Genetic Data in Add Health

Brett C. Haberstick

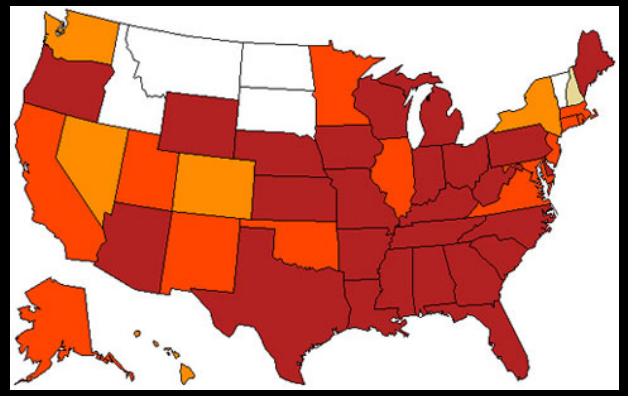
University of Colorado at Boulder

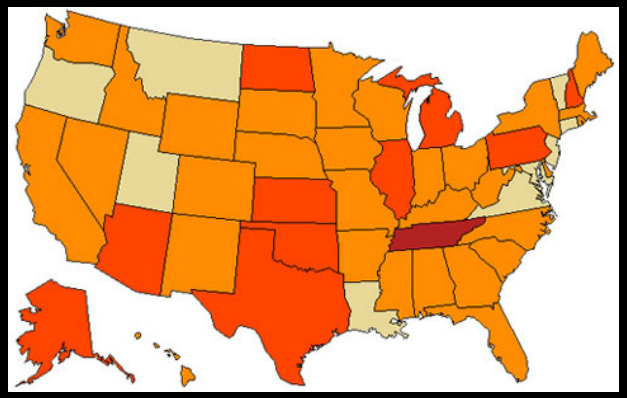
Obesity

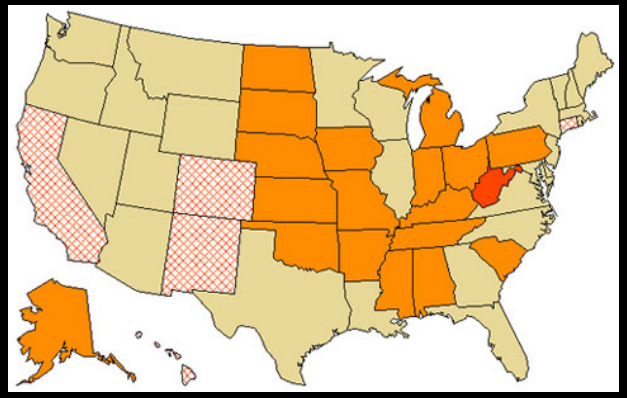
Blacks in America have the highest rate. Hispanic females have the highest rate of any ethnicity.

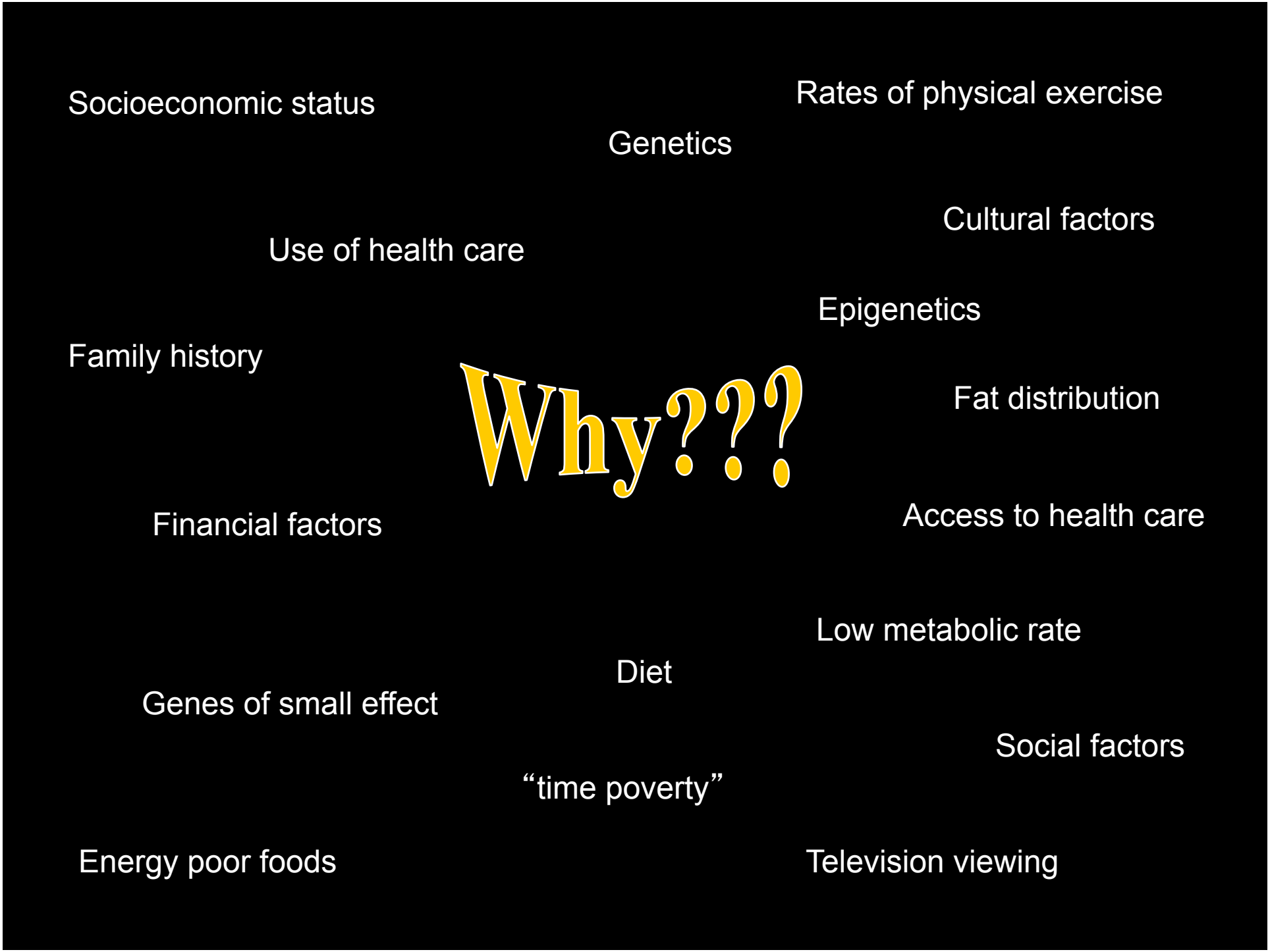
African-American have a higher likelihood of developing cardiovascular disease.

Type II Diabetes is highest in Hispanic females.

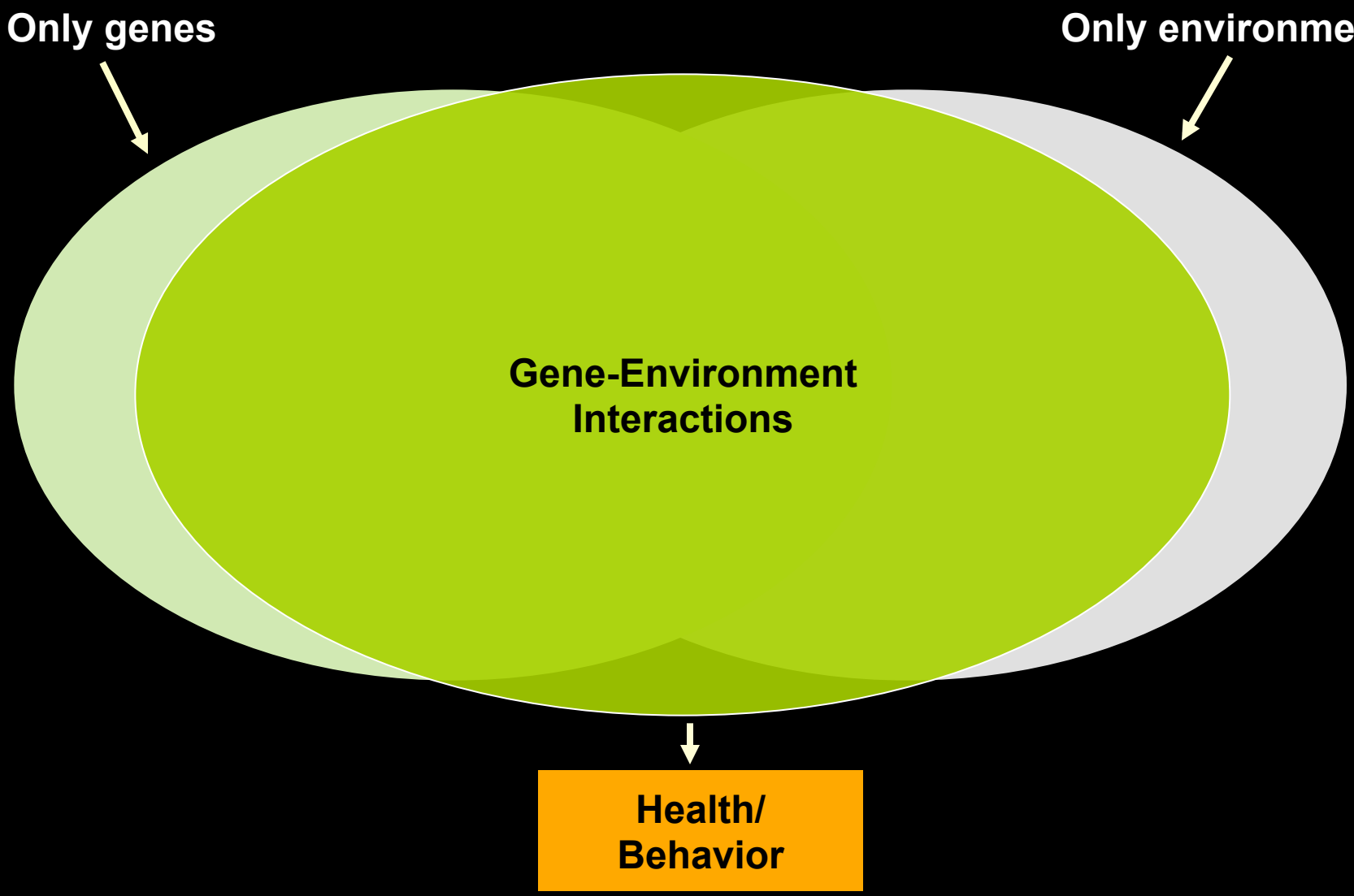






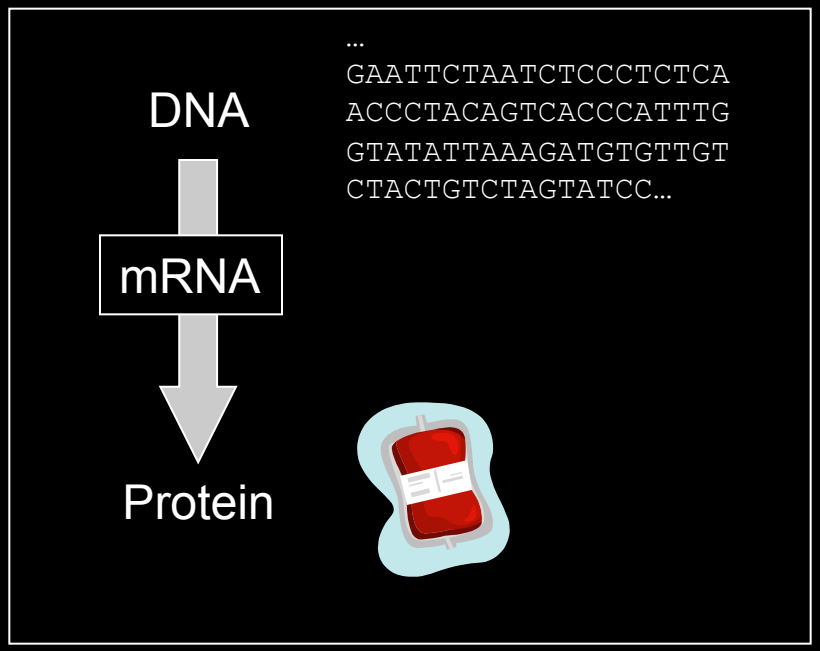

Socioeconomic status
Rates of physical exercise
Genetics
Cultural factors
Use of health care
Epigenetics
Family history
Why???
Fat distribution
Financial factors
Access to health care
Low metabolic rate
Diet
Genes of small effect
Social factors
“time poverty”
Energy poor foods
Television viewing

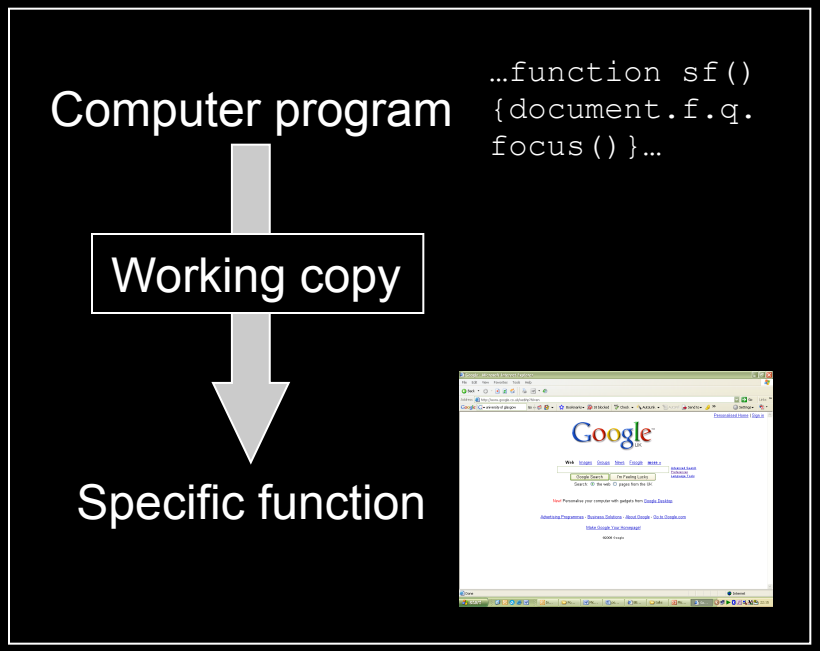
Risk Factors for Health/Behavior: New models



What is a gene?

Gene: is a sequence of DNA (string of letters) that is involved in the synthesis of functional proteins and corresponds to a unit of inheritance





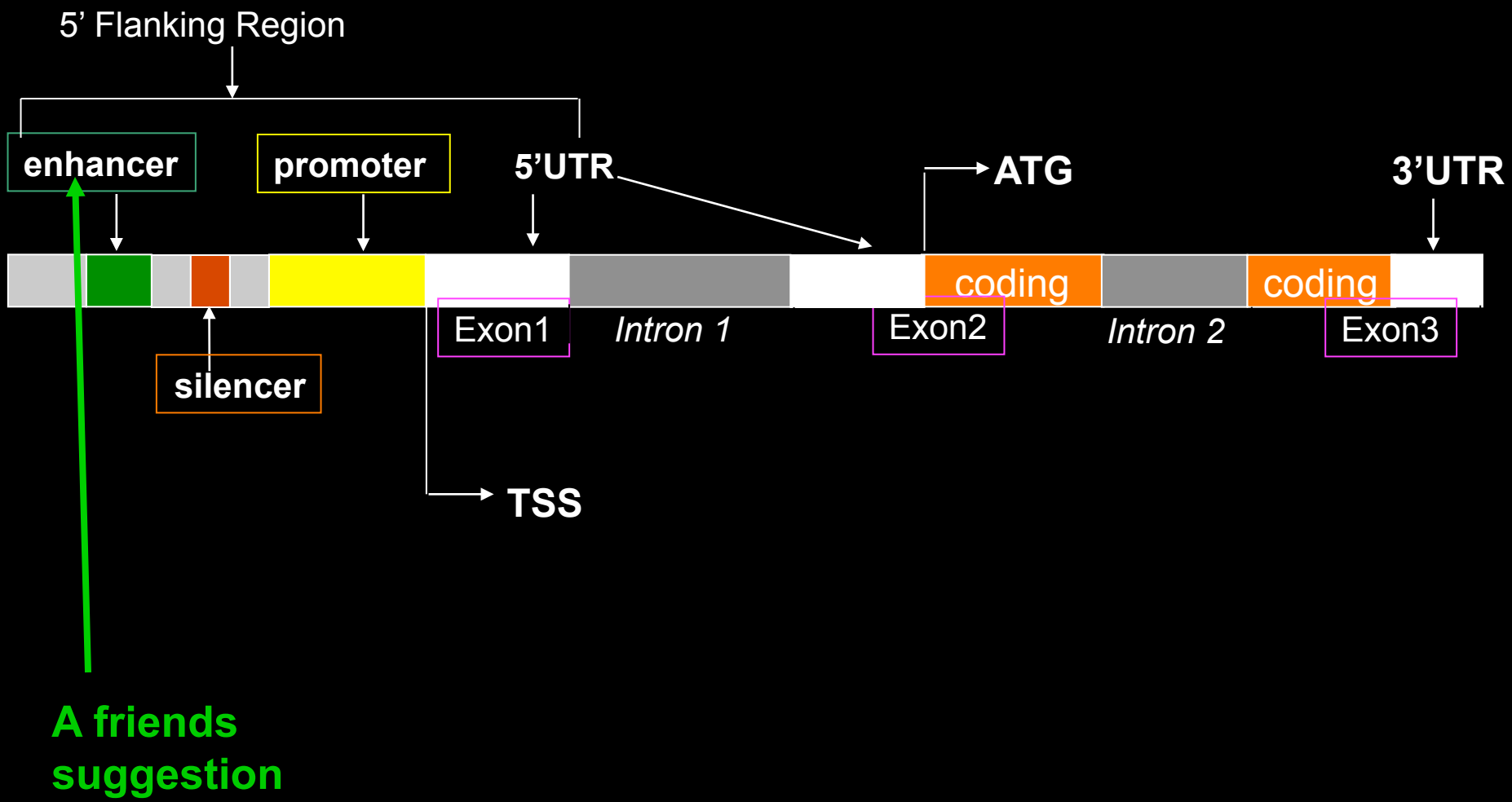
What is a gene?

