjusts for channel-dependent
15 background and global intensity differences, and also scales the data. The X and Y color channels undergo an
16 affine coordinate transformation to make the data appear as canonical as possible with the homozygotes lying
17 along the transformed x- and y-axes. All genotypes will be evaluated using a quantitative quality score called
18 GenCall score. A GenCall score ranges from 0 to 1 and reflects the proximity within a cluster plot of the
19 intensities of that genotype to the centroid of the nearest cluster.
20
21
22
23
24
25
26
27
28
29

30 **Statistical analysis**

31
32 **Quality control:** Individuals showing genotyping rate <93% and SNPs with genotyping rate <95% have been
33 excluded from the dataset. We screened for contaminations and evaluated relationships among individuals
34 using Identity By Descent (IBD) estimations: samples showing extreme heterozygosis and related individuals
35 have been removed. X chromosome data have been used to check for the discordances in terms of gender
36 assignment. SNPs with MAF < 0.01 and markers deviating from the Hardy Weinberg Equilibrium (HWE) in the
37 control population (p -HWD < 0.001) have been excluded from the analysis.
38
39
40
41
42
43

44 **Population structure:** In order to assess the absence of population stratification we applied the Principal
45 Component Analysis (PCA) approach, a method that can capture subtle and extensive variations due to both
46 ethnical, experimental and technical features¹⁴. All autosomal SNPs and individuals that passed the quality
47 control filters (299772 SNPs, 258 centenarians, 178 controls) have been used as input to the EigenSoft 2.0¹⁴
48
49
50
51

1 software using the default parameters except for the number of outliers removal iterations that was set to 0 to
2 obtain estimate for all subjects.
3

4
5 **Imputation:** To impute SNPs from multimarker tags we used Impute v0.5.0¹⁶.
6

7 **Association tests:** We evaluated the association of markers with longevity phenotype by means of allelic
8 association tests (1 d.f.) and applied Pearson χ^2 test (2 d.f.) and Fisher exact test (if cell counts were < 5) in
9 order to compare genotype frequencies between cases and controls. We also performed logistic regression
10 under an additive model and logistic regression comparing homozygous minor vs homozygous major alleles.
11 Correction for population stratification have been performed using logistic regression under additive model
12 adjusting along the top 4 statistical significant principal components. Allele frequencies, genotype counts, OR
13 and 95% confidence intervals (95% CI) have been also estimated for each SNP.
14
15

16 **Meta-analysis:** We used the OR and 95% CI for homozygous minor vs homozygous major alleles from the
17 current analysis and from Willcox BJ *et al.* study to calculate the natural logarithms of the OR (Log-OR) and its
18 standard error (SE). The Log-OR estimates were combined to obtain a summary OR following fixed²⁰ and Der
19 Simonian and Laird random effects²¹ models using inverse variance calculations. Between dataset
20 heterogeneity was identified using the I^2 metric for inconsistency²³ and its statistical significance was
21 assessed by means of the χ^2 distributed Q statistic²⁴ (the number of d.f. is given by $k-1$, where k is the number
22 of analyzed datasets). All calculations have been performed using R v2.7.1 software.
23
24

25 **Haplotype analyses:** We used HaploView 4.0¹⁷ for blocks definition, haplotypes estimation and association
26 tests. We defined haplotype blocks using the Four Gamete Rule, a variant on the algorithm described in Wang
27 study¹⁸ and estimated haplotypes using the default accelerated EM algorithm¹⁷. Differences in haplotype
28 frequencies between centenarians and controls have been tested by means of χ^2 tests, excluding from the
29 analysis haplotypes with a frequency < 0.01.
30
31

32 **Covariate analyses:** Associations between clinical risk factors, cognitive variables with genotypes have been
33 tested among nonagenarians by means of logistic and linear regression using R v2.7.1 software
34 (<http://www.R-project.org>).
35
36

37 All statistical analyses have been performed using PLINK v1.04³² software except when specified.
38
39
40
41
42
43
44
45
46
47
48
49
50
51

AKNOWLEDGMENTS:

We thank all the patients and Medical Doctors of Cilento Health District ASL SA3 that participated to the study. We are grateful to Elixir Pharmaceuticals for access to the DNA samples. This work has been sponsored by Italian Ministry of Health (1% funds), , Italian Ministry of Research and University (FIRB [RBIN04X9XE](#)) and by the FIRB project ITALBIONET, Rete Italiana di Bioinformatica.

Formatted: English (U.K.)

