Cell Stem Cell, Volume *18*

Supplemental Information

A Global Social Media Survey of Attitudes to Human Genome Editing

Tristan McCaughey, Paul G. Sanfilippo, George E.C. Gooden, David M. Budden, Li Fan, Eva Fenwick, Gwyneth Rees, Casimir MacGregor, Lei Si, Christine Chen, Helena Hai Liang, Timothy Baldwin, Alice Pébay, and Alex W. Hewitt

Supplemental Methods:

Ethics Approval:

Ethics approval for this research was obtained from the Monash University Human Research Ethics Committee (MUHREC - CF15/1606). Prior to commencing the questionnaire, participants were directed to a landing page, which in a neutral manner, outlined the purpose of the survey and introduced the prospect of human genome engineering.

Questionnaire Development and Administration:

A questionnaire comprising 17 core items was designed. The initial section of the questionnaire recorded de-identified demographic details that were hypothesised to affect participants' attitudes towards gene editing. These included religious affiliation, country of residency, ethnicity, education, financial status and whether participants were personally affected by genetic disease. Given the scope and international exposure of our survey, financial status was broadly defined by a self-reported financial standing of being below, average or above-average wealth. The second section asked participants' opinion on the application of gene editing using a five-point Likert scale from "strongly agree" to "strongly disagree", with an additional option of 'I don't know'. Three "yes/no" questions appeared if participants 'agreed' or 'strongly agreed' with the application of gene editing technology on non-disease traits. The subject opinion questions were accompanied with a brief explanatory paragraph defining key terms such as 'life threatening', 'debilitating' and 'embryo', whilst also briefly outlining the heritable nature of embryonic gene editing.

Subject opinion question items

Question		Answer options
1	How much do you agree with the use of genetic editing of cells in **children or adults** to cure a **life threatening disease**? This means the disease could still be passed on to their children.	Strongly Agree, Agree, Neutral, Disagree, Strongly Disagree, I don't Know
2	How much do you agree with the use of genetic editing of cells in **children or adults** to cure a **debilitating disease**? This means the disease could still be passed on to their children.	Strongly Agree, Agree, Neutral, Disagree, Strongly Disagree, I don't Know
3	How much do you agree with the use of genetic editing of cells in **embryos** to prevent a **life threatening disease**? This means that all future generations would not have the disease.	Strongly Agree, Agree, Neutral, Disagree, Strongly Disagree, I don't Know
4	How much do you agree with the use of genetic editing of cells in **embryos** to prevent a **debilitating disease**? This means that all future generations would not have the disease.	Strongly Agree, Agree, Neutral, Disagree, Strongly Disagree, I don't Know

Question		Answer options
5	How much do you agree with the use of genetic editing of cells in **embryos** to alter any **non-disease characteristic** - such as memory, eye colour or height? This would mean that all subsequent generations would have the same genetic characteristics.	Strongly Agree, Agree, Neutral, Disagree, Strongly Disagree, I don't Know
5.1*	If you could safely genetically edit your embryo would you use this technology to determine **physical appearance** (eye colour; hair colour; skin colour)?	Yes, No
5.2*	If you could safely genetically edit your embryo would you use this technology to determine **intelligence**?	Yes, No
5.3*	If you could safely genetically edit your embryo would you use this technology to determine **strength or sporting ability**?	Yes, No
5.4*	What other non-health related traits would you edit?	*[Free text response]*
6	How much do you agree with the use of **genetically modified food**?	Strongly Agree, Agree, Neutral, Disagree, Strongly Disagree, I don't Know
7	We are also interested in understanding the reasons for your answers. Please describe the factors that have **influenced your response** or attitude towards human gene engineering.	*[Free text response]*

**Question only appeared to respondents who agreed to the previous question*

A thinking-aloud cognitive phase testing was undertaken in English speaking people to refine the questionnaire (Collins, 2003). Ten participants were interviewed to assess their interpretation and understanding of the questions. Subsequently, changes were iteratively made to two of the application questions, and to the explanatory paragraph, with no additional changes required after the initial four interviews. The final English questionnaire was then translated into Arabic, Chinese, French, German, Hindi, Japanese, Portuguese, Russian, Spanish and Turkish. To ensure consistency across the questionnaires, a separate translator then back translated each language with changes made as appropriate. The questionnaires were then formatted and placed onto an online platform and completed by a third translator prior to dissemination. Readability analysis found that the English questionnaire had an overall Gunning Fog Score of 13.8 (Friedman and Hoffman-Goetz, 2006; Gunning, 1968); and all characters in the Chinese version are in the most commonly used form of modern simplified Chinese.