~ Roughly 23,333 genes in human genome (RefSeq) – more than a chicken, less than a grape

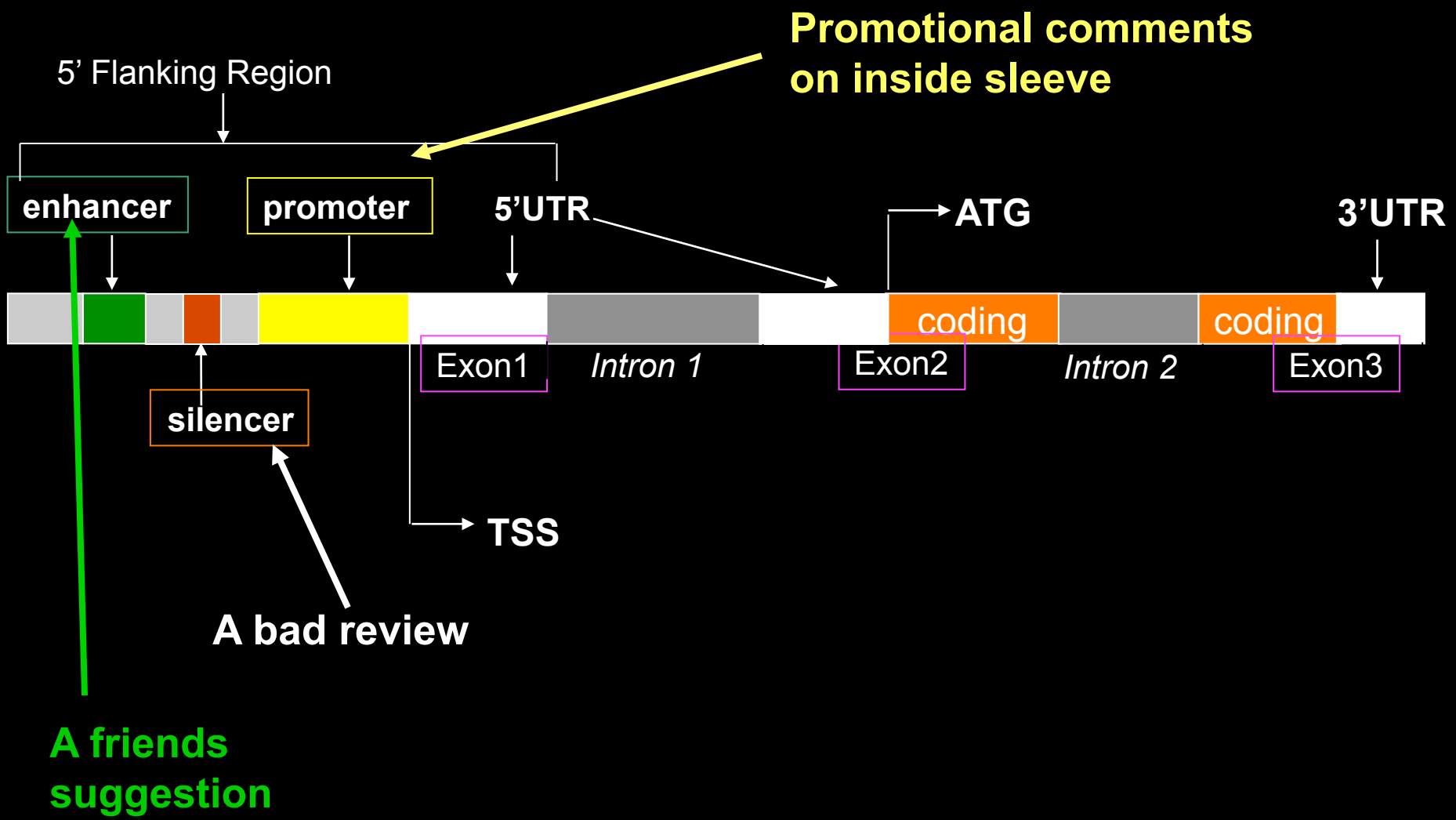
~ Genes account for 1% of the 3 billion base-pairs (A,T,C,G) contained on 23 chromosomes and 2 sex-chromosomes (X,Y)

~ Less than 0.1% of the human genome differs between two people (we are 99.9% similar) – this corresponds to 3 Million differences

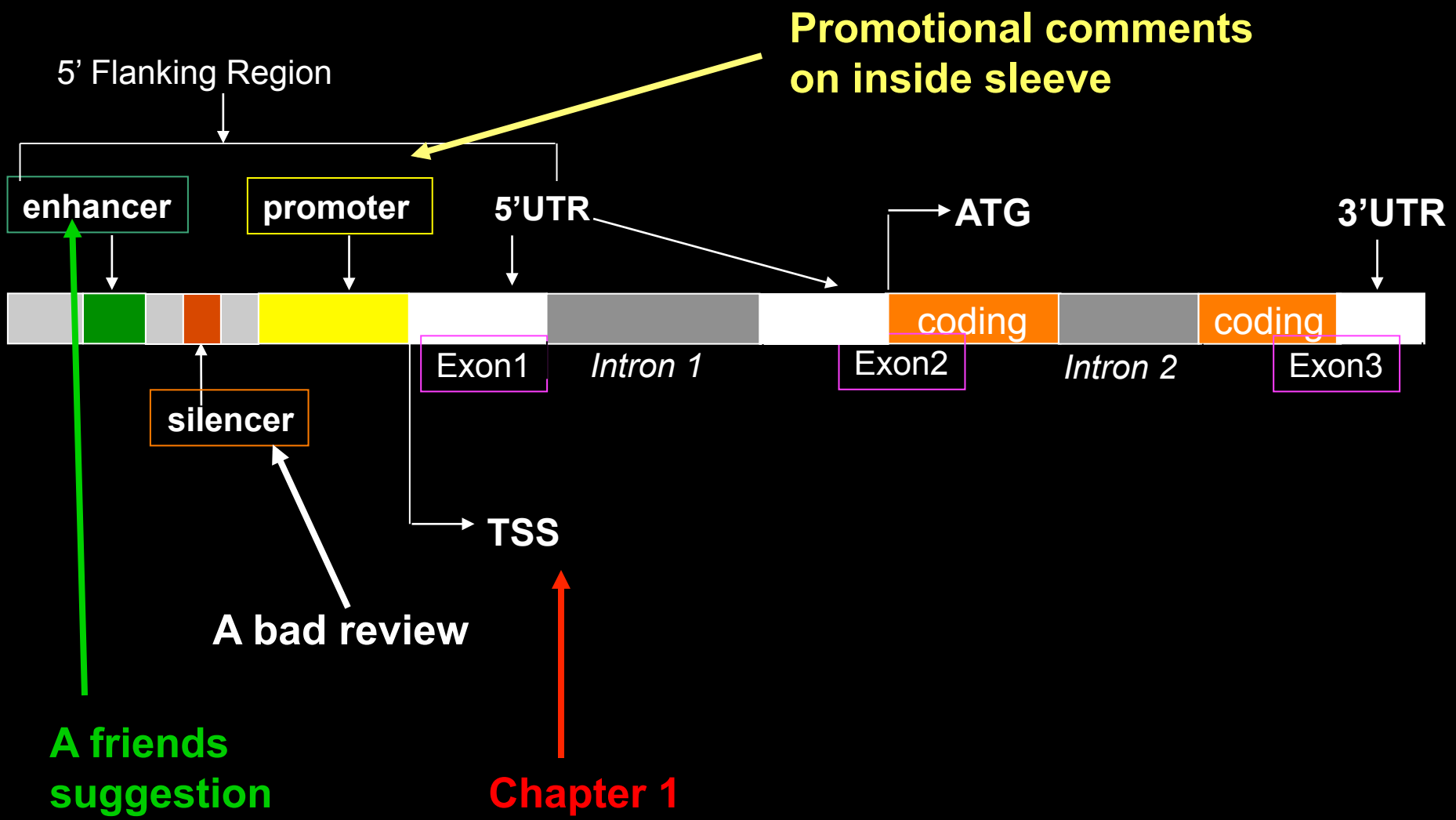
The structure of a gene is much like a book…..



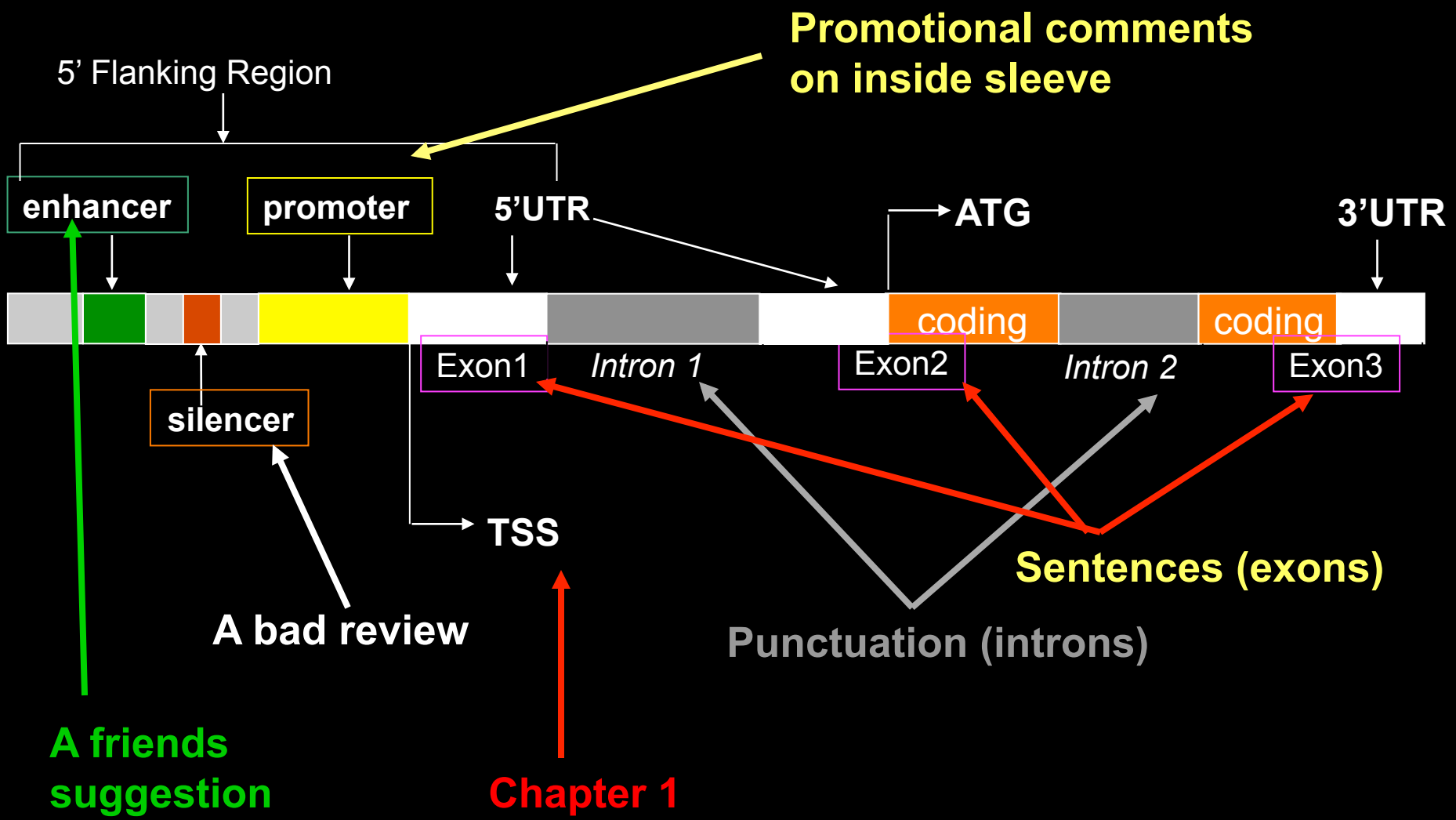
The structure of a gene is much like a book…..



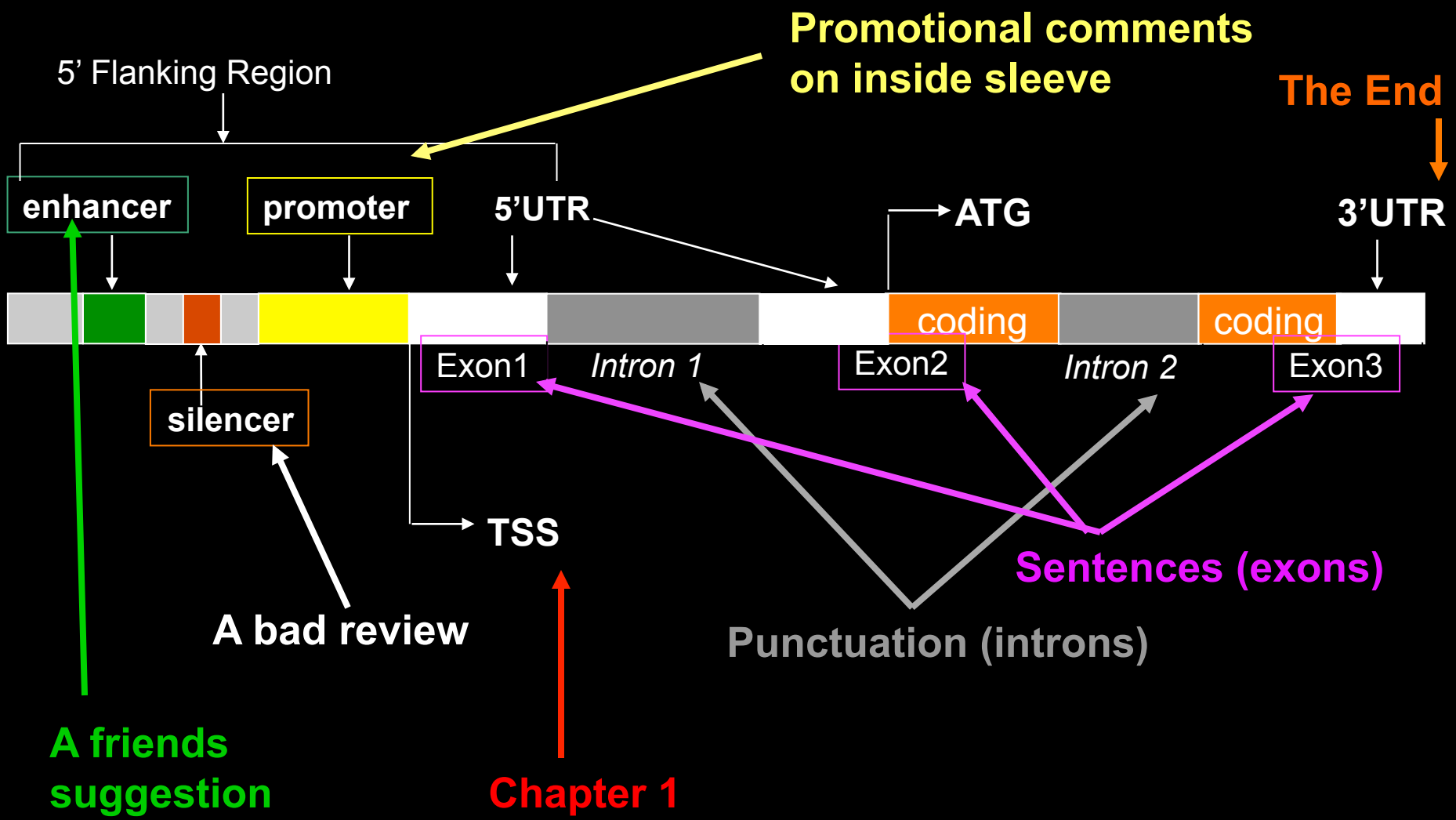
The structure of a gene is much like a book…..



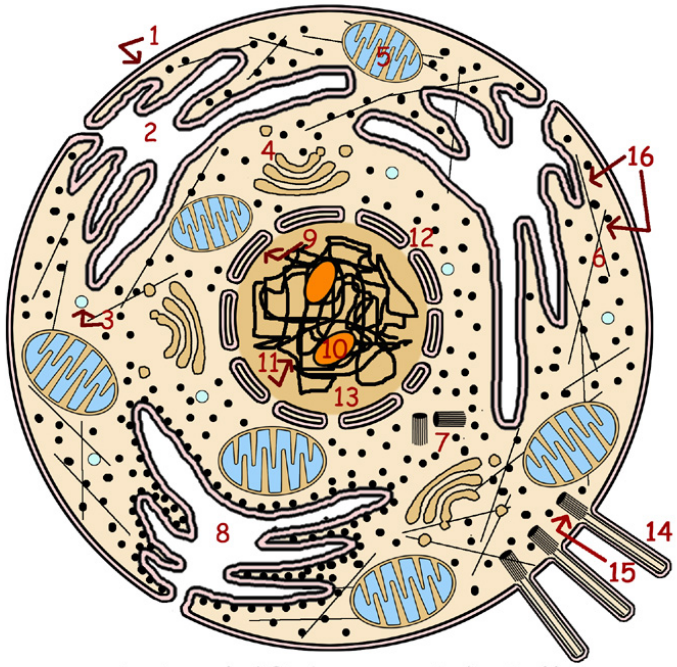
The structure of a gene is much like a book…..



The structure of a gene is much like a book…..



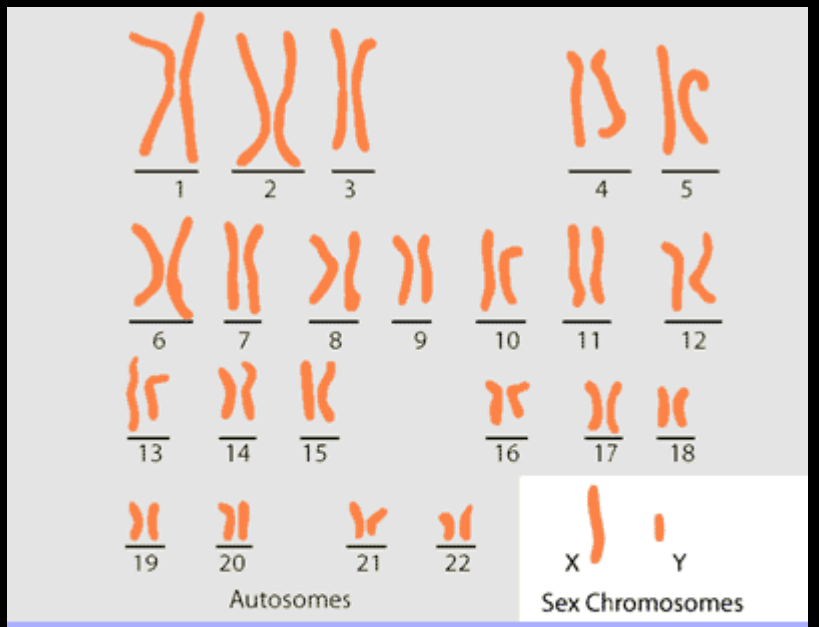
Genes are located in the cell nucleus on chromosomes



Animal (Eukaryotic) Cell

1-Plasma Membrane
2-Smooth Endoplasmic Reticulum
3-Vacuole
4-Golgi Body with Vesicles
5-Mitochondrion
6-Ribosomes
7-Centrioles
8-Rough Endoplasmic Reticulum
9-Double Nuclear Envelope
10-Nucleolus
11-Chromosomes (Chromatin)
12-Nuclear Pore
13-Nucleoplasm
14-Cilium
15-Basal Body
16-Cytoskeleton

Karyotype



The Human Genome

The Human Genome is the total of the genetic information that is held in each human cell. It is usually made up of 46 chromosomes: 22 pairs of autosomes and 1 pair of sex chromosomes, which are usually X and X for females and X and Y for males.

Types of Genetic Variation in the human genome:

Class of variation	Designation/Description	Example	Frequency
Single Nucleotide Polymorphism (SNP)	Single base substitution involving A,T,C, or G	A/T COMT	11,883,685 (6,262,709)
Deletion/Insertion Polymorphisms (DIPs; In/Del)	Designated using the full sequence of the insertion as one allele, and either a fully defined string for the variant allele or a "-" character to specify the deleted allele.	T/-CCTA/G 5-HTTLPR	431,319
Microsatellite or short tandem repeat (STR)	Alleles are designated by providing the repeat motif and the copy number for each allele.	DRD4 MAO-A	~10,000
Copy Number Variants	Duplication, deletion of 1000 up to 1 Mb of DNA		TBD 5% of genome
Gross Chromosomal Variation	Contiguous gene deletion syndromes (> 1 Mb) Aneuploidy	VCFS	small

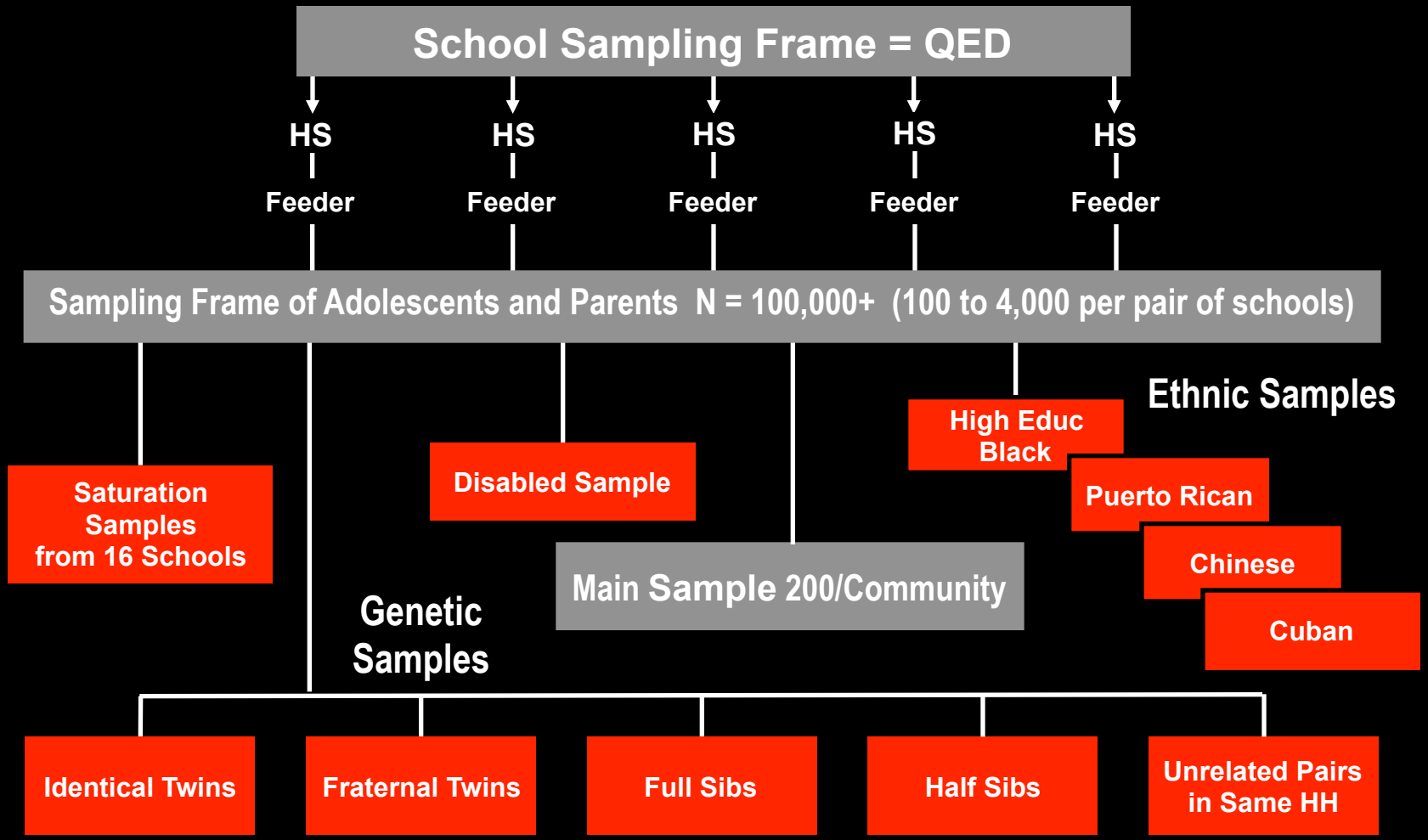
National Longitudinal Study of Adolescent Health