REFERENCES:

- 1 Barzilai, N. et al., Unique lipoprotein phenotype and genotype associated with exceptional longevity. *JAMA* **290** (15), 2030 (2003); Perls, T. and Terry, D., Genetics of exceptional longevity. *Exp Gerontol* **38** (7), 725 (2003); Suh, Y. et al., Functionally significant insulin-like growth factor I receptor mutations in centenarians. *Proc Natl Acad Sci U S A* **105** (9), 3438 (2008).
- 2 Christensen, K., Johnson, T. E., and Vaupel, J. W., The quest for genetic determinants of human longevity: challenges and insights. *Nat Rev Genet* **7** (6), 436 (2006).
- 3 Novelli, V. et al., Lack of replication of genetic associations with human longevity. *Biogerontology* **9** (2), 85 (2008).
- 4 Friedman, D. B. and Johnson, T. E., A mutation in the age-1 gene in *Caenorhabditis elegans* lengthens life and reduces hermaphrodite fertility. *Genetics* **118** (1), 75 (1988).
- 5 Ogg, S. et al., The Fork head transcription factor DAF-16 transduces insulin-like metabolic and longevity signals in *C. elegans*. *Nature* **389** (6654), 994 (1997).
- 6 Kaestner, K. H., Knochel, W., and Martinez, D. E., Unified nomenclature for the winged helix/forkhead transcription factors. *Genes Dev* **14** (2), 142 (2000); McElwee, J., Bubb, K., and Thomas, J. H., Transcriptional outputs of the *Caenorhabditis elegans* forkhead protein DAF-16. *Aging Cell* **2** (2), 111 (2003); Murphy, C. T. et al., Genes that act downstream of DAF-16 to influence the lifespan of *Caenorhabditis elegans*. *Nature* **424** (6946), 277 (2003).
- 7 Brunet, A. et al., Akt promotes cell survival by phosphorylating and inhibiting a Forkhead transcription factor. *Cell* **96** (6), 857 (1999); Burgering, B. M. and Kops, G. J., Cell cycle and death control: long live Forkheads. *Trends Biochem Sci* **27** (7), 352 (2002); Kops, G. J. et al., Forkhead transcription factor FOXO3a protects quiescent cells from oxidative stress. *Nature* **419** (6904), 316 (2002); Kops, G. J. et al., Control of cell cycle exit and entry by protein kinase B-regulated forkhead transcription factors. *Mol Cell Biol* **22** (7), 2025 (2002).
- 8 Xuan, Z. and Zhang, M. Q., From worm to human: bioinformatics approaches to identify FOXO target genes. *Mech Ageing Dev* **126** (1), 209 (2005).
- 9 Guarente, L. and Picard, F., Calorie restriction--the SIR2 connection. *Cell* **120** (4), 473 (2005); Weindruch, R. and Sohal, R. S., Seminars in medicine of the Beth Israel Deaconess Medical Center. Caloric intake and aging. *N Engl J Med* **337** (14), 986 (1997).
- 10 Kenyon, C., The plasticity of aging: insights from long-lived mutants. *Cell* **120** (4), 449 (2005); Tatar, M., Bartke, A., and Antebi, A., The endocrine regulation of aging by insulin-like signals. *Science* **299** (5611), 1346 (2003).
- 11 Bonkowski, M. S. et al., Targeted disruption of growth hormone receptor interferes with the beneficial actions of calorie restriction. *Proc Natl Acad Sci U S A* **103** (20), 7901 (2006).