In June 2015, the questionnaire was launched online http://humanediting.org. At this time we began advertising on social media through Facebook, Twitter and Google. A separate questionnaire was formatted for WeChat to enable dissemination in China, where Twitter, Facebook and Google promotions are restricted.

Data Preparation and Text Cleaning:

Survey responses were collated before being back-translated to English. Data preparation and analyses were performed in the R statistical environment (version 3.1.2) and Python (version 3.4.3). All data and scripts are available at https://github.com/hewittlab/Gene_Edit_Survey.

We used descriptive statistics to summarize demographic features of respondents. Given the likelihood of spurious responses, participants with extremes in reported age, below 10 above 90 years, were dropped (n=141). We utilised multinomial logistic regression to analyse the relationship between each of the demographic variables and subject opinion responses. For dependent variables, we collapsed the Likert scale to a three-point scale ('disagree', 'neutral', 'agree') initially merging the 'neutral' and 'I don't know' response categories. In this way, the multinomial model allows subject opinion to be analysed as a series of binary comparisons ('agree' is the reference category): i.e. agree versus neutral and agree versus disagree. Additional analyses without merging 'neutral' and 'I don't know' were also performed.

For each participant in the survey, where possible, we retrieved geolocation data. IP addresses were queried using a function in R to return location data (country, region/state, longitude and latitude, etc.) from freegeoip.net, a freely available HTTP application-programming interface. Geolocate data were not available for 1,362 participants. Duplicate IP addresses were reviewed and there were a total of 10,564 unique sites. Countries were further categorised into two groups (Advanced Economies and Emerging Market, Developing Economies) based on the International Monetary Fund's (IMF) World Economic Outlook Database, April 2015. Countries that are not members of the IMF were categorised based on GDP per capita data sourced from the World Bank (Web resources).

Data Analyses: Logistic Regression

To determine the effect of the demographic variables on participant attitudes towards gene editing, a series of 11 hierarchical models were fitted to the raw data and the significance of each assessed. The full model consisted of ten demographic variables (sex, age, self-reported ethnicity, GDP per capita, prior knowledge of gene editing, education level, religion, wealth, healthcare worker and personal or family history of a genetic condition) with each nested sub-model dropping one variable (i.e. personal or family history of a genetic condition eliminated from the first sub-model, personal or family history of a genetic condition and health care worker eliminated from the second, etc.). Each of the nested sub-models was then compared to the preceding model by χ^2 tests. Twice the difference between log-likelihood values of the nested and preceding models is asymptotically distributed as χ^2 with degrees of freedom (df) equal to the difference in parameters being estimated. The Akaike information criterion (AIC) was used to determine the best-fitting model by evaluating model parsimony (i.e., the best goodness-of-fit combined with the fewest latent variables). The model with the lowest AIC suggests the best fit. The variance inflation factor was calculated for the final regression model and did not show significant multicollinearity.

Multiple regression models used for determining reason for agreement in utility of human editing.

Model	Covariates Included
1	-
2	sex
3	sex + age
4	sex + age + ethnicity
5	sex + age + ethnicity + GDP
6	sex + age + ethnicity + GDP + heard_about
7	sex + age + ethnicity + GDP + heard_about + edu_level
8	sex + age + ethnicity + GDP + heard_about + edu_level + religion_type
9	sex + age + ethnicity + GDP + heard_about + edu_level + religion_type + wealth
10	sex + age + ethnicity + GDP + heard_about + edu_level + religion_type + wealth + worked_health
11	sex + age + ethnicity + GDP + heard_about + edu_level + religion_type + wealth + worked_health + genetic_cond

Abbreviations: GDP, Gross Domestic Product per capita in U.S dollars; heard_about, Previously heard about human genetic engineering or gene editing; edu_level, highest level of education; religion_type, religious affiliation; worked_health, previous occupation in a health or medical related field; genetic_cond, personnel or family history of an inherited medical condition.