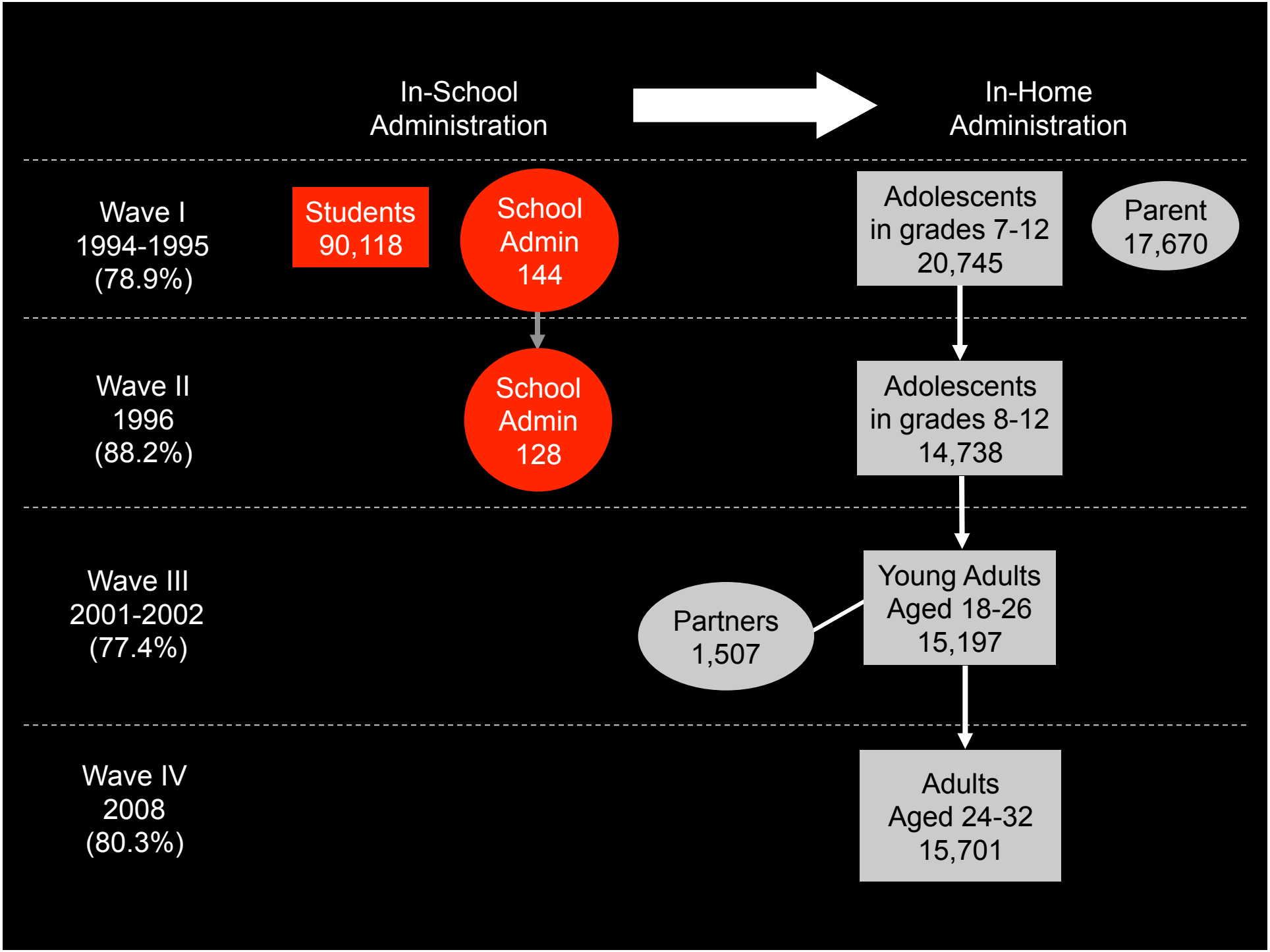
- On-going program project that began in 1994 and located at the University of North Carolina.
- Developed in response to a congressional mandate to fund a study of adolescent health.
- Funded by the National Institute of Child Health and Human Development (NICHD) with co-funding from 23 other federal agencies and foundations.
- Fourth wave of data collection funded in 2006 and collected in 2008.

Key Features of Add Health

- Nationally representative study that explores the causes of health and health-related behaviors of adolescents and their outcomes in adulthood.
- Multi-survey, multi-wave inter-disciplinary design.
- Direct measurement of the social contexts of adolescent life and their effects on health and health behavior.
- Unprecedented racial and ethnic diversity and genetically informed sibling samples.

Add Health Sampling Structure




In-School Administration
In-Home Administration
Wave I 1994-1995 (78.9%)
Students 90,118
School Admin 144
Adolescents in grades 7-12 20,745
Parent 17,670
Wave II 1996 (88.2%)
School Admin 128
Adolescents in grades 8-12 14,738
Wave III 2001-2002 (77.4%)
Partners 1,507
Young Adults Aged 18-26 15,197
Wave IV 2008 (80.3%)
Adults Aged 24-32 15,701

Biological Data Across Waves

Adolescence →	Young Adulthood →	Adulthood
Wave I-II (Ages 12-18)	Wave III (Ages 18-26)	Wave IV (Ages 24-32)
Embedded genetic sample of 3,000 pairs →		
Physical development		
Height, weight	Height, weight	Height, weight, waist
	STI tests (urine)	Metabolic
	HIV test (saliva)	Immune function
	DNA (buccal cell)	Inflammation
		Cardiovascular
		DNA (buccal cell)
		Medications

Add Health Sibling Pairs Sample

- Sort out genetic from environmental effects and determine if there are sex differences in their influence.
- Explore gene-environment interactions and gene-environment correlations.
- Determine the extent that genes and environments are stable and/or change across developmental periods.
- Obtain the genetic and environmental correlations between different, but related, traits.

Add Health Sibling Pairs Sample

The National Longitudinal Study of Adolescent Health (Add Health) Twin Data

Kathleen Mullan Harris,[1,2] Carolyn Tucker Halpern,[1,3] Andrew Smolen,[4] and Brett C. Haberstick[4]

[1] *Carolina Population Center, University of North Carolina at Chapel Hill, North Carolina, United States of America*
[2] *Department of Sociology, University of North Carolina at Chapel Hill, North Carolina, United States of America*
[3] *Department of Maternal and Child Health, School of Public Health, University of North Carolina at Chapel Hill, North Carolina, United States of America*
[4] *Institute for Behavioral Genetics, University of Colorado at Boulder, Colorado, United States of America*

The National Longitudinal Study of Adolescent Health (Add Health) Sibling Pairs Data

Kathleen Mullan Harris,[1,2] Carolyn Tucker Halpern,[1,3] Brett C. Haberstick,[4] and Andrew Smolen[4]

[1] *Carolina Population Center, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA*
[2] *Department of Sociology, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA*
[3] *Department of Maternal and Child Health, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA*
[4] *Institute for Behavioral Genetics, University of Colorado at Boulder, Boulder, CO, USA*

Twin Research and Human Genetics, 2006 & 2013

Add Health Sibling Pairs Sample

Relatedness	Wave I	Wave III	Both siblings at Wave III	Wave IV	Both Siblings at Wave IV
MZ Twins	307	273	239	282	233
DZ Twins	452	402	320	418	336
ND Twin type	25	19	14	22	18
Full-Siblings	1251	1152	924	1173	924
Half-Siblings	442	388	287	399	289
Nonrelated siblings	662	576	361	611	388
Total N	**3139**	**2810**	**2145**	**2905**	**2188**

Add Health Sibling Pairs Sample

Heritability

Heritability: the proportion of individual differences in a particular trait or disorder in a particular population that results from genetic differences between individuals.

Phenotypic variation is calculated as: $V_P = V_A + V_D + V_C + V_E$

Genetic influences include: additive (A) and non-additive (D) genetic influences
Environmental influences include: shared (C) and non-shared or unique (E)

Heritability estimates range from 0 to 1
Heritability estimates are aggregate estimates – provides no insight into which genes or variants

Most traits or diseases have some degree of heritable influence

Heritability	Behavioral /Psychiatric	Other Important Familial Traits
~zero		Language, Religion
20-40%	Anxiety disorders, Depression, Bulimia, Personality Disorders	Myocardial Infarction, Normative Personality, Breast Cancer, Hip Fracture
40-60%	Substance Use Disorders, ADHD, Physical Activity, Aggressive behaviors	Blood Pressure, Asthma Plasma cholesterol, Prostate Cancer, Adult-onset diabetes
60-80%	Schizophrenia Bipolar Illness	Weight, Bone Mineral Density
80-100%	Autism	Height, Total Brain Volume

Add Health Sibling Pairs Sample

Most traits or diseases have some degree of heritable influence, though

~ heritability is not an absolute property of a physical or behavioral characteristic

~ it is a function of the *genetic and environmental variation* for a *given population in particular circumstances*, and at a *particular developmental stage*

~ if environment is homogenous, influence of genotype on phenotype will increase (heritability will be higher)

~ as environmental variations (includes measurement error) increase, influence of genotypes on phenotypes will decrease (heritability will be lower)

~ estimates will typically be quite imprecise, with large standard errors or estimation AND *will depend on how accurate and precise the phenotypic definition is assessed*

~ EXAMPLE: magnitude of genetic effects on nicotine dependence, DSM/ICD vs. FTND

~ heritability of nicotine dependence differs between the two measurement instruments

Add Health Sibling Pairs Sample

~ Heritable influences can remain stable and change over time, increasing and decreasing in their contribution to variation

~ age-related AND stage-related

~ Factors within heterogeneous phenotypes may evidence *differential heritability* and/or *serve as 'better' phenotypes for genetic investigations*

~ EXAMPLE: nicotine dependence (FTND), subjective experiences to nicotine

(Haberstick et al, Addiction, 2006)

~ Sex can have impact on genetic or environmental effects in two ways:

~ QUANTITATIVE: are absolute magnitude differences between males and females. Genetic effects are stronger in one sex than the other

~ QUALITATIVE: are different A or C influences for males or females

(Haberstick et al, Behavior Genetics, 2010)

Stable Genes and Changing Environments: Body Mass Index Across Adolescence and Young Adulthood

Brett C. Haberstick · Jeffery M. Lessem · Matthew B. McQueen · Jason D. Boardman · Christian J. Hopfer · Andrew Smolen · John K. Hewitt

Key Findings: Quantitative Sex Differences, early genetic influences on BMI are lasting where as early environmental influences are not

Table 4 Parameter estimates (95% confidence intervals) for the additive genetic and individual specific environmental influences on body mass index during adolescence and young adulthood

	Wave 1 (16 years)		Wave 2 (17 years)		Wave 3 (22 years)	
	Males	Females	Males	Females	Males	Females
a^2	.84 (.81–.87)		.86 (.81–.89)	.75 (.68–80)	.85 (.80–.89)	.83 (.78–.87)
e^2	.16 (.13–.19)		.14 (.11–.19)	.25 (.20–.32)	.15 (.11–20)	.17 (.13–.22)

Note: a^2 Total genetic influence, e^2 total individual-specific environmental influence

Genetic correlation

Environmental correlation

Table 5 Genetic and individual-specific environmental correlations (95% confidence intervals) for body mass index

		Wave1 → Wave 2	Wave1 → Wave3	Wave2 → Wave3
r_g	Males	.96 (.94–.98)	.85 (.82–.89)	.89 (.86–.93)
	Females	.97 (.95–.99)	.96 (.92–.99)	.97 (.93–1.0)
r_e	Males	.51 (.38–.62)	.30 (.11–.48)	.29 (.10–.47)
	Females	.66 (.56–.74)	.21 (.05–.36)	.33 (.20–.46)

Note: r_g Additive genetic correlation, r_e individual specific correlation

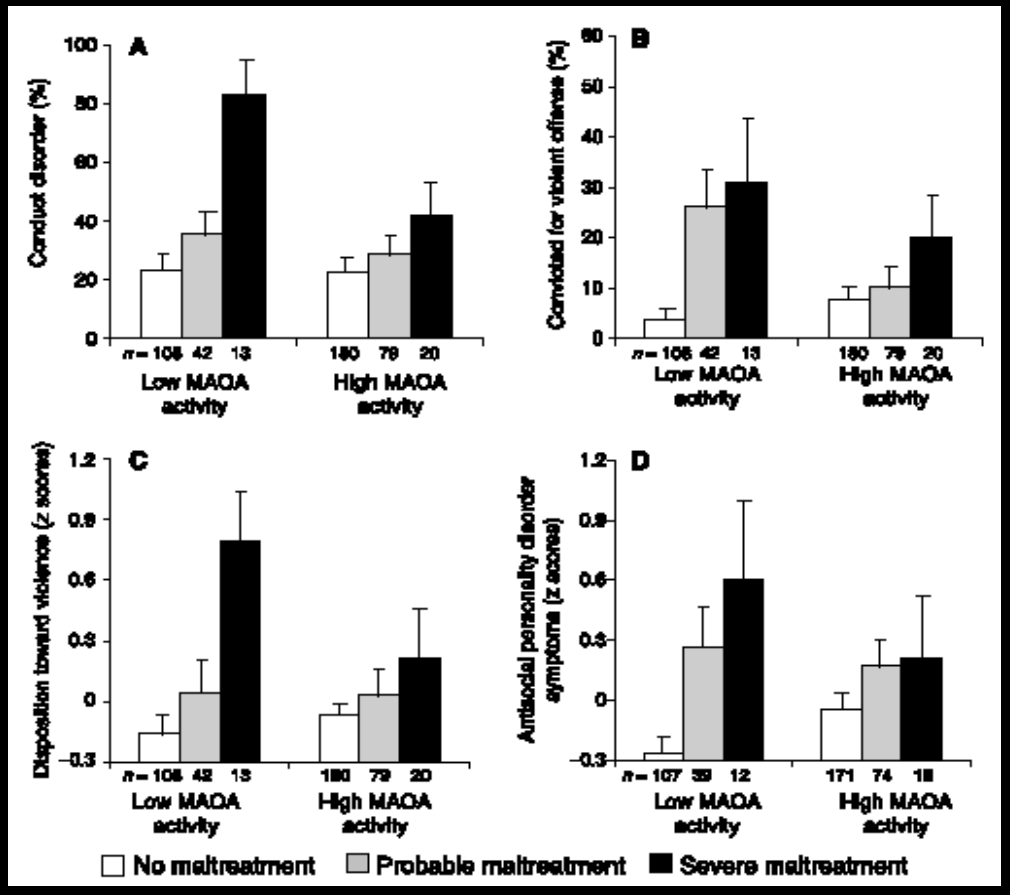
Role of Genotype in the Cycle of Violence in Maltreated Children

Avshalom Caspi,[1,2] Joseph McClay,[1] Terrie E. Moffitt,[1,2]* Jonathan Mill,[1] Judy Martin,[3] Ian W. Craig,[1] Alan Taylor,[1] Richie Poulton[3]

We studied a large sample of male children from birth to adulthood to determine why some children who are maltreated grow up to develop antisocial behavior, whereas others do not. A functional polymorphism in the gene encoding the neurotransmitter-metabolizing enzyme monoamine oxidase A (*MAOA*) was found to moderate the effect of maltreatment. Maltreated children with a genotype conferring high levels of *MAOA* expression were less likely to develop antisocial problems. These findings may partly explain why not all victims of maltreatment grow up to victimize others, and they provide epidemiological evidence that genotypes can moderate children's sensitivity to environmental insults.

Caspi et al. (2002), *Science* 297: 851-854

Key Findings: (1) Higher rates of conduct disorder, antisocial personality disorder, and a disposition towards violence was observed among those with the low functioning MAOA genotype; (2) low MAOA functioning was a risk factor for adult antisocial behavior among those who experienced more frequent and pervasive maltreatment before age 11.



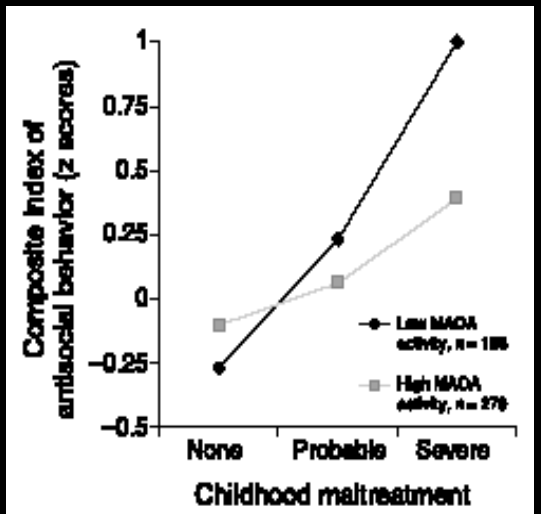


Fig. 1. Means on the composite index of antisocial behavior as a function of *MAOA* activity and a childhood history of maltreatment (*27*). *MAOA* activity is the gene expression level associated with allelic variants of the functional promoter polymorphism, grouped into low and high activity; childhood maltreatment is grouped into 3 categories of increasing severity. The antisocial behavior composite is standardized (z score) to a M = 0 and SD = 1; group differences are interpretable in SD unit differences (d).

Caspi et al. (2002), *Science* 297: 851-854

Role of Genotype in the Cycle of Violence in Maltreated Children

Avshalom Caspi,[1,2] Joseph McClay,[1] Terrie E. Moffitt,[1,2]* Jonathan Mill,[1] Judy Martin,[3] Ian W. Craig,[1] Alan Taylor,[1] Richie Poulton[3]

We studied a large sample of male children from birth to adulthood to determine why some children who are maltreated grow up to develop antisocial behavior, whereas others do not. A functional polymorphism in the gene encoding the neurotransmitter-metabolizing enzyme monoamine oxidase A (*MAOA*) was found to moderate the effect of maltreatment. Maltreated children with a genotype conferring high levels of *MAOA* expression were less likely to develop antisocial problems. These findings may partly explain why not all victims of maltreatment grow up to victimize others, and they provide epidemiological evidence that genotypes can moderate children's sensitivity to environmental insults.

Caspi et al. (2002), *Science* 297: 851-854

Childhood Adversity, Monoamine Oxidase A Genotype, and Risk for Conduct Disorder

Debra L. Foley, PhD; Lindon J. Eaves, PhD, DSc; Brandon Wormley, BS; Judy L. Silberg, PhD; Hermine H. Maes, PhD; Jonathan Kuhn, PhD; Brien Riley, PhD

Monoamine Oxidase A (MAOA) and Antisocial Behaviors in the Presence of Childhood and Adolescent Maltreatment

Brett C. Haberstick,[1]* Jeffrey M. Lessem,[1] Christian J. Hopfer,[2] Andrew Smolen,[1] Marissa A. Ehringer,[1] David Timberlake,[1] and John K. Hewitt[1]

[1]*Institute for Behavioral Genetics, University of Colorado, Boulder, Colorado*
[2]*Department of Psychiatry, University of Colorado Health Sciences Center, Denver, Colorado*

Role of Monoamine Oxidase A Genotype and Psychosocial Factors in Male Adolescent Criminal Activity

Kent W. Nilsson, Rickard L. Sjöberg, Mattias Damberg, Jerzy Leppert, John Öhrvik, Per Olof Alm, Leif Lindström, and Lars Oreland

Childhood Maltreatment, Subsequent Antisocial Behavior, and the Role of Monoamine Oxidase A Genotype

David Huizinga, Brett C. Haberstick, Andrew Smolen, Scott Menard, Susan E. Young, Robin P. Corley, Michael C. Stallings, Jennifer Grotpeter, and John K. Hewitt

MAOA and the "Cycle of Violence:" Childhood Abuse and Neglect, MAOA Genotype, and Risk for Violent and Antisocial Behavior

Cathy Spatz Widom and Linda M. Brzustowicz

MAOA, maltreatment, and gene–environment interaction predicting children's mental health: new evidence and a meta-analysis

J Kim-Cohen[1,2], A Caspi[2,3], A Taylor[2], B Williams[2], R Newcombe[2], IW Craig[2] and TE Moffitt[2,3]

[1]*Department of Psychology, Yale University, New Haven, CT, USA;* [2]*Social, Genetic, and Developmental Psychiatry Centre, Institute of Psychiatry, King's College London, London, UK and* [3]*Department of Psychology, University of Wisconsin at Madison, Madison, WI, USA*

Role of Genotype in the Cycle of Violence in Maltreated Children

Avshalom Caspi,[1,2] Joseph McClay,[1] Terrie E. Moffitt,[1,2]* Jonathan Mill,[1] Judy Martin,[3] Ian W. Craig,[1] Alan Taylor,[1] Richie Poulton[3]

We studied a large sample of male children from birth to adulthood to determine why some children who are maltreated grow up to develop antisocial behavior, whereas others do not. A functional polymorphism in the gene encoding the neurotransmitter-metabolizing enzyme monoamine oxidase A (*MAOA*) was found to moderate the effect of maltreatment. Maltreated children with a genotype conferring high levels of *MAOA* expression were less likely to develop antisocial problems. These findings may partly explain why not all victims of maltreatment grow up to victimize others, and they provide epidemiological evidence that genotypes can moderate children's sensitivity to environmental insults.