- 1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
- 12 Kuningas, M. et al., Haplotypes in the human Foxo1a and Foxo3a genes; impact on disease and mortality at old age. *Eur J Hum Genet* **15** (3), 294 (2007).
- 13 Willcox, B. J. et al., FOXO3A genotype is strongly associated with human longevity. *Proc Natl Acad Sci U S A* (2008).
- 14 Price, A. L. et al., Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* **38** (8), 904 (2006).
- 15 A haplotype map of the human genome. *Nature* **437** (7063), 1299 (2005).
- 16 Marchini, J. et al., A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat Genet* **39** (7), 906 (2007).
- 17 Barrett, J. C., Fry, B., Maller, J., and Daly, M. J., Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* **21** (2), 263 (2005).
- 18 Wang, N. et al., Distribution of recombination crossovers and the origin of haplotype blocks: the interplay of population history, recombination, and mutation. *Am J Hum Genet* **71** (5), 1227 (2002).
- 19 Gabriel, S. B. et al., The structure of haplotype blocks in the human genome. *Science* **296** (5576), 2225 (2002).
- 20 Pettiti, D., *Meta-analysis, decision analysis, and cost-effectiveness analysis*. (1999).
- 21 Fleiss, J.L., The statistical basis of meta-analysis. *. Stat Methods Med Res* **2**: 121–45. (1993).
- 22 Evangelou, E., Maraganore, D. M., and Ioannidis, J. P., Meta-analysis in genome-wide association datasets: strategies and application in Parkinson disease. *PLoS ONE* **2** (2), e196 (2007).
- 23 Higgins, J. P., Thompson, S. G., Deeks, J. J., and Altman, D. G., Measuring inconsistency in meta-analyses. *BMJ* **327** (7414), 557 (2003).
- 24 Lau, J., Ioannidis, J. P., and Schmid, C. H., Quantitative synthesis in systematic reviews. *Ann Intern Med* **127** (9), 820 (1997).
- 25 Brunet, A., [The multiple roles of FOXO transcription factors]. *Med Sci (Paris)* **20** (10), 856 (2004); Brunet, A. et al., Stress-dependent regulation of FOXO transcription factors by the SIRT1 deacetylase. *Science* **303** (5666), 2011 (2004); Skurk, C. et al., The Akt-regulated forkhead transcription factor FOXO3a controls endothelial cell viability through modulation of the caspase-8 inhibitor FLIP. *J Biol Chem* **279** (2), 1513 (2004).
- 26 Altshuler, D., Daly, M. J., and Lander, E. S., Genetic mapping in human disease. *Science* **322** (5903), 881 (2008).
- 27 Chanock, S. J. et al., Replicating genotype-phenotype associations. *Nature* **447** (7145), 655 (2007).
- 28 Okawara, M. et al., Resveratrol protects dopaminergic neurons in midbrain slice culture from multiple insults. *Biochem Pharmacol* **73** (4), 550 (2007).
- 29 Houstis, N., Rosen, E. D., and Lander, E. S., Reactive oxygen species have a causal role in multiple forms of insulin resistance. *Nature* **440** (7086), 944 (2006).
- 30 Guarente, L. and Kenyon, C., Genetic pathways that regulate ageing in model organisms. *Nature* **408** (6809), 255 (2000).
- 31 Colonna, V. et al., Campora: a young genetic isolate in South Italy. *Hum Hered* **64** (2), 123 (2007).
- 32 Purcell, S. et al., PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* **81** (3), 559 (2007).

46 **TABLE TITLE:**

47
48
49 **Table 1 : Age (years) distribution by gender and by centenarian/control phenotype**

Deleted: male

Table 2: FOXO3A locus in long-living individuals and controls.

Table 3: Haplotype association test of block 1 using HaploView.

Deleted: Table 4: Meta-analysis for rs2802292 combining Okinawa study and SICS.[¶]

FIGURE LEGENDS:

Figure 1. Population structure of centenarians (red) and control subjects (black). Each scatter plot shows the first two principal components that were estimated using genotype data for more than 300K SNPs in centenarians and control subjects using the program EigenSoft 2.0. The two populations are ethnically identical as shown in the right scatter plot.

Figure 2. Pattern of LD of FOXO3A region displayed using D' with HaploView. LD displays were generated using the D' colour scheme. The different shades of red indicate different D' value ($0 < D' < 1$). Haplotype blocks were generated using the Four Gamete Rule as implemented in HaploView software. Boxed in green is rs2802288, boxed in blue are the SNPs explored by Willcox BJ *et al.*

Figure 3. Meta-analysis of rs2802292 variant in the current investigation and in Willcox study.

Each study is represented by the estimate of the homozygous minor vs homozygous major alleles OR and its 95% CI (circle). Summary effects by random effects (rhombus) and fixed effects (rectangle) calculations are also shown.

Table 1 : Age (years) distribution by gender and by centenarian/control phenotype

Phenotype	Males			Females		
	n	Mean \pm SD	Min-Max	n	Mean \pm SD	Min-Max
Centenarians	281	94.96 \pm 2.97	90-108	199	98.05 \pm 3.25	90-109
Controls	195	33.53 \pm 7.33	18-48	140	31.39 \pm 7.31	18-48

Footnote: n= number of individuals; SD= standard deviation; Min= minimum age; M= Maximum age