Supplementary References:

Collins, D. (2003). Pretesting survey instruments: an overview of cognitive methods. Qual Life Res *12*, 229-238.

Friedman, D.B., and Hoffman-Goetz, L. (2006). A systematic review of readability and comprehension instruments used for print and web-based cancer information. Health Educ Behav *33*, 352-373.

Gunning, R.C. (1968). The technique of clear writing, 2 edn (New York: McGraw-Hill).

Web resources:

International Monetary Fund http://www.imf.org *Data accessed July 22, 2015*

World Bank: http://data.worldbank.org *Data accessed July 22, 2015*

<u>Supplementary Table 1.</u>

Summary odds ratios (OR) from the most parsimonious multiple regression models used for determining reason for agreement in the utility of human editing. The table displays the point estimates and 95% confidence intervals. Full model details are presented in Supplementary Methods. The OR compares the relative odds of the occurrence of the outcome (*e.g.* a participant's view on genetically modified food), given 'exposure' to the independent variables in the left column. In this way it is possible to determine whether an exposure (*e.g.* age) is associated with a particular outcome and the strength of that association. The OR can be interpreted as a % change in the odds of the outcome per unit change in the independent variable. For continuous variables (age, GDP) the unit is specified, and for categorical variables the change in odds of the outcome for the stated category of the variable is compared against a 'Reference Category,' which was the alternate response option.

Variable		**Adult/Child Life Threatening Disease**	**Adult/Child Debilitating Disease**	**Embryo Life Threatening Disease**	**Embryo Debilitating Disease**	**Embryo Non-Health Purposes**	**Genetically Modified Food**
	Sex (Male)	0.780 [0.662-0.918]**	0.778 [0.664-0.912]**	0.814 [0.699-0.949]**	0.773 [0.665-0.899]***	0.482 [0.424-0.547]***	0.413 [0.363-0.470]***
	Age#	1.057 [1.002-1.115]*	-	1.063 [1.011-1. 118]*	-	1.093 [1.048-1.141]***	1.339 [1.282-1.399]***
	Asian vs. Caucasian	1.615 [1.300-2.005]***	1.822 [1.476-2.250]***	1.718 [1.405-2.100]***	1.702 [1.396-2.076]***	0.375 [0.319-0.442]***	1.364 [1.153-1.614]***
	GDP per Capita^	0.987 [0.982-0.991]***	0.988 [0.984-0.993]***	0.990 [0.986-0.994]***	0.991 [0.987-0.995]***	1.015 [1.012-1.019]***	0.985 [0.982-0.989]***
Prior Knowledge	Heard a little	0.801 [0.645-0.996]*	0.749 [0.608-0.924]**	-	-	1.351 [1.139-1.603]***	-
	Heard a lot	0.777 [0.604-0.999]*	0.685 [0.536-0.874]**	-	-	-	0.616 [0.508-0.747]***
Education level	Attended School	0.508 [0.274-0.945]*	-	0.327 [0.184-0.581]***	0.310 [0.176-0.547]***	-	-
	Tertiary educated	0.342 [0.183-0.639]***	0.373 [0.203-0.685]**	0.305 [0.171-0.545]***	0.321 [0.181-0.569]***	-	-
Religion	Any Religion	1.325 [1.132-1.550]***	1.348 [1.154-1.574]***	1.521 [1.307-1.770]***	1.443 [1.239-1.680]***	1.554 [1.371-1.761]***	1.452 [1.275-1.653]***
	Christian	1.440 [1.176-1.764]***	1.523 [1.249-1.855]***	1.682 [1.391-2.035]***	1.687 [1.402-2.030]***	1.768 [1.502-2.082]***	1.430 [1.220-1.675]***
	Muslim	-	-	-	-	1.712 [1.406-2.085]***	1.871 [1.521-2.302]***
	Worked in HealthCare	-	-	1.356 [1.138-1.615]***	1.317 [1.108-1.565]**	-	-

***P<0.001 **P<0.01 *P<0.05

Percentage change with each decade older

^Percentage change with each 1000 USD increase in GDP per capita

N.b.: Only statistically significant results are shown ('-' indicates the OR was not significant and the confidence interval includes 1).