Caspi et al. (2002), *Science* 297: 851-854

Childhood Adversity, Monoamine Oxidase A Genotype, and Risk for Conduct Disorder

Debra L. Foley, PhD; Lindon J. Eaves, PhD, DSc; Brandon Wormley, BS; Judy L. Silberg, PhD; Hermine H. Maes, PhD; Jonathan Kuhn, PhD; Brien Riley, PhD

Monoamine Oxidase A (MAOA) and Antisocial Behaviors in the Presence of Childhood and Adolescent Maltreatment

Brett C. Haberstick,[1]* Jeffrey M. Lessem,[1] Christian J. Hopfer,[2] Andrew Smolen,[1] Marissa A. Ehringer,[1] David Timberlake,[1] and John K. Hewitt[1]

[1]*Institute for Behavioral Genetics, University of Colorado, Boulder, Colorado*
[2]*Department of Psychiatry, University of Colorado Health Sciences Center, Denver, Colorado*

Role of Monoamine Oxidase A Genotype and Psychosocial Factors in Male Adolescent Criminal Activity

Kent W. Nilsson, Rickard L. Sjöberg, Mattias Damberg, Jerzy Leppert, John Öhrvik, Per Olof Alm, Leif Lindström, and Lars Oreland

Childhood Maltreatment, Subsequent Antisocial Behavior, and the Role of Monoamine Oxidase A Genotype

David Huizinga, Brett C. Haberstick, Andrew Smolen, Scott Menard, Susan E. Young, Robin P. Corley, Michael C. Stallings, Jennifer Grotpeter, and John K. Hewitt

MAOA and the "Cycle of Violence:" Childhood Abuse and Neglect, MAOA Genotype, and Risk for Violent and Antisocial Behavior

Cathy Spatz Widom and Linda M. Brzustowicz

MAOA, maltreatment, and gene–environment interaction predicting children's mental health: new evidence and a meta-analysis

J Kim-Cohen[1,2], A Caspi[2,3], A Taylor[2], B Williams[2], R Newcombe[2], IW Craig[2] and TE Moffitt[2,3]

[1]Department of Psychology, Yale University, New Haven, CT, USA; [2]Social, Genetic, and Developmental Psychiatry Centre, Institute of Psychiatry, King's College London, London, UK and [3]Department of Psychology, University of Wisconsin at Madison, Madison, WI, USA

Monoamine Oxidase A (MAOA) and Antisocial Behaviors in the Presence of Childhood and Adolescent Maltreatment

Brett C. Haberstick,[1]* Jeffrey M. Lessem,[1] Christian J. Hopfer,[2] Andrew Smolen,[1] Marissa A. Ehringer,[1] David Timberlake,[1] and John K. Hewitt[1]

[1]*Institute for Behavioral Genetics, University of Colorado, Boulder, Colorado*

[2]*Department of Psychiatry, University of Colorado Health Sciences Center, Denver, Colorado*

American Journal of Medical Genetics, 135B: 59-64

Sample: Sibling-pairs in the Add Health sample between ages of18 and 15. Total sample size 771

Measures: Lifetime conduct disorder (longitudinal measures), violent convictions, physical and sexual abuse, neglect, visits from and removal by social services

Analysis: Family-based (increased statistical power)

Demographics: 8.5% low offending, 0.3 high offending; 76.1% no maltreatment, 18.6% one experience of maltreatment.

Monoamine Oxidase A (MAOA) and Antisocial Behaviors in the Presence of Childhood and Adolescent Maltreatment

Brett C. Haberstick,[1]* Jeffrey M. Lessem,[1] Christian J. Hopfer,[2] Andrew Smolen,[1] Marissa A. Ehringer,[1] David Timberlake,[1] and John K. Hewitt[1]

[1]*Institute for Behavioral Genetics, University of Colorado, Boulder, Colorado*

[2]*Department of Psychiatry, University of Colorado Health Sciences Center, Denver, Colorado*

American Journal of Medical Genetics, 135B: 59-64

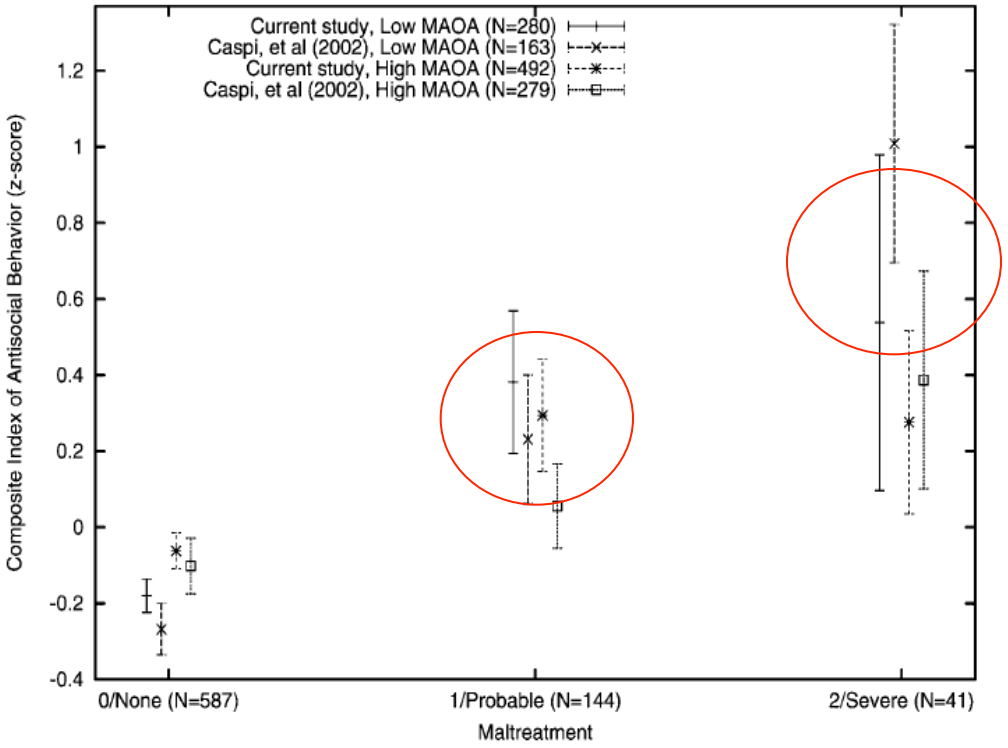


Fig. 1. Mean levels (z-scored) of antisocial behaviors for those experiencing none, probable, and severe amounts of childhood maltreatment as a function of monoamine oxidase A (MAOA) genotype for the current and Caspi et al. [2002] studies. For each of the three childhood maltreatment groups within the current study, standard errors around the mean indicate that the means did not differ significantly by genotype. For readability points have been offset slightly from each other, but are centered around the appropriate tick mark. Means and standard errors kindly provided by Caspi et al. [personal communication, 2004].

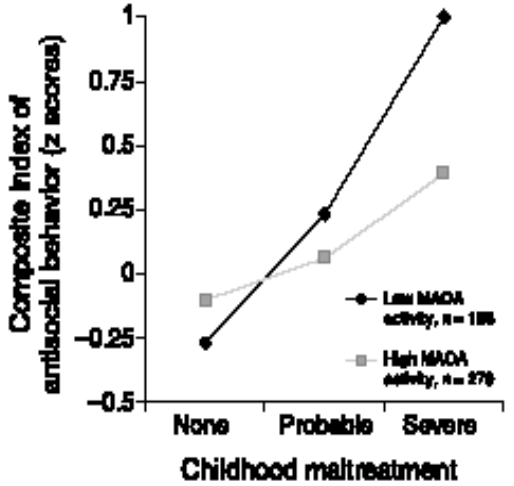


Fig. 1. Means on the composite index of antisocial behavior as a function of *MAOA* activity and a childhood history of maltreatment (*27*). *MAOA* activity is the gene expression level associated with allelic variants of the functional promoter polymorphism, grouped into low and high activity; childhood maltreatment is grouped into 3 categories of increasing severity. The antisocial behavior composite is standardized (z score) to a M = 0 and SD = 1; group differences are interpretable in SD unit differences (d).

Caspi et al. (2002), *Science* 297: 851-854

Add Health 'Program Project' Sample

Simple Sequence Repeats in the National Longitudinal Study of Adolescent Health: An Ethnically Diverse Resource for Genetic Analysis of Health and Behavior

Brett C. Haberstick · Andrew Smolen · Gary L. Stetler · Joyce W. Tabor · Taylor Roy · H. Rick Casey · Alicia Pardo · Forest Roy · Lauren A. Ryals · Christina Hewitt · Eric A. Whitsel · Carolyn T. Halpern · Ley A. Killeya-Jones · Jeffrey M. Lessem · John K. Hewitt · Kathleen Mullan Harris

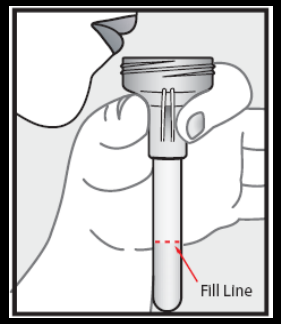
***Behavior Genetics**, 2014*

Add Health 'Program Project' Sample

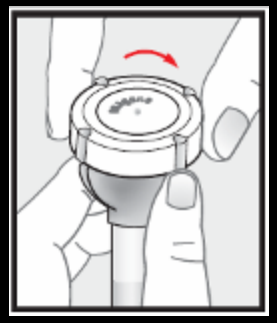
- Examine the degree of association between health/behavior phenotypes and genotypes.
- Explore gene-environment interactions.
- Determine the extent that genetic influences are stable and/or change across developmental periods.
- Conduct replication studies of *a priori* genetic associations findings from existing literature.
- Case-control & quantitative, not family based analyses

Wave IV DNA Collection

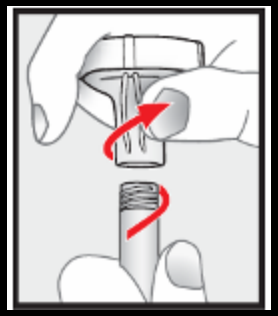
- DNA was collected using Oregene ® DNA Self-Collection kit using a 5-step process:



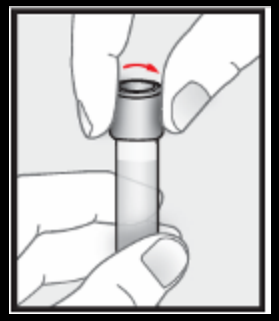
Spit until liquid saliva reaches the fill line.



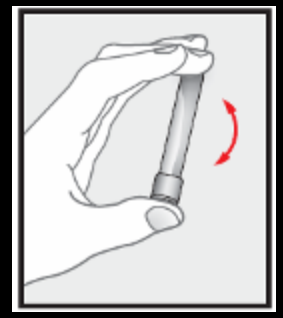
Close funnel. Liquid in big cap will mix with saliva.



Unscrew funnel from tube.



Close tube with small cap.



Mix 5 times. Throw out funnel and big cap.


EXEMPT HUMAN
SPECIMENS


BIOHAZARD

Add Health 'Program Project' Consent

Race/Ethnicity	Consent to saliva collection	Consent to archive
Whites	8051	6822
Blacks	3348	2523
Hispanic	2293	1883
Other	1335	997
Total N	**15,140**	**12,234**

Add Health Sibling Pairs Consent

Saliva Collection at Wave IV	Pairs who Consented	Pairs who did not consent	Total N
Monozygotic Twins	225	8	233
Dizygotic Twins	321	25	336
ND Twin type	15	3	18
Full-Siblings	880	44	924
Half-Siblings	278	11	289
Nonrelated siblings	371	17	388
Total N	**2090**	**98**	**2188**

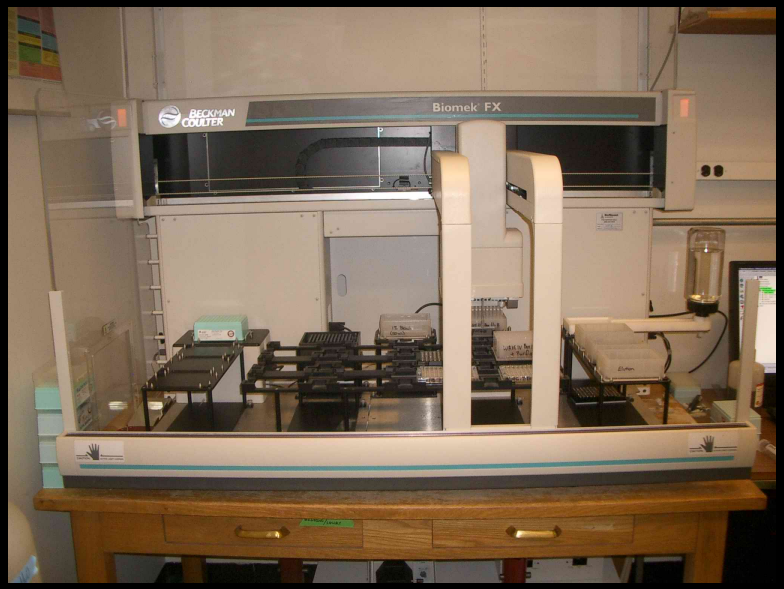
Full Sample – DNA Consent

Ethnicity	Gender	Consent – with archiving (%)	Consent – no archiving (%)	Refusal (%)
Hispanic	Male	78.1	17.2	4.6
	Female	72.9	23.3	3.8
Black	Male	74.4	23.0	4.6
	Female	70.5	25.0	3.5
Asian	Male	70.6	25.0	4.4
	Female	66.3	28.2	5.4
Native American	Male	83.4	17.1	1.4
	Female	80.6	18.7	0.7
Other	Male	70.0	22.0	7.1
	Female	73.3	17.3	9.3
White	Male	81.0	15.7	3.3
	Female	83.4	14.0	2.5
Total (%)		**77.0**	**18.5**	**3.6**

Reasons for refusal to provide saliva sample	Total (%)
Too invasive	40.1
Doesn't give biological samples	20.3
Concerned about confidentiality	17.4
Doesn't want to give/not comfortable	11.5
Religious reasons	2.9
Illness/disability/pregnant	1.4
Incentive wasn't enough	2.3
Concerned about future use of sample	1.3
Interview in a public place	2.0
Time constraints	2.7
Other	7.9
Total N	**557**

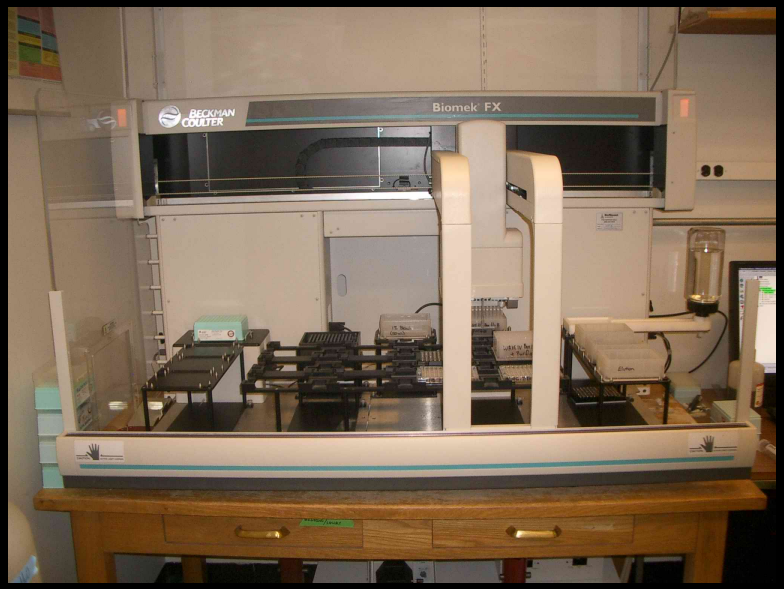
Add Health DNA Samples…next steps

- Extract DNA: used Zymo Research ZR-96 plates or Beckman-Coulter systems. Total #: **15,249 samples**



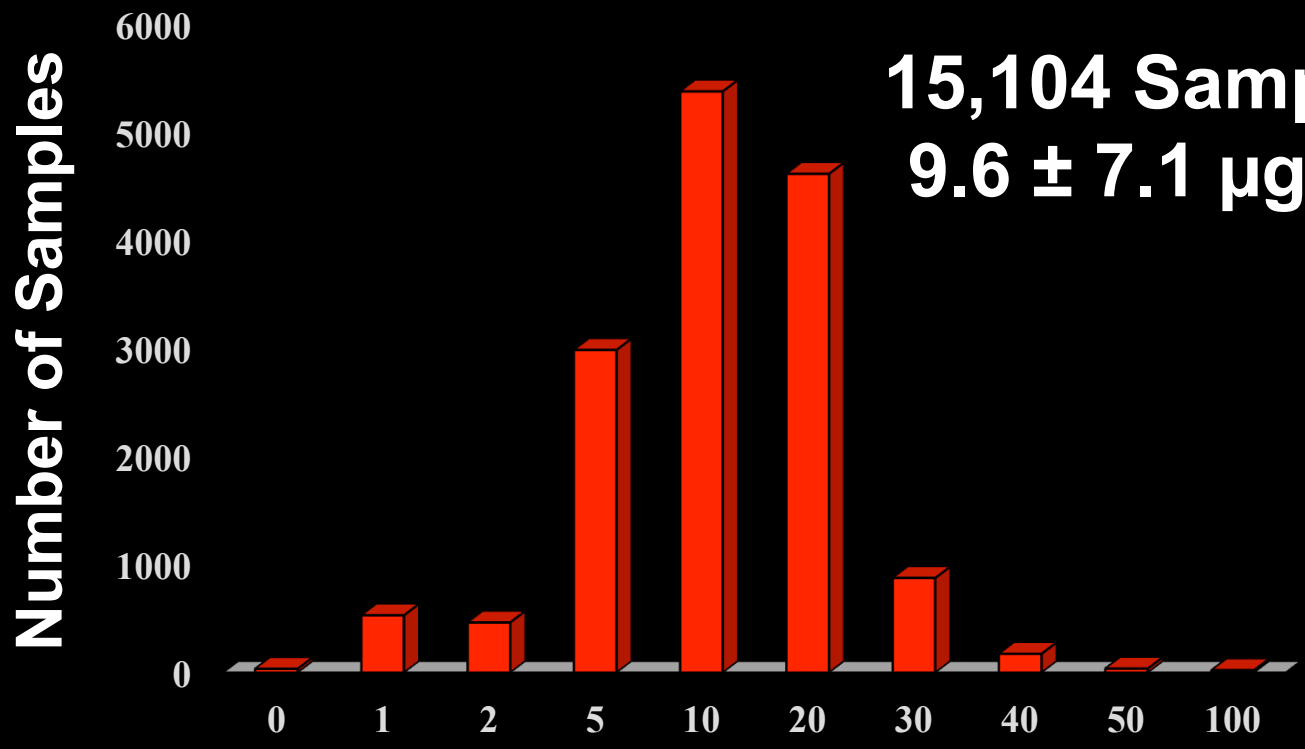
Add Health DNA Samples…next steps

- Extract DNA: used Zymo Research ZR-96 plates or Beckman-Coulter systems. Total #: **15,249 samples**