Table 2: FOXO3A locus in long-living individuals and controls

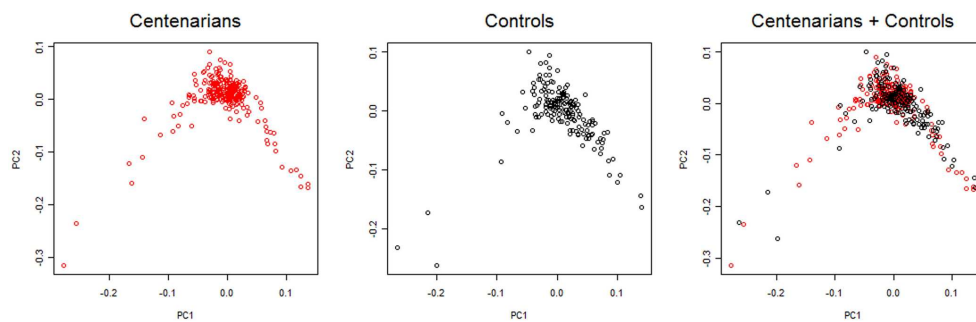
SNP	gender	MAF			Allelic		Additive		Homozygous	
		Overall	Cases	Controls	OR (95%CI)	p	OR (95%CI)	p	OR (95% CI)	p
rs9486902	M	0.2	0.22	0.18	1.27(0.90-1.80)	0.1659	1.30 (0.91-1.85)	0.1524	0.92 (0.30-2.80)	0.8786
	F	0.23	0.21	0.27	0.69 (0.47-1.02)	0.0634	0.67 (0.45-1.01)	0.0553	0.52 (0.16-1.70)	0.2800
rs2802288	M	0.49	0.53	0.43	1.51 (1.15-1.98)	0.003	1.53 (1.16-2.03)	0.0026	2.44 (1.38-4.34)	0.0023
	F	0.48	0.47	0.5	0.86 (0.62-1.21)	0.3975	0.87 (0.62-1.21)	0.4034	0.75 (0.39-1.47)	0.4048
rs10499051	M	0.12	0.14	0.1	1.57 (1.02-2.41)	0.04	1.56 (1.01-2.41)	0.0409	NA	0.9982
	F	0.12	0.11	0.12	0.95 (0.57-1.60)	0.8497	0.95 (0.56-1.61)	0.8474	0.32 (0.03-3.62)	0.3598
rs2802292*†	M	0.47	0.51	0.41	1.53 (1.16-2.03)	0.0028	1.57 (1.17-2.10)	0.0022	2.58 (1.42-4.70)	0.0019
	F	NA	NA	NA	NA	NA	NA	NA	NA	NA
rs13220810	M	0.2	0.18	0.22	0.79 (0.563-1.11)	0.1682	0.79 (0.56-1.10)	0.1653	0.80 (0.29-2.20)	0.6619
	F	0.22	0.23	0.2	1.17 (0.78-1.77)	0.4482	1.17 (0.78-1.78)	0.4471	1.12 (0.35-3.57)	0.8441
rs2764264†	M	0.4	0.43	0.36	1.37 (1.04-1.81)	0.0267	1.40 (1.05-1.86)	0.0227	1.99 (1.07-3.70)	0.0285
	F	0.41	0.4	0.42	0.93 (0.66-1.30)	0.6575	0.92 (0.65-1.31)	0.6553	1.03 (0.49-2.18)	0.9312
rs7341233	M	0.17	0.17	0.15	1.16 (0.80-1.67)	0.4364	1.15 (0.80-1.64)	0.4512	2.14 (0.67-6.79)	0.1963
	F	0.14	0.16	0.12	1.51 (0.92-2.48)	0.1047	1.51 (0.92-2.49)	0.1070	3.61 (0.41-31.41)	0.2454
rs17598747	M	0.15	0.15	0.15	1.05 (0.71-1.53)	0.8172	1.04 (0.72-1.50)	0.8237	1.72 (0.53-5.60)	0.3683
	F	0.12	0.13	0.1	1.44 (0.84-2.46)	0.1825	1.43 (0.84-2.43)	0.1911	2.79 (0.31-25.35)	0.3626
rs9285397	M	0.22	0.24	0.2	1.29 (0.93-1.80)	0.1274	1.34 (0.94-1.90)	0.1054	1.76 (0.53-5.88)	0.3561
	F	0.26	0.24	0.31	0.71 (0.49-1.03)	0.0699	0.69 (0.47-1.02)	0.0637	0.71 (0.26-1.99)	0.5190
rs12202209	M	0.15	0.15	0.15	1.07 (0.73-1.56)	0.739	1.06 (0.74-1.53)	0.7479	1.74 (0.53-5.60)	0.359
	F	0.12	0.13	0.1	1.44 (0.84-2.46)	0.1825	1.43 (0.84-2.43)	0.1911	2.79 (0.31-25.35)	0.3626
rs9480866	M	0.21	0.22	0.19	1.24 (0.87 1.75)	0.206	1.27 (0.89-1.80)	0.1881	1.71 (0.52-5.71)	0.3781
	F	0.24	0.21	0.29	0.67 (0.46-0.99)	0.0412	0.65 (0.44-0.97)	0.0370	0.51 (0.17-1.53)	0.2279
rs12207868	M	0.15	0.15	0.15	1.06 (0.72 1.54)	0.7751	1.05 (0.73-1.52)	0.7829	1.72 (0.53-5.60)	0.3683
	F	0.12	0.13	0.1	1.43 (0.84-2.44)	0.1905	1.42 (0.83-2.41)	0.1992	2.77 (0.30-25.16)	0.3662
rs13217795*†	M	0.35	0.38	0.3	1.42 (1.06 1.90)	0.0176	1.44 (1.07-1.945)	0.0159	2.30 (1.14-4.64)	0.0196
	F	NA	NA	NA	NA	NA	NA	NA	NA	NA
rs2153960	F	0.4	0.38	0.42	0.83 (0.59-1.18)	0.3009	0.83 (0.59-1.18)	0.2979	0.83 (0.39-1.73)	0.6107
	M	0.38	0.41	0.34	1.31 (0.99 1.74)	0.0578	1.33 (1-1.78)	0.0521	1.87 (0.99-3.53)	0.0542