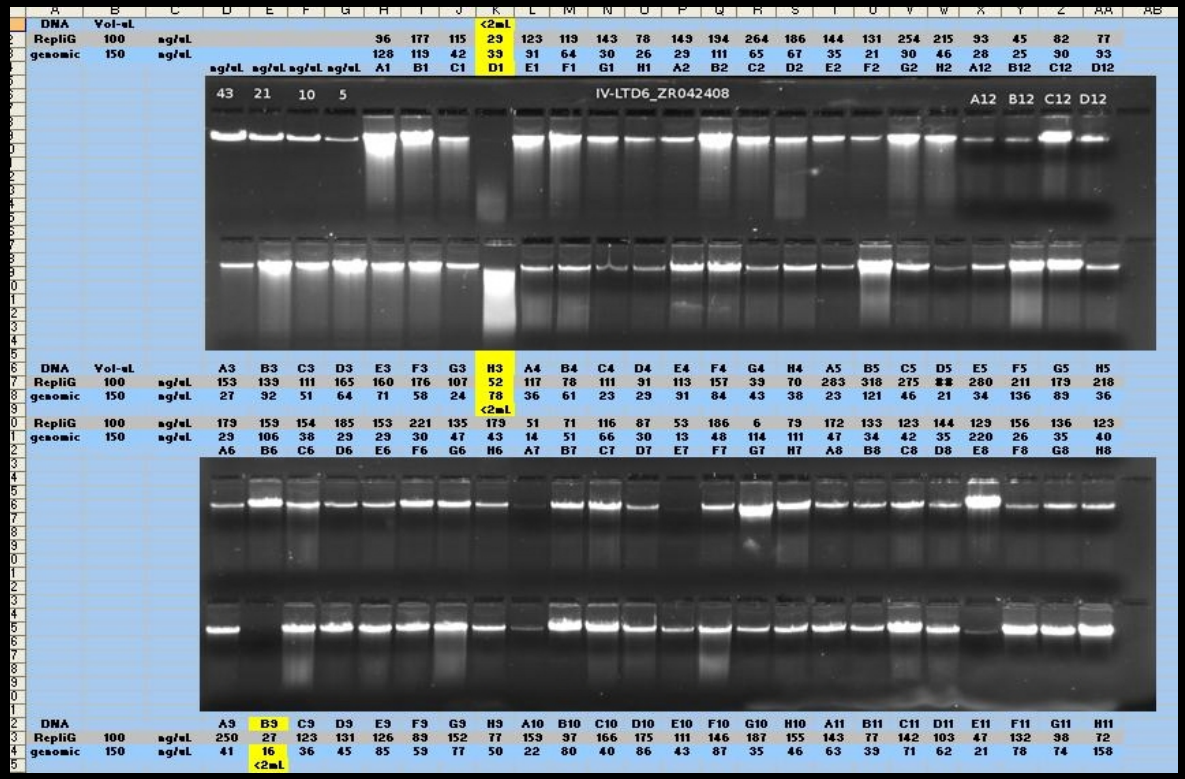
- Quantify DNA: PicoGreen® Fluorescence

DNA Yield for Oragene™ Saliva Method



Add Health DNA Samples…next steps

- Visualize DNA: Aragrose gel with SYBR®Safe



Measured Genotypes in Add Health

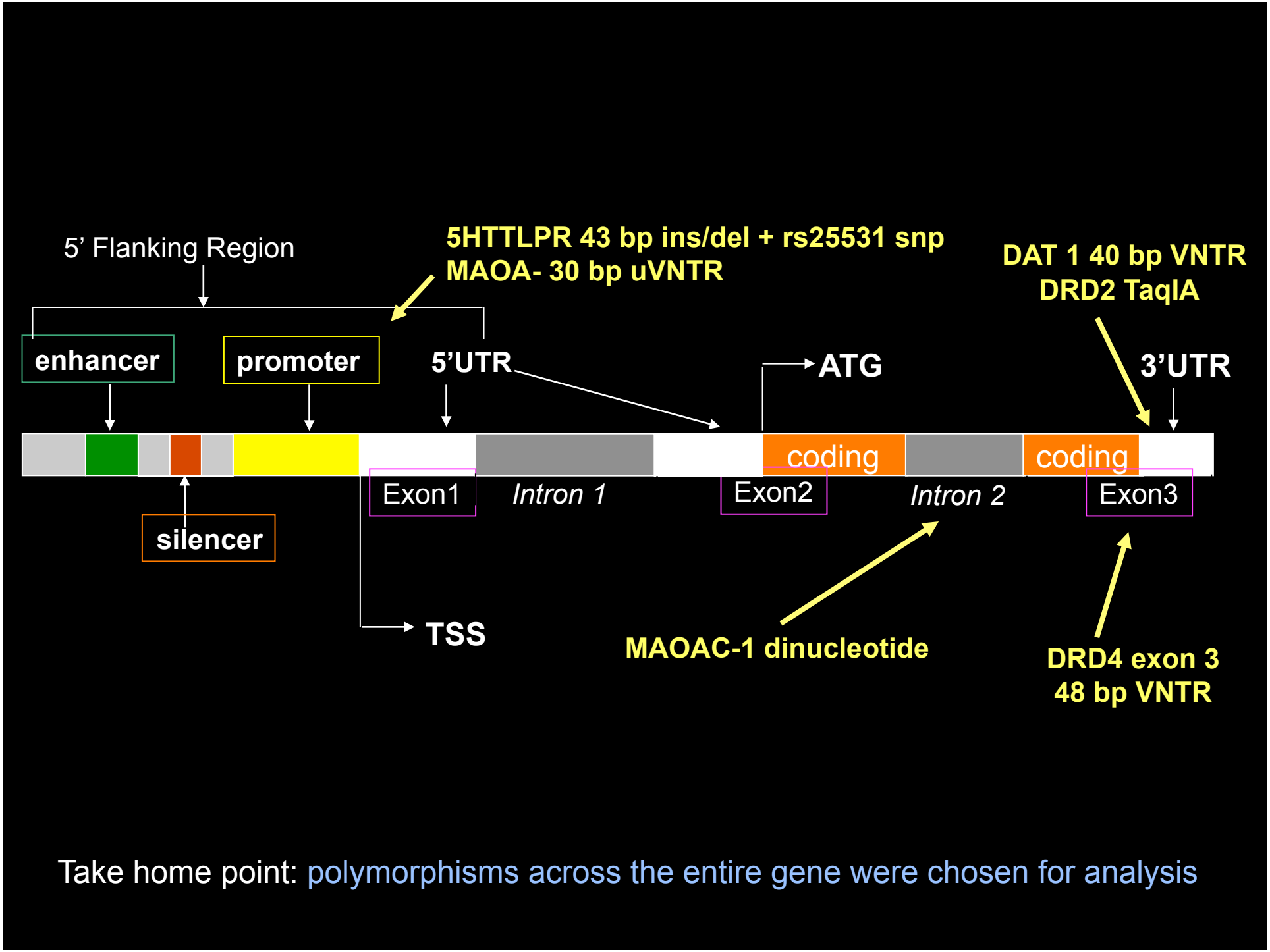
Name	Symbol	Type
Dopamine Transporter	DAT1	40 bp VNTR
Dopamine Receptor D4	DRD4	48 bp VNTR
Monoamine Oxidase A	MAOA	30 bp VNTR
Dopamine Receptor D2	DRD2	Taq IA SNP
Serotonin Tranpsorter	5HTTLPR	43 bp ins/del
Cytochrome P450 - 2A6	CYP2A6	SNP
Catechol-O-Methyl Transferase	COMT	Val/Met SNP
Dopamine D5 receptor	DRD5	Dinucleotide SSR
Monamine Oxidase A	MAOAC-1	Dinucleotide SSR

+ 93 Ancestral Informative Markers (AIMS)

VNTR: Variable number Tandem Repeat; SNP: Single Nucleotide polymorphism

Add Health Measured Genotypes

Genetic Marker	Run 1 Null Result (N, %)	Run 2 Null Result (N, %)	Null Result in Common (N, %)	Discordant Genotype Call (N, %)
DRD4	64 (0.02)	85 (0.03)	45 (0.44)	27 (0.0094)
DAT1	103 (0.04)	140 (0.05)	55 (0.43)	9 (0.0031)
5HTTLPR	128 (0.04)	118 (0.04)	55 (0.41)	49 (0.0021)
MAOA-uVNTR	133 (0.05)	91 (0.03)	51 (0.57)	5 (0.0017)
DRD5	267 (0.09)	399 (0.14)	135 (0.51)	4 (0.0014)
MAOAC-1	**221 (0.08)**	**325 (0.11)**	**110 (0.50)**	**47 (0.0164)**

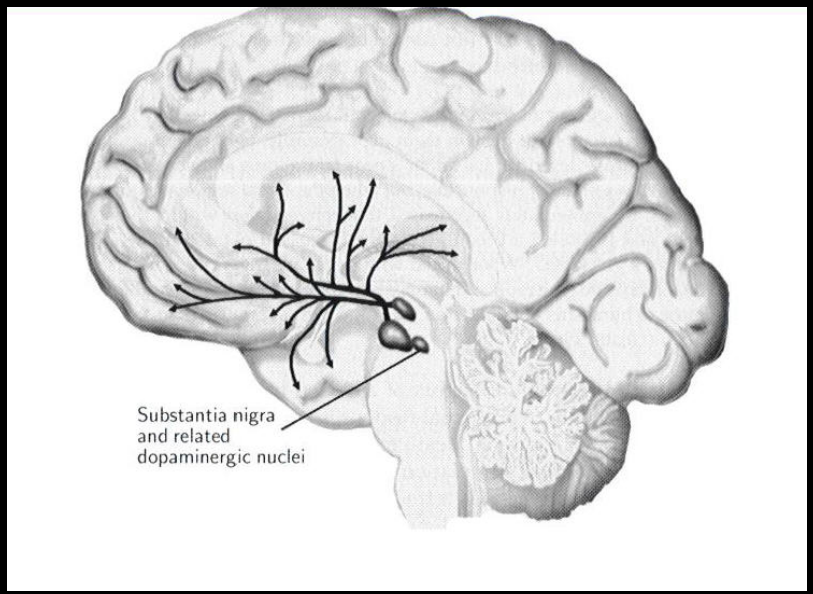


Take home point: polymorphisms across the entire gene were chosen for analysis

Dopaminergic System

~ Dopamine found in neurons of nigrostriatal, mesocortical and mesolimbic systems.

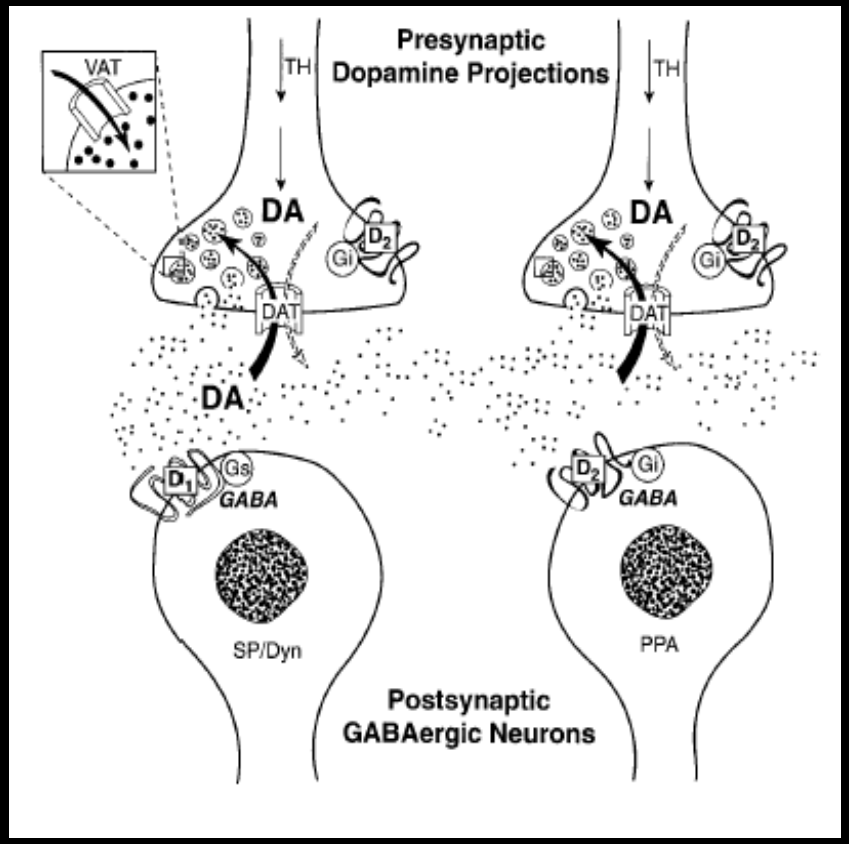
~ Role in motor function, reward, reinforcement, emotional expression, neuroendocrine release and behavioural homoeostasis.



from Dr. Ursula. M. D' Souza

Dopaminergic Receptor System

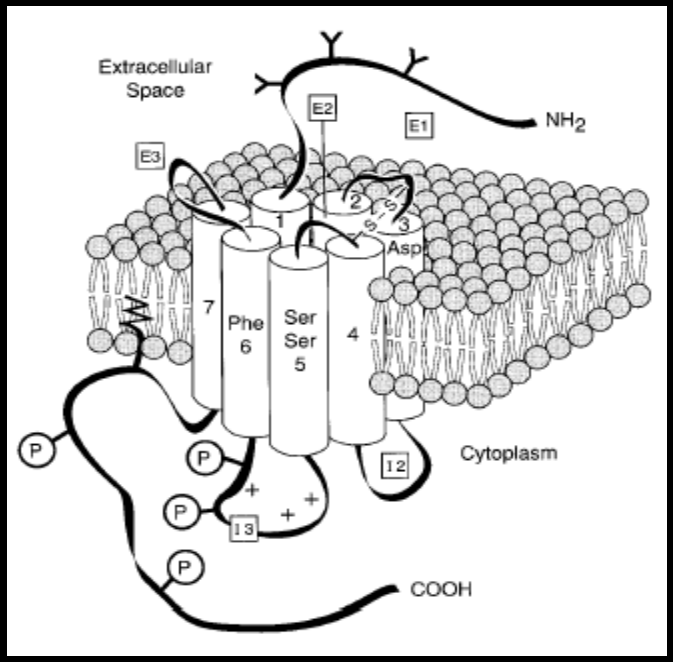
Transporters & Auto-transporters



Receptors

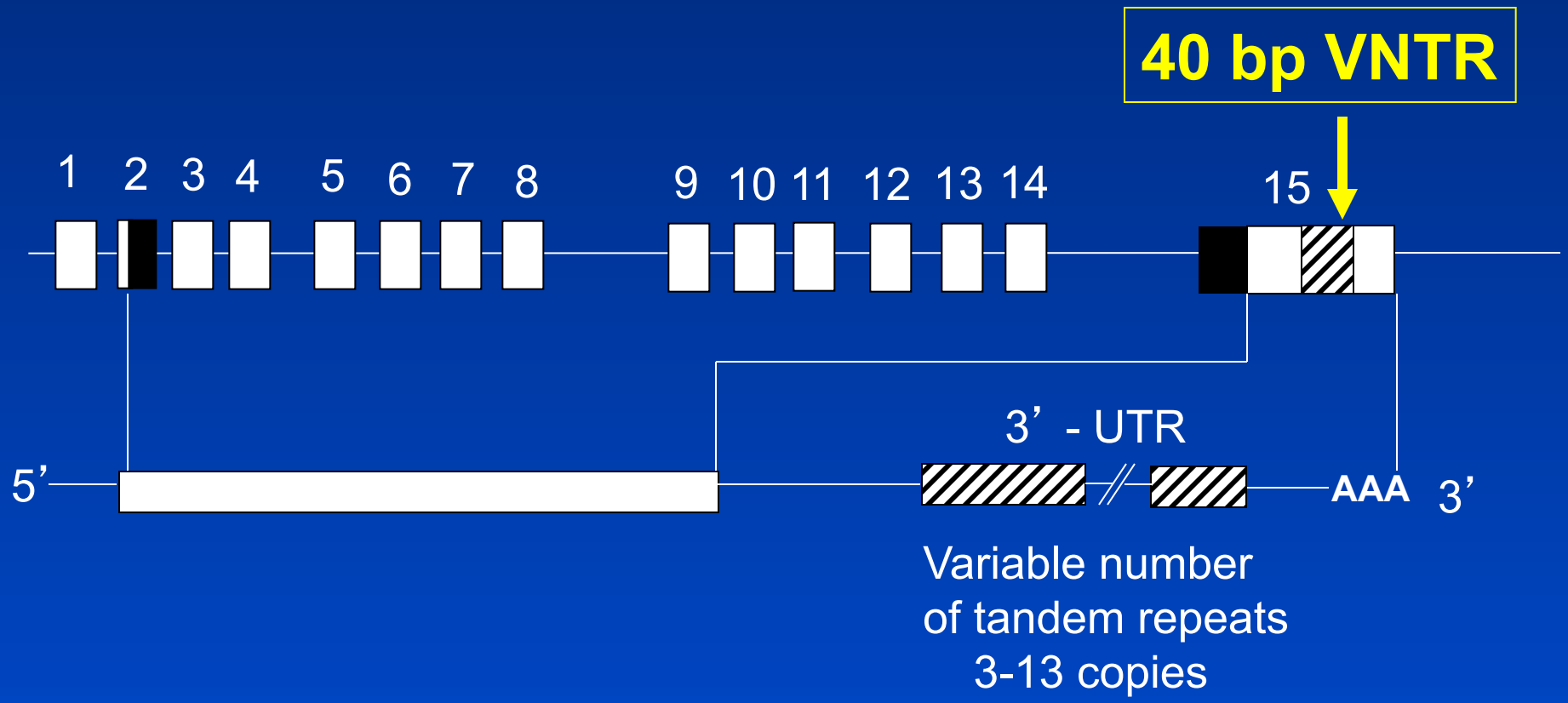
D_1-like (D1, D5)

D_2-like (D2, D3 and D4)



Missale et al., 1998: Physiology Reviews 78(1): 189-225

VNTR polymorphism of Dopamine Transporter

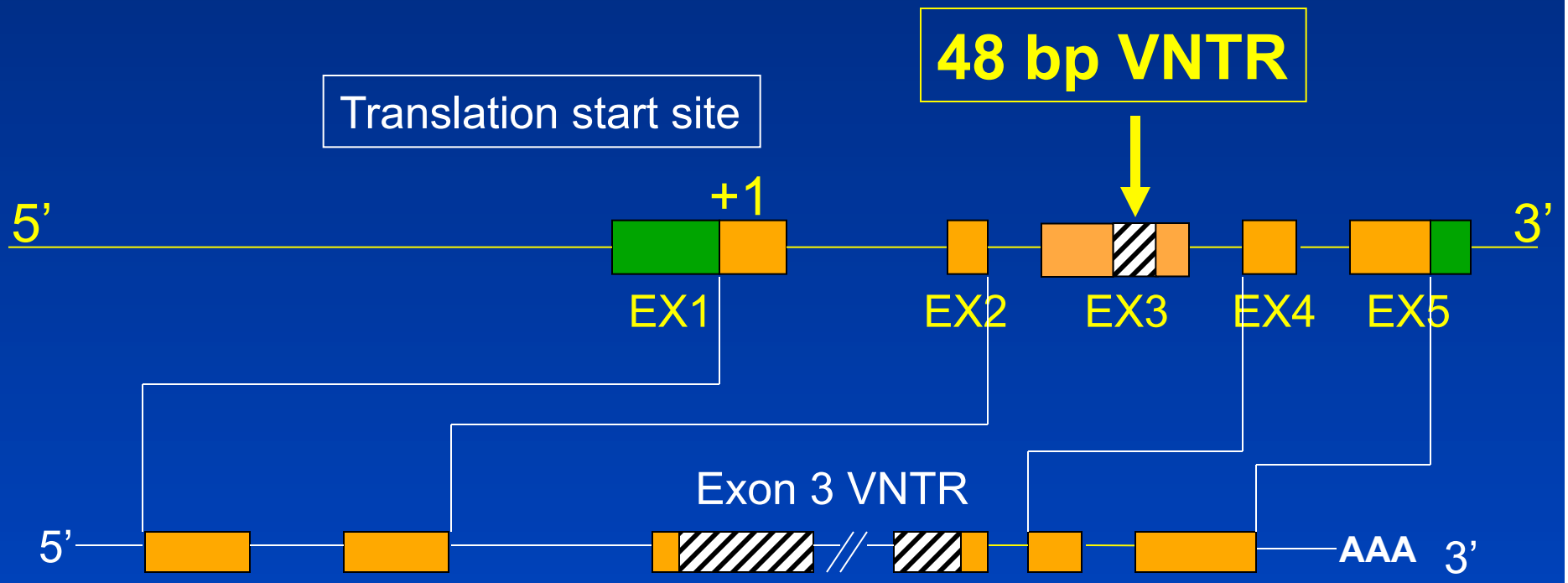


Adapted from Miller & Madras, 2002; Fuke et al., (2001), Vandenberg et al., 1992

Functional & behavioral effects of the 40 bp DAT1 VNTR

- Transcriptional efficiency
  - Michelhaugh et al., (2001)- 9 repeat enhances transcription in cells and also in dopamine neurons in neonatal rat midbrain slices.
  - Miller, GM & Madras, BK. (2002). 9 repeat shows significantly higher levels of luciferase production than 10 repeats in HEK293 cells. Cloned downstream
  - Fuke et al., (2001) – 10 repeat showed higher expression than the other repeats in COS-7 cells and human glioblastoma A172 cells.
  - Mill et al., (2002) - 10 repeat showed higher expression than other commonly express genes (housekeeping genes)
- Behavioral associations
  - Alcohol, tobacco, illicit drug use/misuse.
  - gambling
  - conduct problems, violence, delinquency
  - attention-deficit/hyperactivity

DRD4 VNTR Polymorphism in Exon 3



Variable number
of tandem repeats
2-11 copies

Adapted from: D’Souza; van Tol et al., (1992)

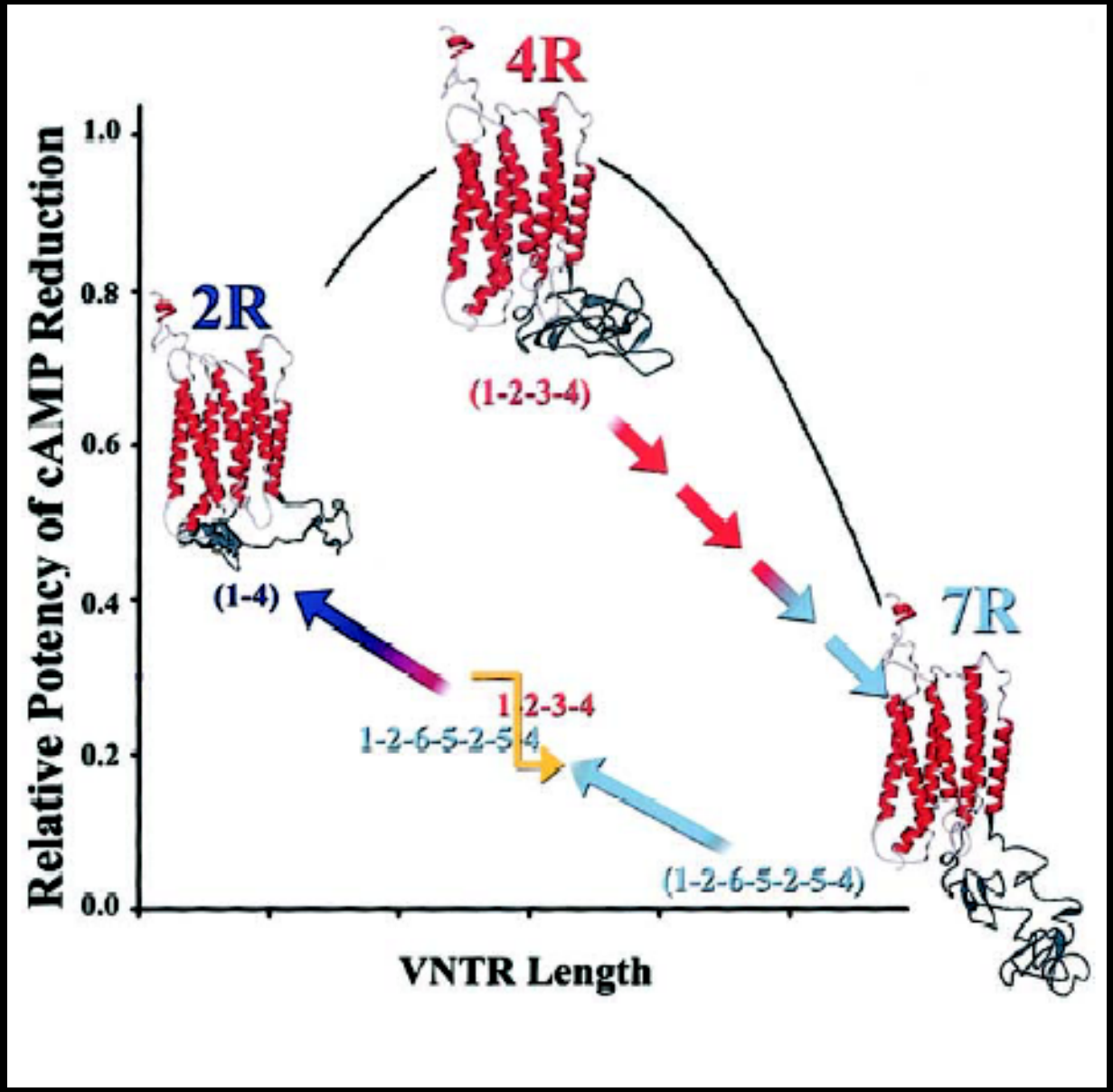
DRD4 VNTR Polymorphism of Exon 3

- DRD4 is A D_2-like receptor
- Fragment lengths vary between 379 bp - 811 bp
- *4R* (475 bp) & *7R* (619 bp) alleles are most common (64.3% and 20.6%, respectively)
- Results in variation in the 3rd cytoplasmic loop of the receptor affecting G-protein binding

- DRD4 is activated by dopmamine
- Inhibits adenylate cyclase and reduces cAMP levels

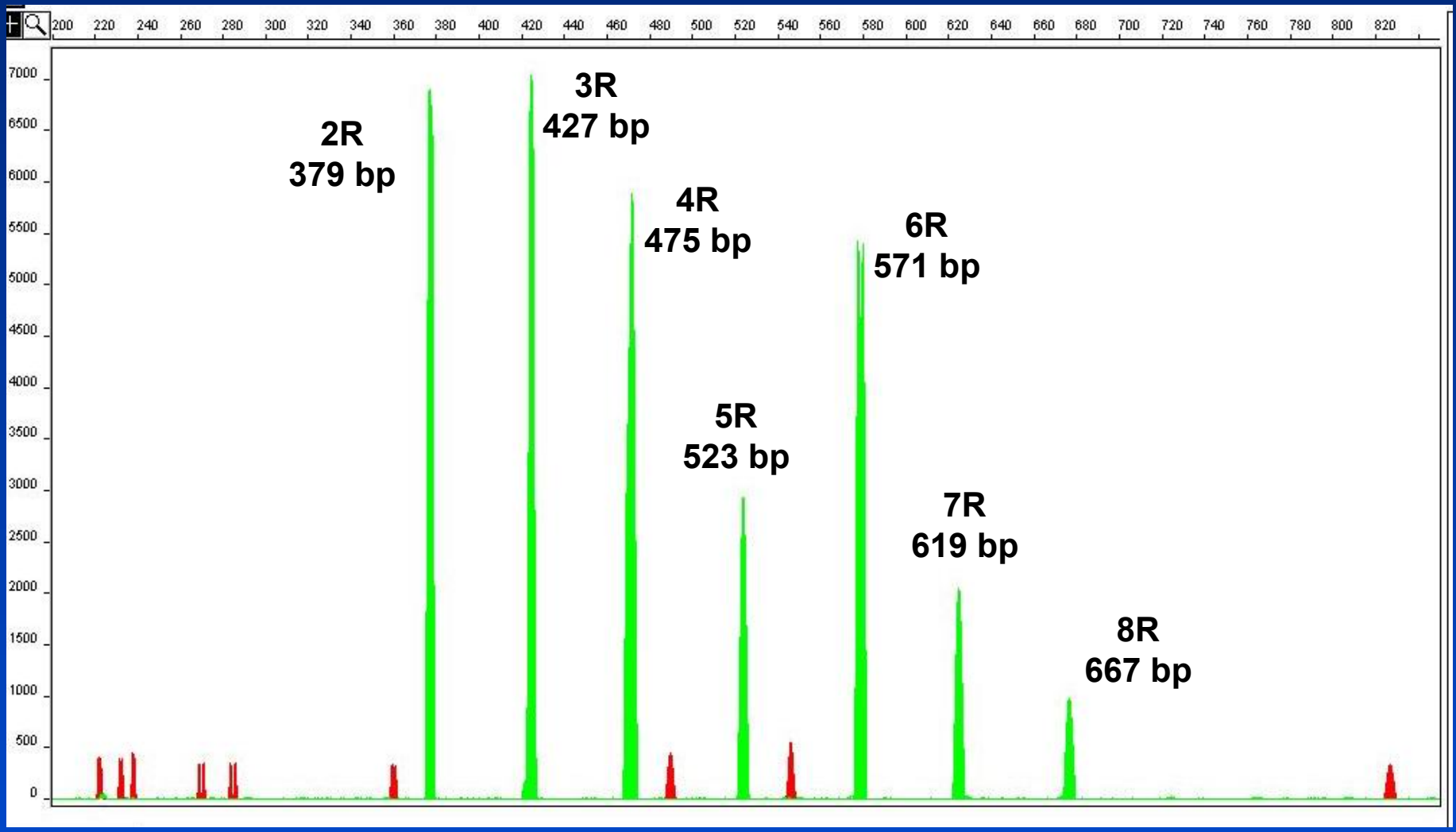
- *7R* allele exhibits blunted ability to reduce cAMP levels

Wang et al., (2004); DiMaio et al., (2003); Anchordoquy et al., (2003)



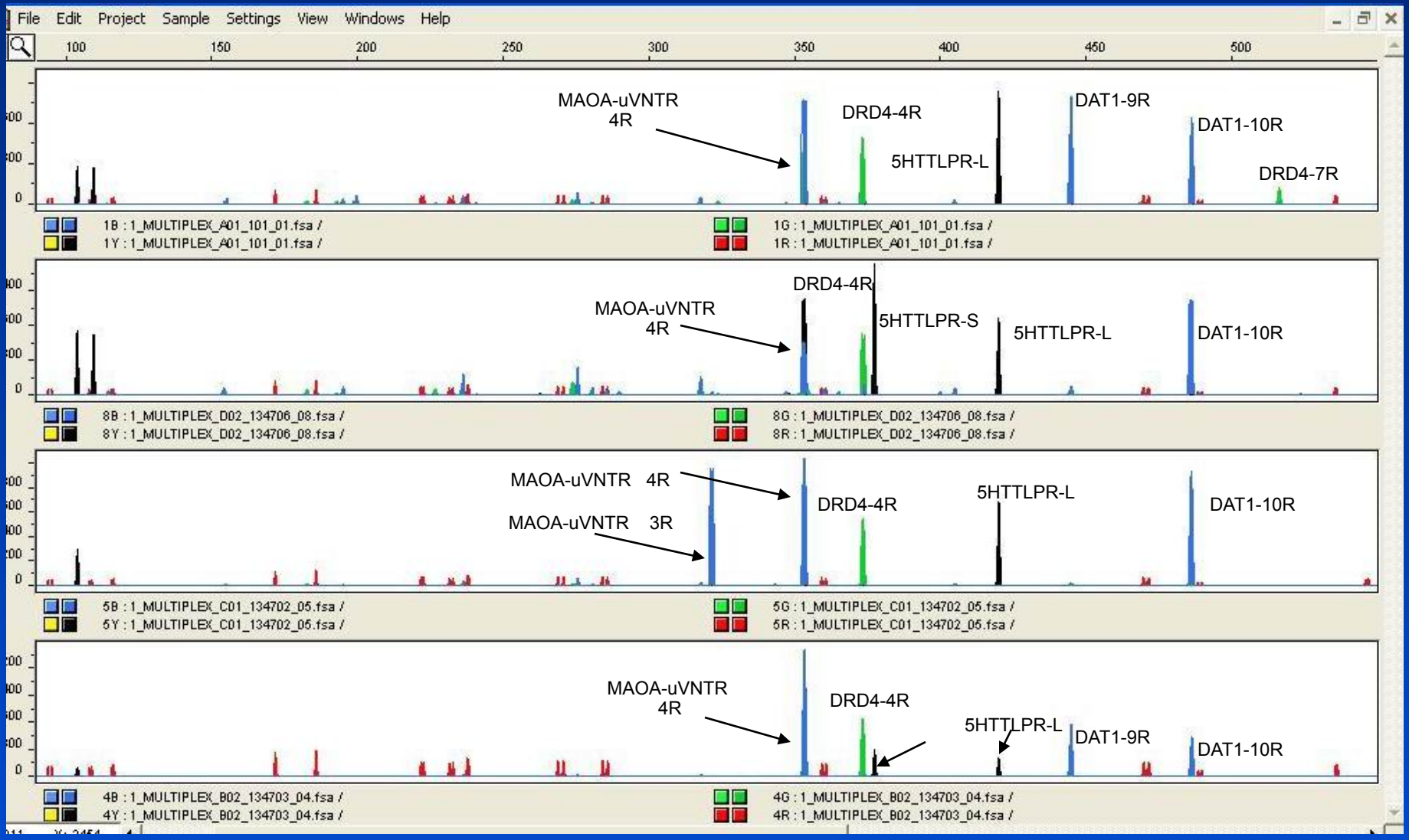
Wang et al., (2004), Am. J. Hum. Genet. 74: 931-944

DRD4 Exon-3 48 bp VNTR



VNTR and Ins/del Genotyping on ABI 3130xl

A panel of four Candidate Genes in a single multiplex PCR assay



Genotyping four typical VNTR Candidate Genes

Add Health 'Program Project' Genetics

Supplemental Table 3. Allele counts and frequencies (N, %) for DRD4 polymorphism by race/ethnicity.

	Race/Ethnicity				
Allele	White (N= 8188)	Black (N= 3235)	Native American (N= 116)	Asian (N= 905)	Hispanic (N= 2337)
2R	1440 (0.09)	326 (0.05)	18 (0.08)	469 (0.26)	280 (0.06)
3R	711 (0.04)	100 (0.02)	7 (0.03)	34 (0.02)	103 (0.02)
3.39R	--	14 (<0.01)	--	--	1 (<0.01)
4R	10622 (0.65)	4046 (0.63)	146 (0.63)	1165 (0.64)	2945 (0.63)
5R	136 (0.01)	279 (0.04)	2 (0.01)	28 (0.02)	83 (0.02)
6R	74 (<0.01)	82 (0.01)	11 (0.05)	40 (0.02)	106 (0.02)
7R	3239 (0.20)	1401 (0.22)	44 (0.19)	71 (0.04)	1119 (0.24)
8R	138 (0.01)	209 (0.03)	4 (0.02)	2 (<0.01)	32 (0.01)
9R	11 (<0.01)	4 (<0.01)	--	1 (<0.01)	5 (<0.01)
10R	5 (<0.01)	9(<0.01)	--	--	--

Add Health 'Program Project' Genetics

Supplemental Table 7. Allele counts and frequencies (N, %) for MAOA-uVNTR polymorphism by race/ethnicity for **Males**.

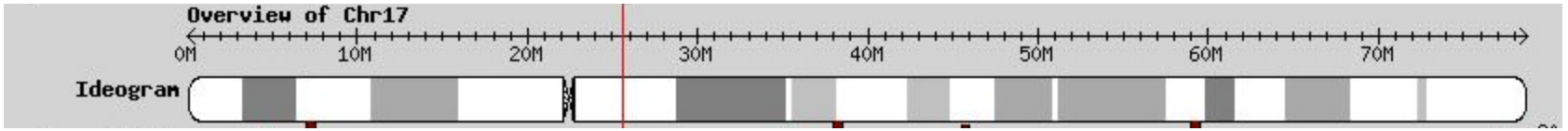
	Race/Ethnicity				
Allele	White (N= 8369)	Black (N= 1403)	Native American (N= 117)	Asian (N= 468)	Hispanic (N= 2360)
2R	16 (<0.01)	142 (0.05)	--	--	18 (<0.01)
3R	2662 (0.34)	1408 (0.50)	40 (0.36)	526 (0.56)	808 (0.34)
3.5R	124 (0.02)	4 (<0.01)	2 (0.02)	--	26 (0.02)
4R	4850 (0.62)	1238 (0.44)	70 (0.62)	408 (0.44)	1350 (0.62)
5R	116 (0.01)	14 (<0.01)	--	2 (<0.01)	12 (<0.01)

Note: R, repeat; N, sample size.

Serotonin Transporter, SLC6A4

17q11.1-q12

minus strand



[25,392,812

[25,685,191

EFCAB5

SLC6A4

TMIGD1

CCDC55

BLMH

rs2066713

25,586,831]

25,549,032]

5’

3’

5HTTLPR

1a

1b 2 3 4 5 6 7 8 9 10 11 12 13 14

rs12945042

rs25531
MspI
L_A/L_G

rs6354

rs2020942

rs140700

rs2054847

rs1042173

Coding Region

Untranslated Region

Simple Sequence Repeats in the National Longitudinal Study of Adolescent Health: An Ethnically Diverse Resource for Genetic Analysis of Health and Behavior

Brett C. Haberstick · Andrew Smolen · Gary L. Stetler · Joyce W. Tabor · Taylor Roy · H. Rick Casey · Alicia Pardo · Forest Roy · Lauren A. Ryals · Christina Hewitt · Eric A. Whitsel · Carolyn T. Halpern · Ley A. Killeya-Jones · Jeffrey M. Lessem · John K. Hewitt · Kathleen Mullan Harris

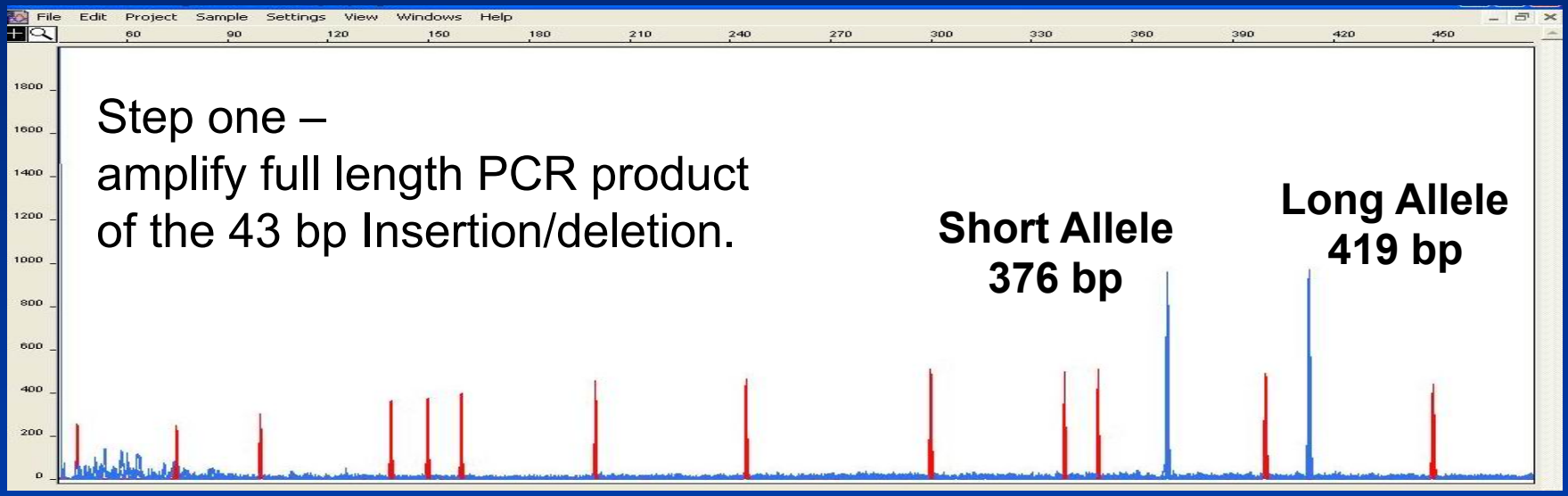
Table 5 Allele and genotype frequencies for binned 5HTTLPR genotypes

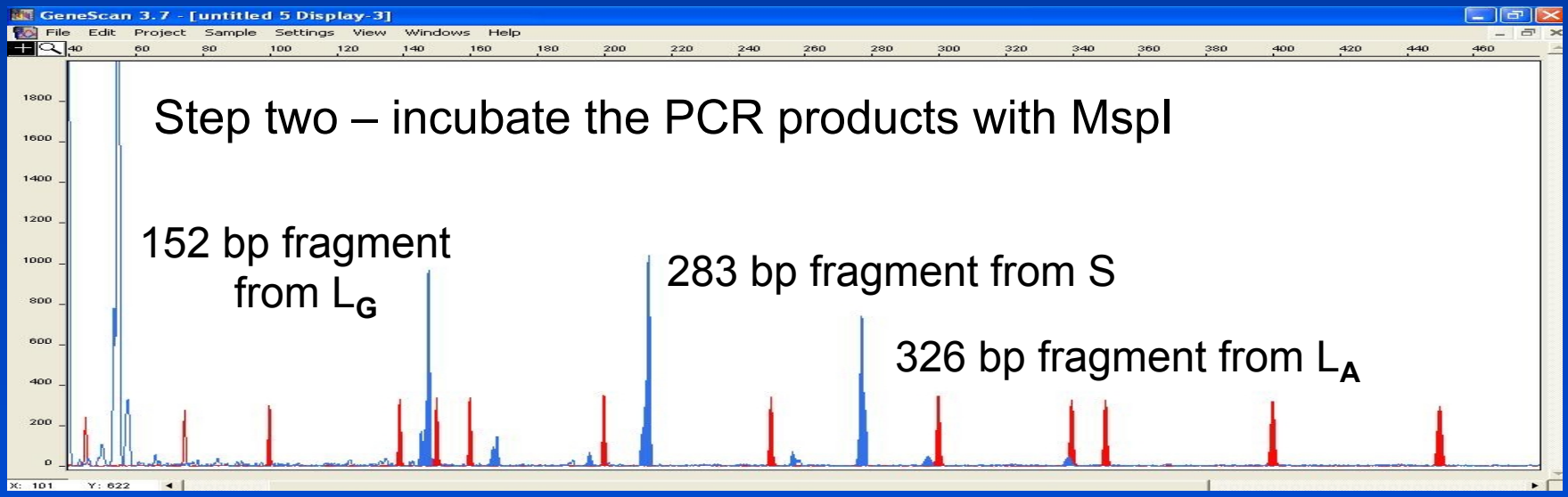
Population	Allele (*N*, %)		Genotype (*N*, %)		
	14R	16R	14R/14R	14R/16R	16R/16R
White ($n = 8215$)	7055 (0.43)	9375 (0.57)	1516 (0.18)	4023 (0.49)	2676 (0.33)
Black ($n = 3259$)	1712 (0.27)	4806 (0.74)	254 (0.08)	1204 (0.37)	1801 (0.55)
Native American ($n = 116$)	125 (0.54)	107 (0.46)	36 (0.31)	53 (0.45)	28 (0.24)
Asian ($n = 904$)	1210 (0.67)	598 (0.33)	405 (0.45)	400 (0.44)	99 (0.11)
Hispanic ($n = 2346$)	2386 (0.51)	2306 (0.49)	623 (0.27)	1140 (0.49)	583 (0.25)