rs3800229	M	0.36	0.38	0.33	1.25 (0.94 1.67)	0.1187	1.26 (0.94-1.68)	0.1162	1.65 (0.87-3.12)	0.1222
	F	0.38	0.36	0.41	0.81 (0.57-1.14)	0.2285	0.81 (0.57-1.15)	0.2310	0.75 (0.36-1.58)	0.4478
rs1935949	M	0.36	0.38	0.33	1.25 (0.94 1.67)	0.118	1.26 (0.95-1.68)	0.1143	1.65 (0.87-3.12)	0.1222
	F	0.38	0.36	0.41	0.81 (0.57-1.14)	0.2200	0.81 (0.57-1.14)	0.2222	0.77 (0.37-1.62)	0.4911

Footnotes: Gender = Gender-specific analysis (M = males, F = females), MAF Overall = minor allele frequency based on whole sample, MAF Cases = minor allele frequency in centenarians, MAF Controls = minor allele frequency in the control population; Odds Ratio (OR), 95% confidence intervals (95% C.I.) and p-value (p) respectively for allelic association test (Allelic), logistic regression under additive model (Additive) and logistic regression comparing homozygous minor (mm) vs homozygous major alleles (MM) (Homozygous); * = imputed SNPs; † = SNPs explored by Willcox BJ *et al.*

Table 3: Haplotype association test of block 1 using HaploView.

Block 1	Gender	Frequency	Freq Cases	Freq Controls	p-value
CGA	M	0.51	0.48	0.57	0.003
	F	0.51	0.53	0.49	0.3683
TAA	M	0.2	0.21	0.17	0.1588
	F	0.23	0.21	0.27	0.0634
CAA	M	0.17	0.18	0.16	0.4968
	F	0.13	0.15	0.11	0.2556
CAG	M	0.12	0.14	0.09	0.0381
	F	0.12	0.11	0.12	0.8689

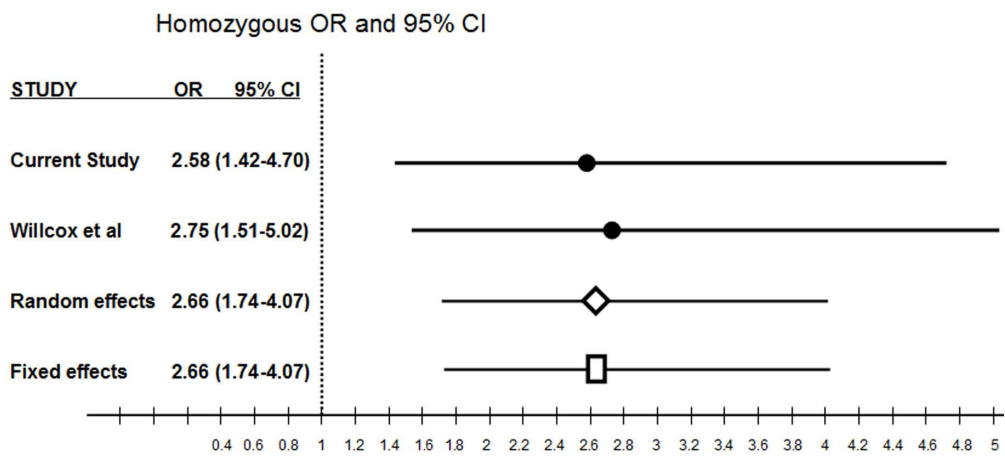


180x60mm (600 x 600 DPI)

Pre Peer Review

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60



179x87mm (600 x 600 DPI)

Peer Review