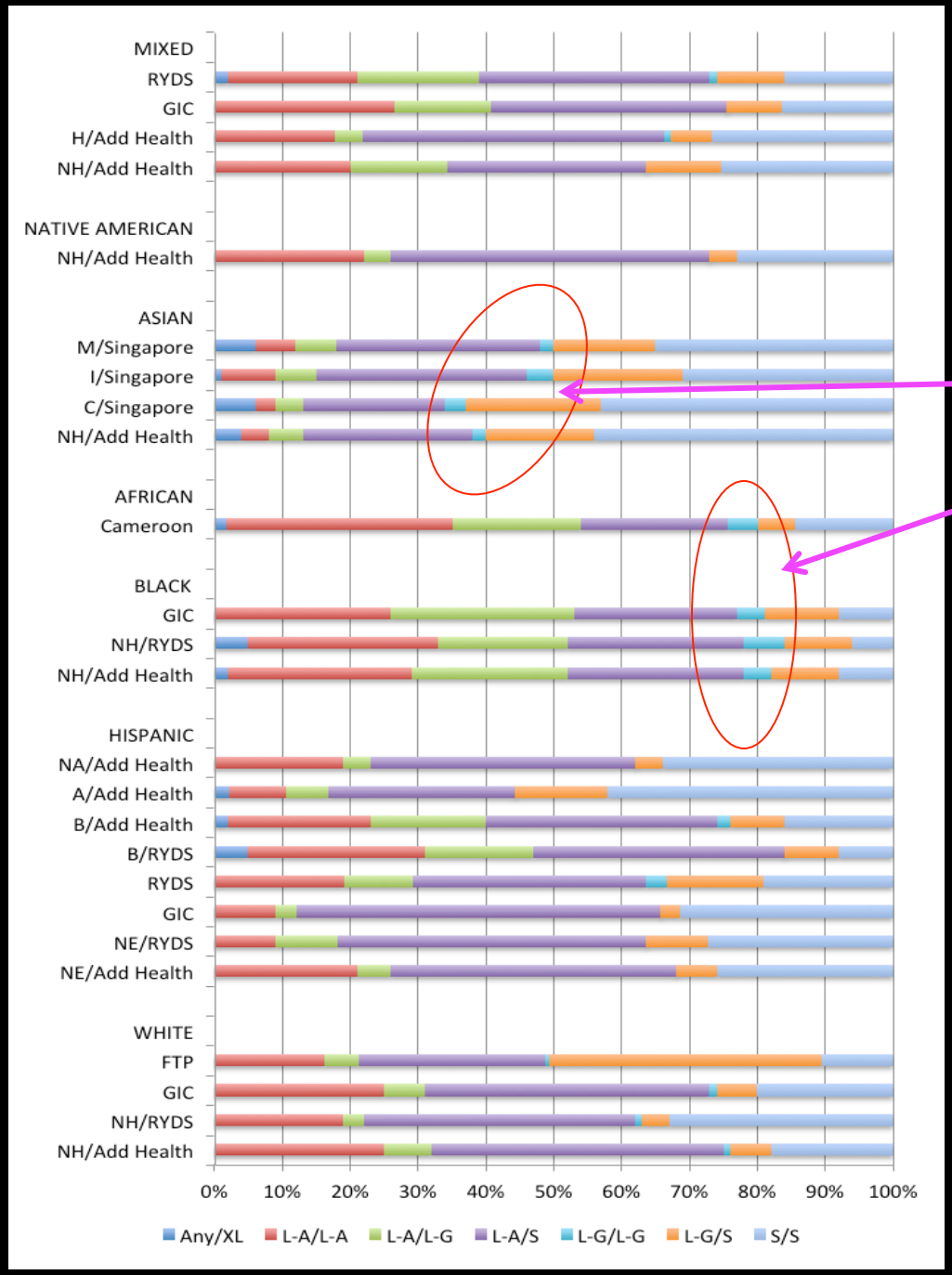
HWE White, $\chi2 = 0.0051$, df = 2, $p = 0.9974$; Black, $\chi2 = 6.8780$, df = 2, $p = 0.0321$; Native Americans, $\chi2 = 1.0638$, df = 2, $p = 0.5874$; Asians, $\chi2 = 0.0102$, df = 2, $p = 0.9949$; Hispanic, $\chi2 = 1.8016$, df = 2, $p = 0.4062$

N sample size, *R* repeat

5HTTLPR – "tri-allelic" determination of SNP rs25531







5HTTLPR + rs25531

Lg-Lg

Most prevalent in Asians and Blacks, *very low or not at all in Whites, Hispanics, and Native Americans*

Haberstick et al, under review

Simple Sequence Repeats in the National Longitudinal Study of Adolescent Health: An Ethnically Diverse Resource for Genetic Analysis of Health and Behavior

Brett C. Haberstick · Andrew Smolen · Gary L. Stetler · Joyce W. Tabor · Taylor Roy · H. Rick Casey · Alicia Pardo · Forest Roy · Lauren A. Ryals · Christina Hewitt · Eric A. Whitsel · Carolyn T. Halpern · Ley A. Killeya-Jones · Jeffrey M. Lessem · John K. Hewitt · Kathleen Mullan Harris

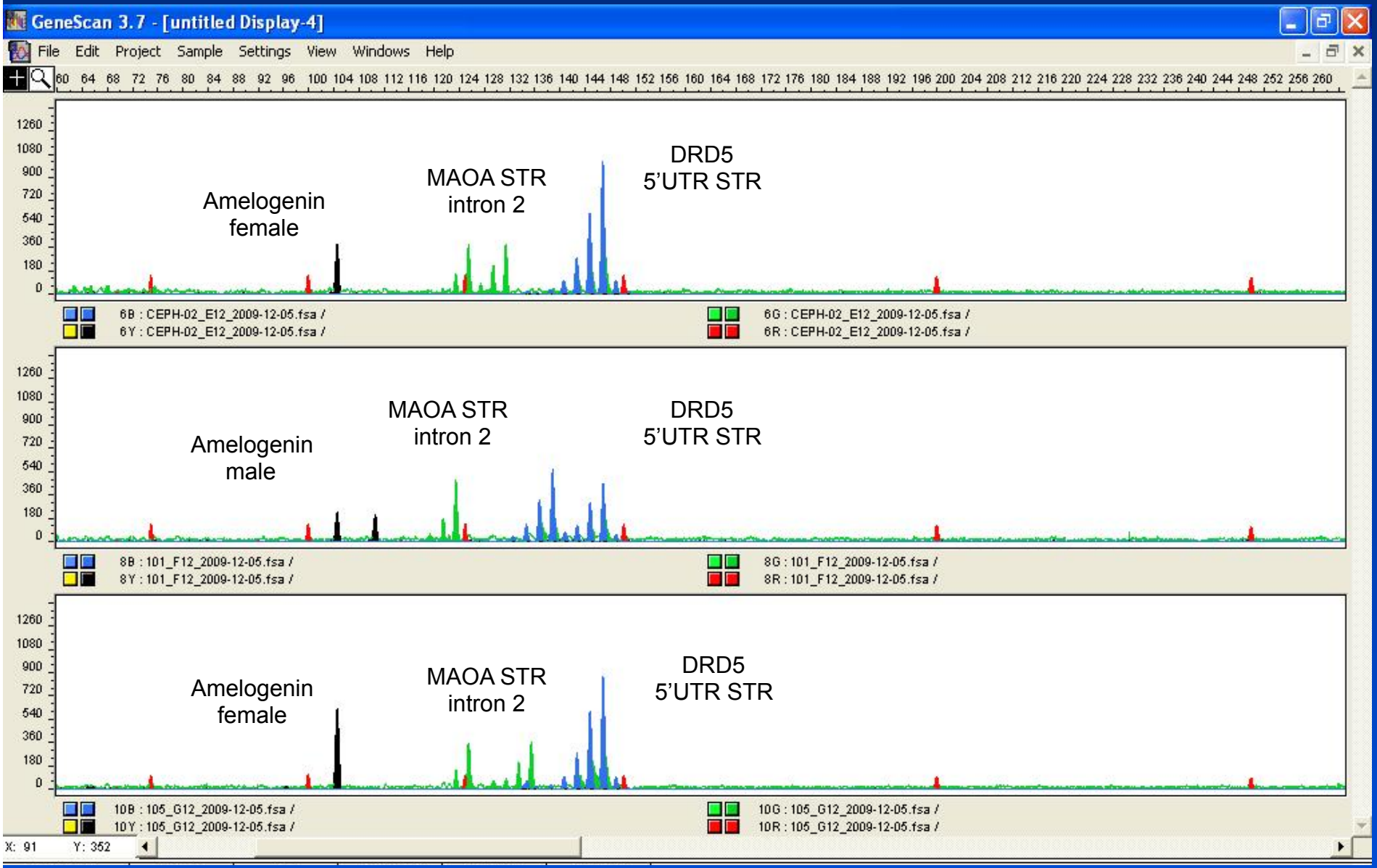
Table 6 Allele and genotype frequencies for binned 5HTTLPR genotypes

	Allele (*N*, %)		Genotype (*N*, %)		
Population	S'	L'	S'/S'	S'/L'	L'/L'
White ($n = 8183$)	8178 (0.50)	8188 (0.50)	2054 (0.25)	4071 (0.50)	2058 (0.25)
Black ($n = 3199$)	3036 (0.47)	3362 (0.53)	710 (0.22)	1616 (0.50)	873 (0.27)
Native American ($n = 116$)	137 (0.59)	95 (0.41)	39 (0.34)	59 (0.51)	18 (0.16)
Asian ($n = 867$)	1400 (0.81)	334 (0.19)	567 (0.04)	266 (0.33)	34 (0.63)
Hispanic ($n = 2330$)	2646 (0.57)	2014 (0.43)	767 (0.33)	1112 (0.48)	451 (0.19)

Reflects the reclassification of L-alleles based on rs25531 status. *HWE* White, $\chi2 = 0.1955$, df $= 2$, $p = 0.9068$; Black, $\chi2 = 0.5286$, df $= 2$, $p = 0.7677$; Native Americans, $\chi2 = 0.2383$, df $= 2$, $p = 0.8876$; Asians, $\chi2 = 0.1913$, df $= 2$, $p = 0.9087$; Hispanic, $\chi2 = 1.8245$, df $= 2$, $p = 0.4016$

N sample size, *R* repeat

Dinucleotide Repeat Genotyping on ABI 3130xl



Supplemental Table 9. Allele counts and frequencies (N, %) for DRD5 polymorphism by race/ethnicity.

	Race/Ethnicity				
Allele	White (N= 8046)	Black (N= 3167)	Native American (N= 115)	Asian (N= 879)	Hispanic (N= 2293)
124	--	--	--	--	1 (<0.01)
126	--	6 (<0.01)	--	2 (<0.01)	1 (<0.01)
128	4 (<0.01)	3 (<0.01)	--	--	1 (<0.01)
130	16 (<0.01)	304 (0.05)	--	2 (<0.01)	33 (0.01)
132	24 (<0.01)	35 (0.01)	--	2 (<0.01)	14 (<0.01)
134	285 (0.02)	38 (0.01)	--	17 (0.01)	42 (0.01)
136	364 (0.02)	115 (0.02)	4 (0.02)	62 (0.04)	95 (0.02)
138	1347 (0.08)	161 (0.03)	26 (0.11)	139 (0.08)	499 (0.11)
140	661 (0.04)	321 (0.05)	15 (0.07)	172 (0.10)	220 (0.05)
142	622 (0.04)	588 (0.09)	31 (0.13)	102 (0.06)	399 (0.09)
144	624 (0.04)	734 (0.12)	8 (0.03)	187 (0.11)	263 (0.06)
146	1009 (0.06)	1227 (0.19)	22 (0.10)	160 (0.09)	286 (0.06)
148	7172 (0.45)	1609 (0.25)	72 (0.31)	681 (0.39)	1475 (0.32)
150	2063 (0.13)	325 (0.05)	35 (0.15)	133 (0.08)	762 (0.17)
152	1315 (0.08)	503 (0.08)	10 (0.04)	82 (0.05)	342 (0.07)
154	440 (0.03)	119 (0.02)	4 (<0.01)	7 (<0.01)	91 (0.02)
156	116 (0.01)	37 (0.01)	1 (<0.01)	6 (<0.01)	32 (0.01)
158	20 (<0.01)	9 (<0.01)	--	3 (<0.01)	5 (<0.01)
160	7 (<0.01)	21 (<0.01)	--	1 (<0.01)	2 (<0.01)
162	--	6 (<0.01)	--	--	--
164	--	18 (<0.01)	--	--	--
166	1 (<0.01)	39 (<0.01)	1 (<0.01)	--	6 (<0.01)
168	1 (<0.01)	62 (<0.01)	1 (<0.01)	--	11 (<0.01)
170	--	23 (<0.01)	--	--	4 (<0.01)
172	--	27 (<0.01)	--	--	2 (<0.01)
174	1 (<0.01)	4 (<0.01)	--	--	--
176	--	--	--	--	--

The size of the repeat itself may not matter
The location and function are important

Position of polymorphism could have effects when in:

1. Regulatory regions : influences the amount of protein made.
2. Coding regions : influences the type of protein made.

Non-synonymous ⟷
- Nonsense mutations: causes premature stop signal
- Missense mutations: changes in protein sequence

Synonymous "silent" → Mutation *does not* alter protein, could affect mRNA stability & translation

Human Catechol-*O*-Methyltransferase Haplotypes Modulate Protein Expression by Altering mRNA Secondary Structure

A. G. Nackley,[1] S. A. Shabalina,[2] I. E. Tchivileva,[1] K. Satterfield,[1] O. Korchynskyi,[3] S. S. Makarov,[4] W. Maixner,[1] L. Diatchenko[1]*

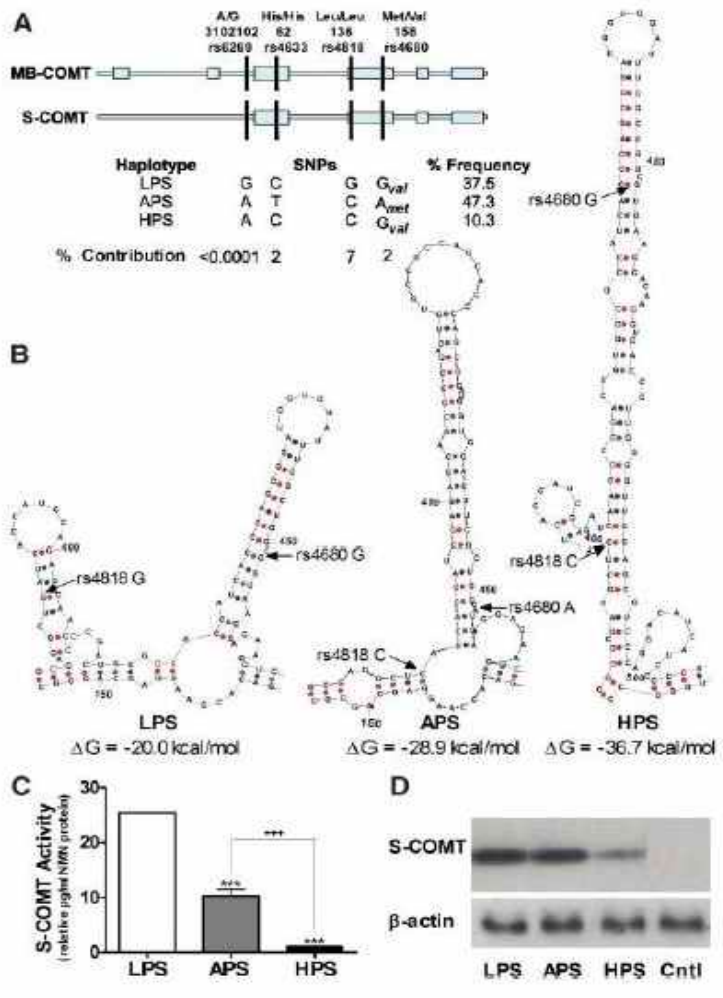


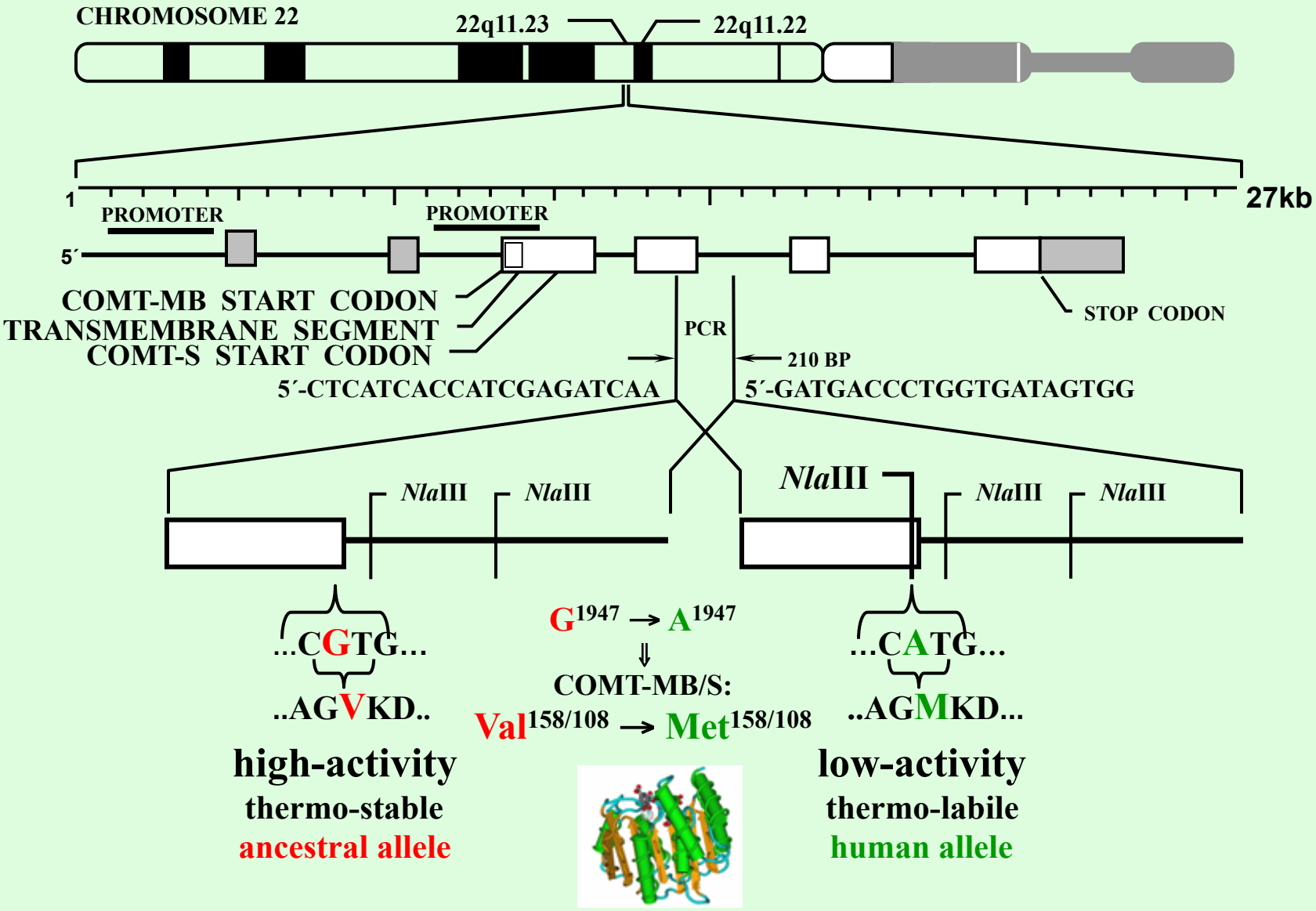
Fig. 1. Common haplotypes of the human *COMT* gene differ with respect to mRNA secondary structure and enzymatic activity. (**A**) A schematic diagram illustrates *COMT* genomic organization and SNP composition for the three haplotypes. Percent frequency of each haplotype in a cohort of healthy Caucasian females, and percent independent SNP contribution to pain sensitivity, are indicated. (**B**) The local stem-loop structures associated with each of the three haplotypes are shown. Relative to the LPS and APS haplotypes, the HPS local stem-loop structure had a higher folding potential. (**C** and **D**) The LPS haplotype exhibited the highest, while the HPS haplotype exhibited the lowest enzymatic activity and protein levels in cells expressing COMT. $^{***}P < 0.001$, ≠ LPS. $^{+++}P < 0.001$, ≠ APS.

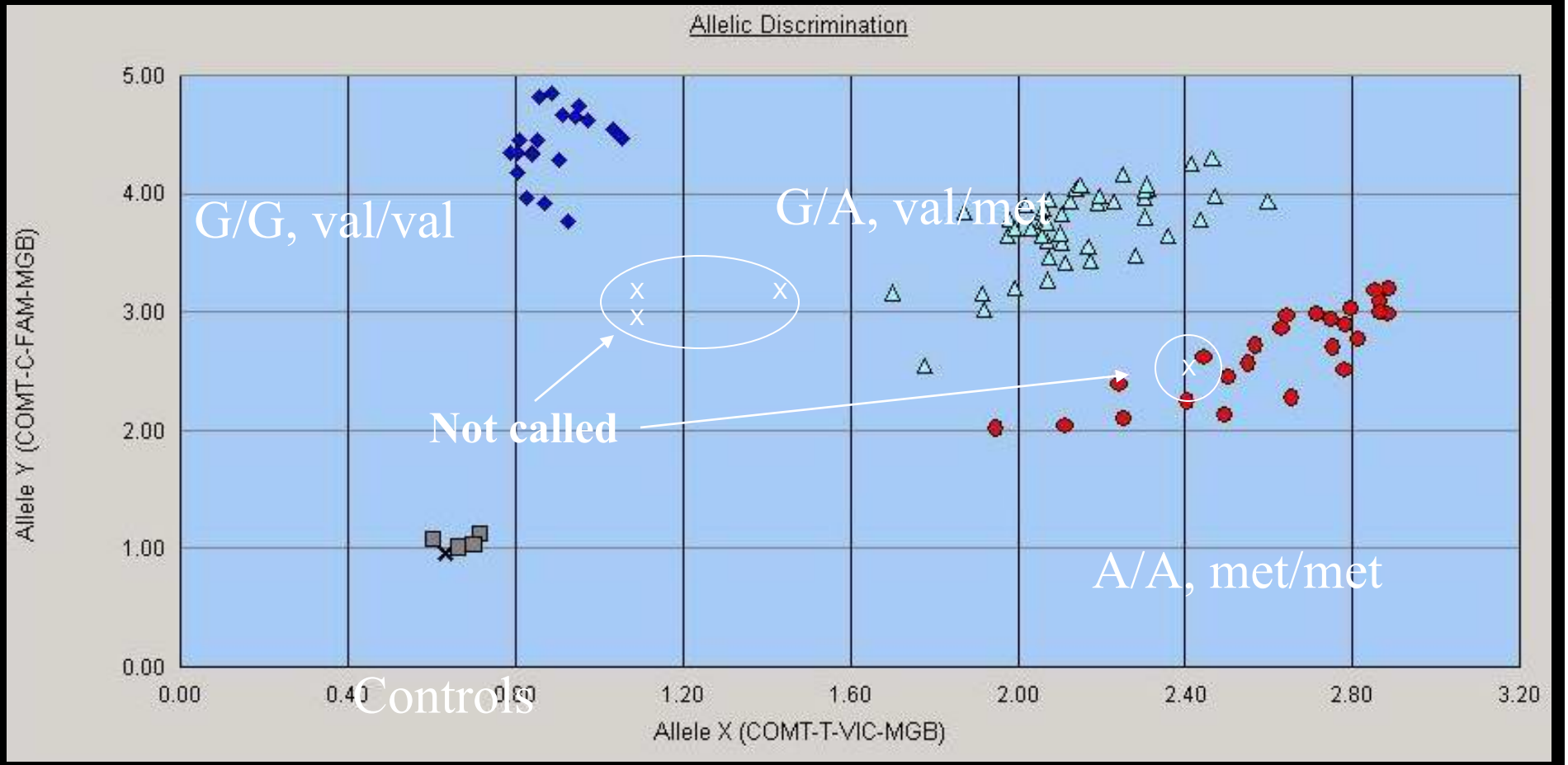
~ COMT accounts for more than 60% of DA degradation in PFC but <15% in striatum (Karoum et al 1994)

~ COMT inhibitors enhance working memory (Khromova et al 1995; Liljequist et al 1997)

~ minimal role of prefrontal dopamine transporters (Sesack et al 1998, Lewis et al 2001, Moron et al 2002, Mazei et al 2002)

The COMT Gene



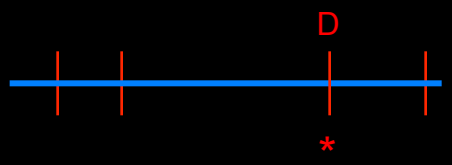

Allelic Discrimination
Allele Y (COMT-C-FAM-MGB)
Allele X (COMT-T-VIC-MGB)
5.00
4.00
3.00
2.00
1.00
0.00
0.00
0.40
0.80
1.20
1.60
2.00
2.40
2.80
3.20
G/G, val/val
G/A, val/met
A/A, met/met
Not called
Controls

Testing Genetic Associations in Add Health

- Add Health sampling design allows different types of association testing.
- Simplest type is correlating phenotype with genotype.
- Other types:
  - Candidate gene: common practice in medical genetics
  - Pharmacogenetics: genotyping clinically relevant samples
  - Genome-wide association: exhaustive characterization of many SNPs to allow a search across the genome.

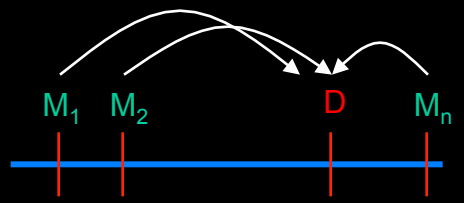
Three Types of Genetic Associations

1. Direct Association



Measure disease relevance (*) directly, ignoring correlated markers nearby

2. Indirect Association & LD



Assess trait effects on D via correlated markers (M_i) rather than susceptibility/etiologic variants.

3. Spurious Association

Apparent association not related to genetic etiology

Association Studies: Track Record

Pubmed: February 2011. "*Genetic association*" gives 77,861 hits
— in Mar 2006: 36,908

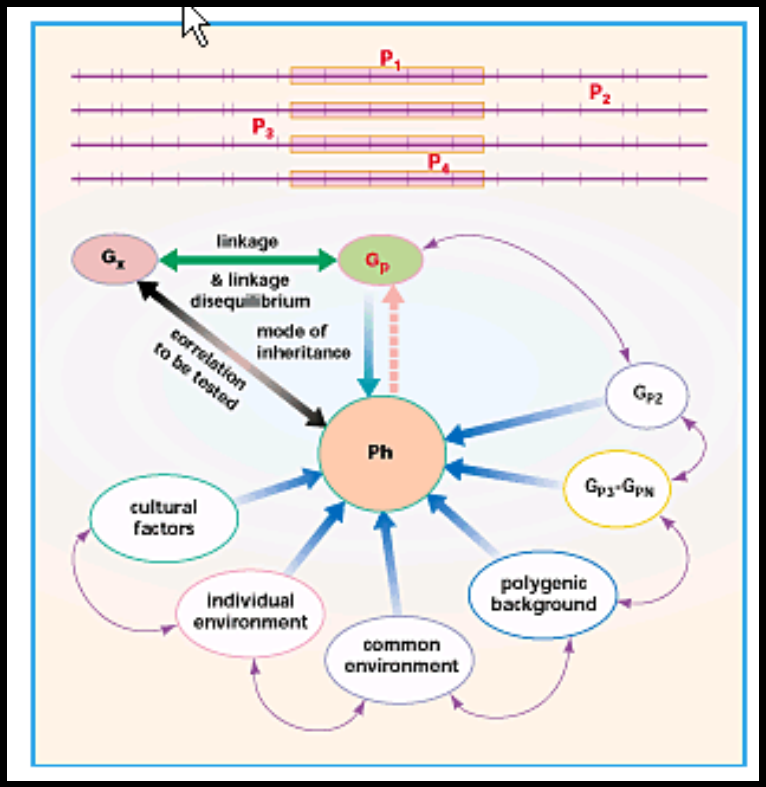
Question: How many are real?

Answer: < 1%

- Claims of "replicated genetic association" → 1602 (0.02%)
- Claims of "validated genetic association" → 780 hits (0.01%)

Reasons for limited success with Association studies

- Small sample sizes, limited power
- Phenotypes are complex and not measured well. Candidate genes thus difficult to choose.
- Allelic/genotypic contributions are complex. Even true associations are difficult to identify.
- Population stratification is thought to be a primary reason



Race/Ethnicity	N	%
Mexico	1,767	8.5
Cuba	508	2.5
Central-South America	647	3.1
Puerto Rico	570	2.8
China	341	1.7
Philippines	643	3.1
Other Asia	601	2.9
Black (Africa/Afro-Caribbean)	4,601	22.2
Non-Hispanic White (Eur/Canada)	10,760	52.0
Native American (non-Hispanic)	248	1.2
Total N	**20,686**	**100.0**

Detecting and Controlling for Population Stratification

- Population stratification leads to spurious associations
- Occurs when there are genetic differences in allele frequencies AND genetic differences in outcomes.

- To control for the possibility: use FAMILY-BASED CONTROLS found in the Sibling Pairs sample OR
- Use Ancestral Informative Markers (AIM) genotyped at Wave IV

Current Association Study Challenges – Which association approach is appropriate??

Candidate gene or genome-wide?

Candidate gene

- Hypothesis-driven
- Low-cost: small genotyping requirements
- Multiple-testing less important
  - Possible many misses, fewer false positives

Genome-wide screen

- Hypothesis-free
- High-cost: large genotyping requirements
- Multiple-testing issues
  - Possible many false positives, fewer misses

Current Association Study Challenges – What constitutes a replication??

Replicating association results in different laboratories is often seen as most compelling piece of evidence for 'true' finding

BUT…. in any sample, we measure

Multiple traits
Multiple genes
Multiple markers in genes

and we analyse all this using multiple statistical tests

Criteria of what is and is not a replication needs to be determined in advance

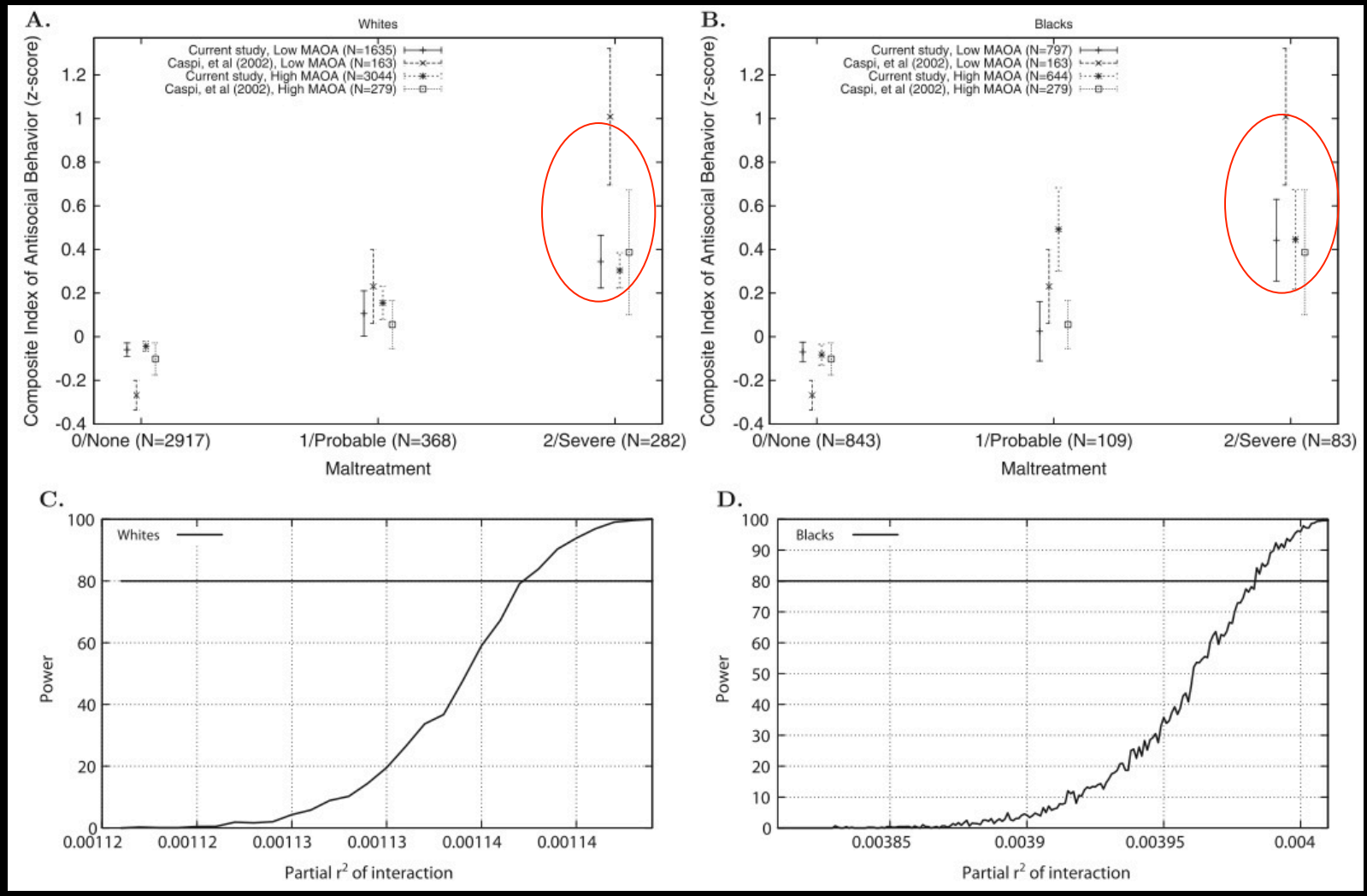
MAOA Genotype, Childhood Maltreatment, and Their Interaction in the Etiology of Adult Antisocial Behaviors

Brett C. Haberstick, Jeffrey M. Lessem, John K. Hewitt, Andrew Smolen, Christian J. Hopfer, Carolyn T. Halpern, Ley A. Killeya-Jones, Jason D. Boardman, Joyce Tabor, Ilene C. Siegler, Redford B. Williams, and Kathleen Mullan Harris

Biological Psychiatry, 2014

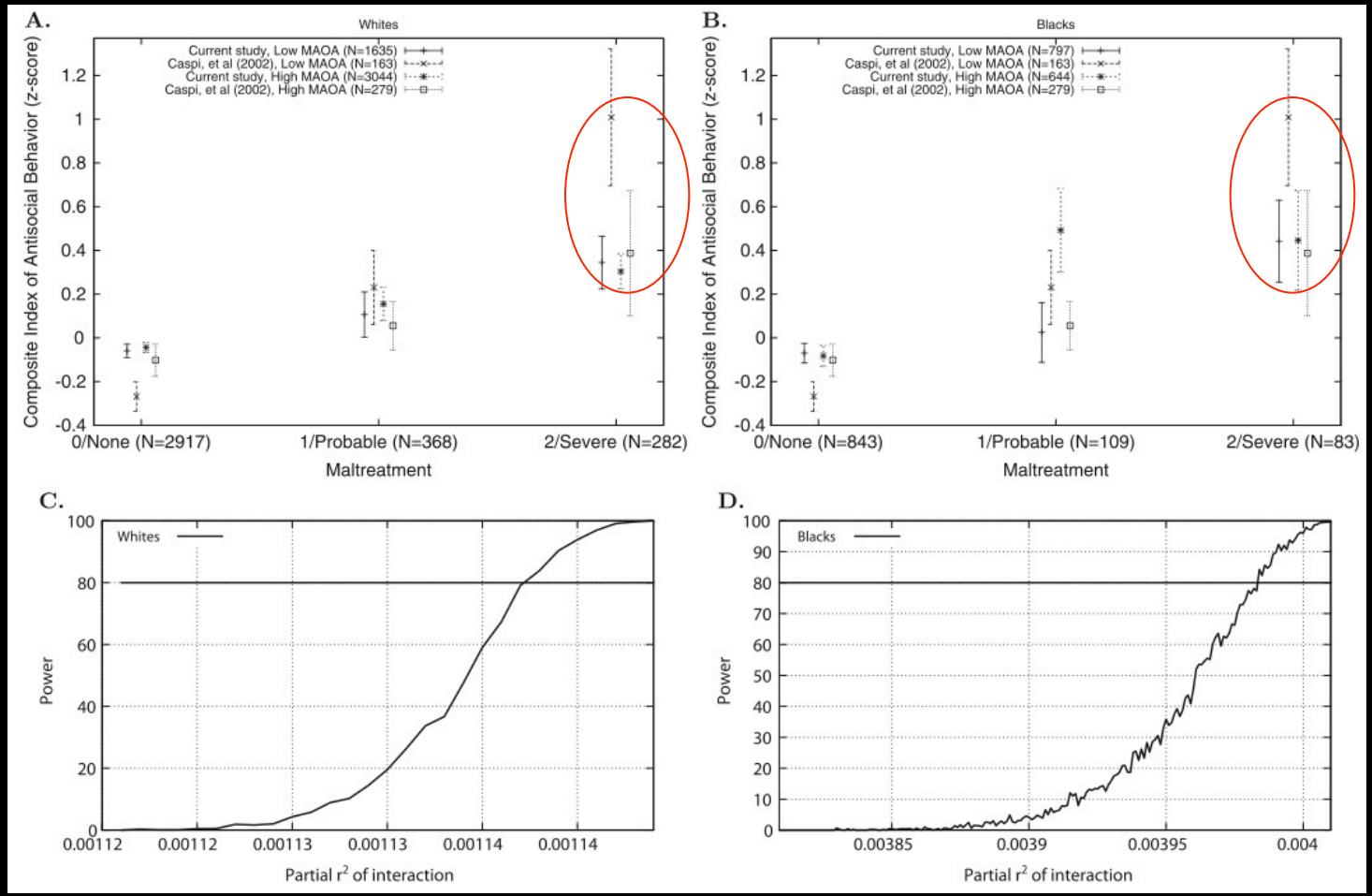
Sample: Adults between ages 24-34. Total sample size 3356 (White), 960 (Black)
Measures: CD: Lifetime conduct disorder (4 waves, repeated measures), violent convictions (2 waves, repeated measures), disposition towards violence;
Maltreatment: physical and sexual abuse, neglect, visits from and removal by social services (2 waves)
Analysis Approach: Direct replication
Demographics: CD: 66.0% no offending, 32.7% low offending, 1.3% high offending;
Maltreatment: (White sample) 81.8% no maltreatment, 10.3% 'probable' maltreatment; 7.9% 'severe' maltreatment. Similar frequencies in our Black sample.

Key Findings: (1) Higher rates of conduct disorder, antisocial personality disorder, and a disposition towards violence was observed among those with more experiences of maltreatment; (2) No main effect of genotype, (3) No significant moderation of long-term relationship by low MAOA functioning.



Key Findings: (1) Higher rates of conduct disorder, antisocial personality disorder, and a disposition towards violence was observed among those with more experiences of maltreatment; (2) No main effect of genotype, (3) No significant moderation of long-term relationship by low MAOA functioning.

Power of study could detect an effect size as small as 0.001138 (White sample)

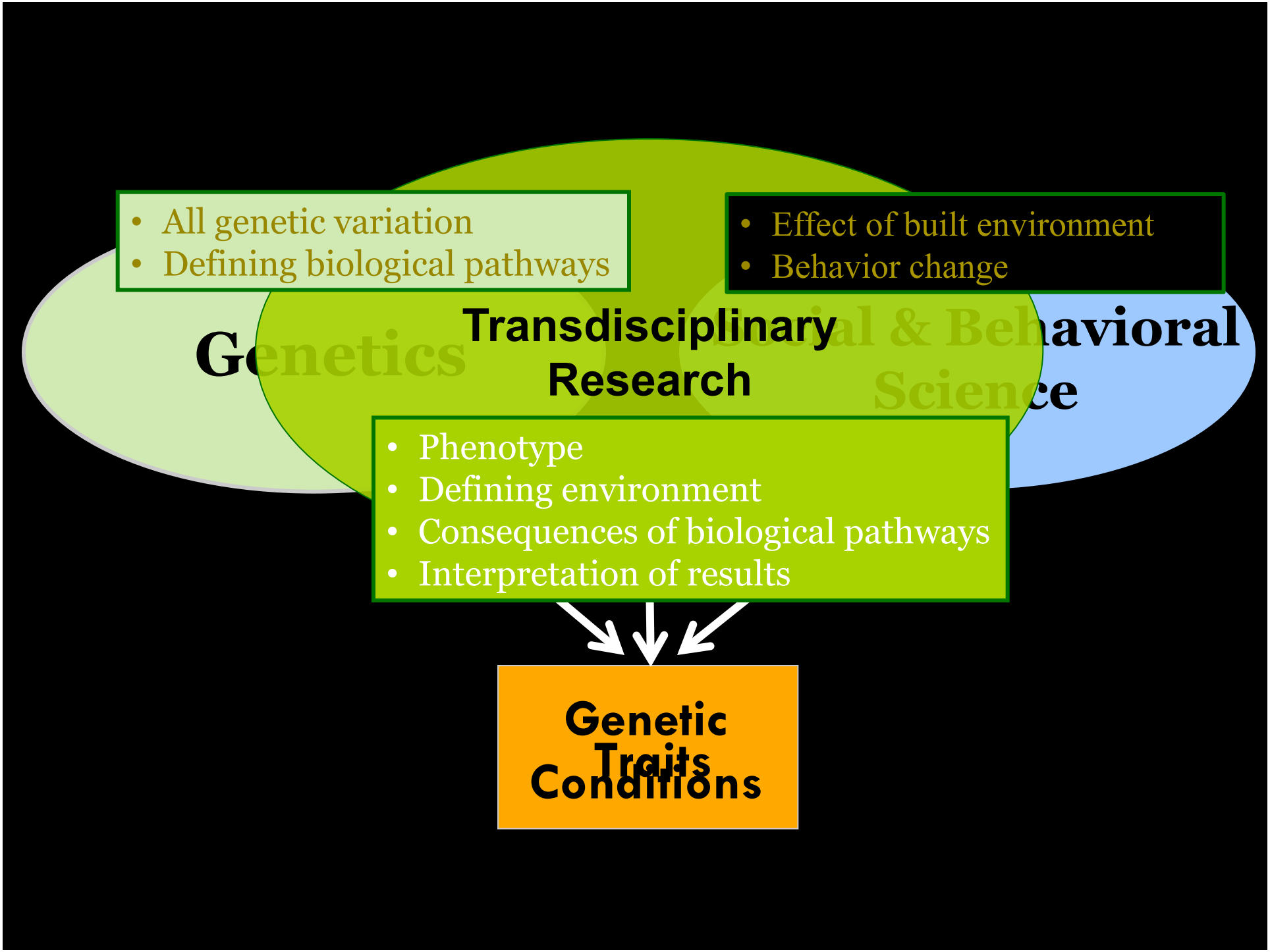


Current Association Study Challenges - Replication

Replication Outcome: Association to same trait, different gene
Explanation: Genetic heterogeneity
Replication Outcome: Association to same trait, same gene, different SNPs
Explanation: Allelic heterogeneity
Replication Outcome: Association to different but correlated phenotype(s)
Explanation: Phenotypic heterogeneity
Replication Outcome: No association at all
Explanation: Insufficient sample size (low statistical power)

Final Thoughts……

- Add Health presents an optimal study design for examining genotype-phenotype relationship.
- Collecting DNA on a large-scale is feasible and can be done.
- Watch for population stratification and derive empirical p-values to correct for multiple testing.
- Establish a collaborative relationship with someone outside your field of expertise - Social Scientist & a Geneticist.


• All genetic variation
• Defining biological pathways
• Effect of built environment
• Behavior change
Genetics
Transdisciplinary Research
Social & Behavioral Science
• Phenotype
• Defining environment
• Consequences of biological pathways
• Interpretation of results
Genetic Traits Conditions


GOOD LUCK
45
M.P.H.

Special Thanks to:

Kathie Mullan Harris, Add Health Study Director
Andrew Smolen, IBG, Boulder, Colorado
John K. Hewitt. IBG, Boulder, Colorado

Brett C. Haberstick
brett.haberstick@colorado.edu

ibgwww.colorado.edu
www.cpc.unc.edu/projects/addhealth

QUESTIONS? COMMENTS?