Impact of Copy number Variations in Genome organization, Diseases and Evolution

Nallur B. Ramachandra

Professor and Chairman

Department of Studies in Genetics and Genomics

University of Mysore, Manasagangotri, Mysuru - 570 006

nallurbr@gmail.com, ramachandra@genetics.uni-mysore.ac.in

www.ramachandralab.com

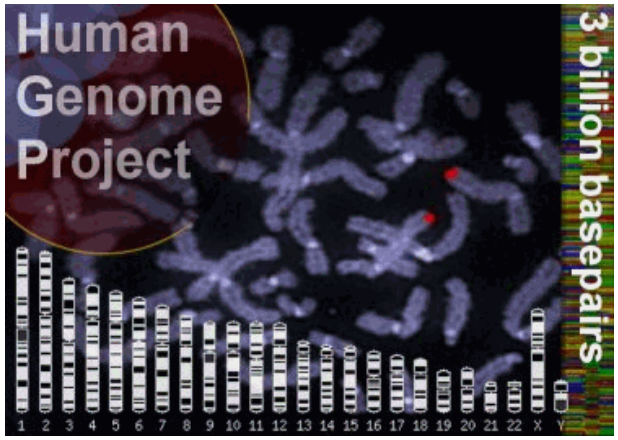
Outline

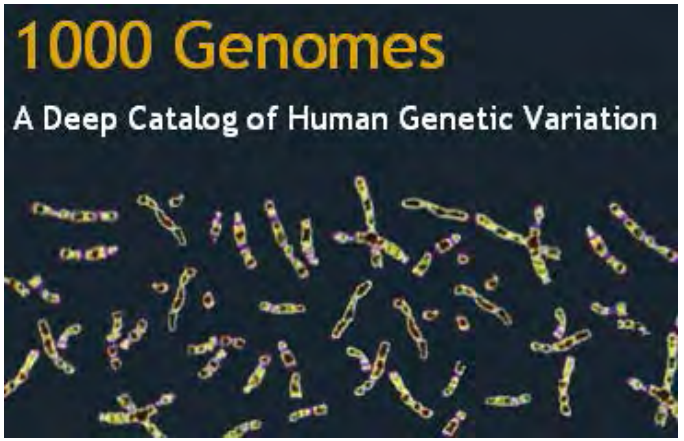
1) Genomic variations
2) CNVs in Genome organization
3) CNVs in disease risk genes in normal cohorts
4) CNVs involvement in Evolution
5) Conclusions

1) Genomic variations

What is a genome?

Exploring the Human Genome





Sanger sequencing, targeted genotyping

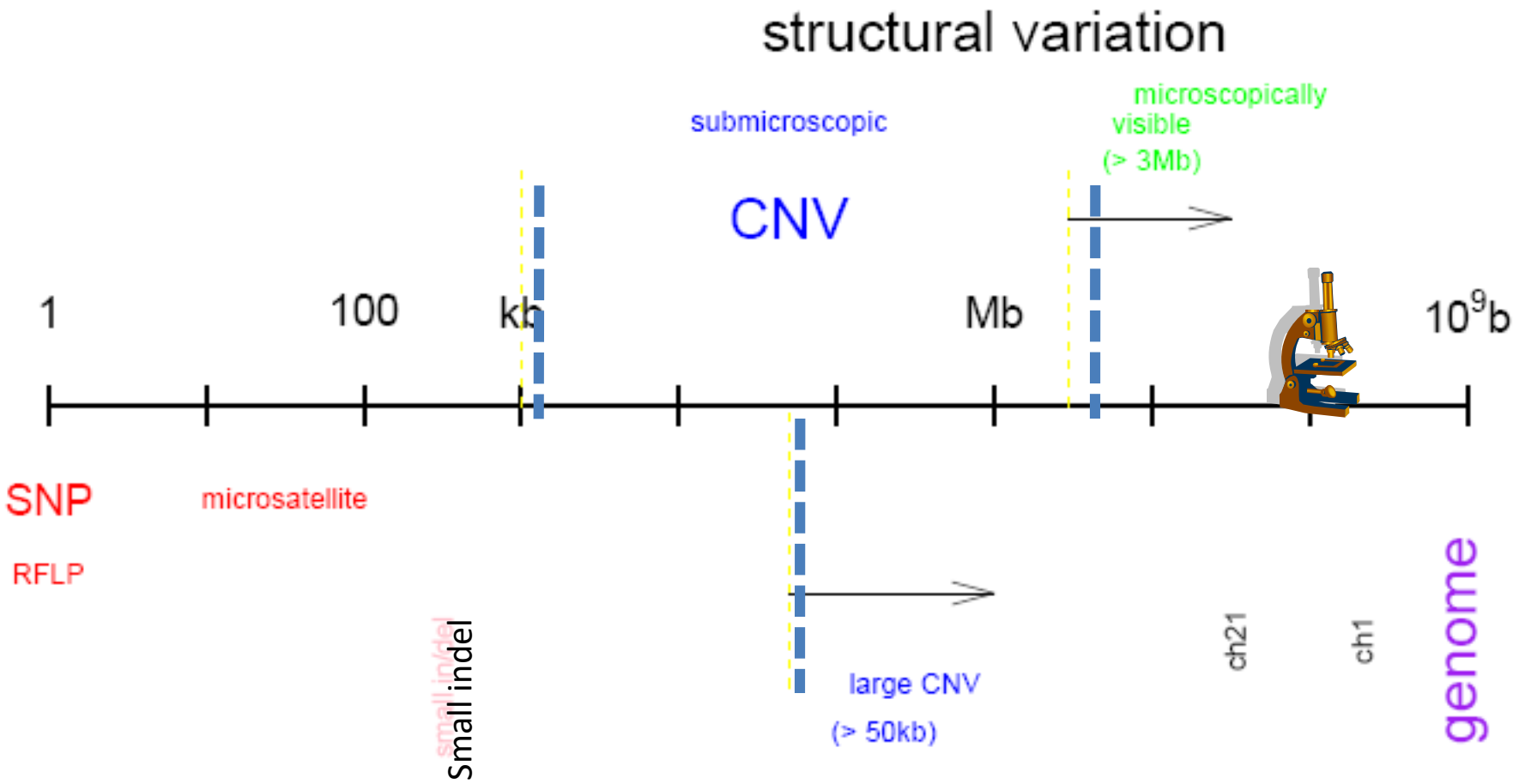
Genome-wide genotyping (GWAS)

Exome sequencing

Genome sequencing

99.6% inter-individual identity (yet 4 millions differences)

Genomic variations



Genomic variations

- **Normal or wild type**
  - **is the most frequent in a population**

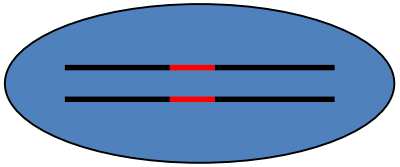
- **Polymorphism**
  - **is the variant if its frequency is › 1 % in the population**
  - **(formerly:having no effect on phenotype)**

- **Mutation**
  - **is the variant if its frequency is ‹ 1 % in the population**
  - **(formerly: disease causing negative connotation)**

- **Increased by – Mutation and Sexual reproduction**

Copy Number Variations
(CNVs)

Normal cell



CN=2

Homologous repeats

Segmental duplications

Chromosomal rearrangements

Duplicative transpositions

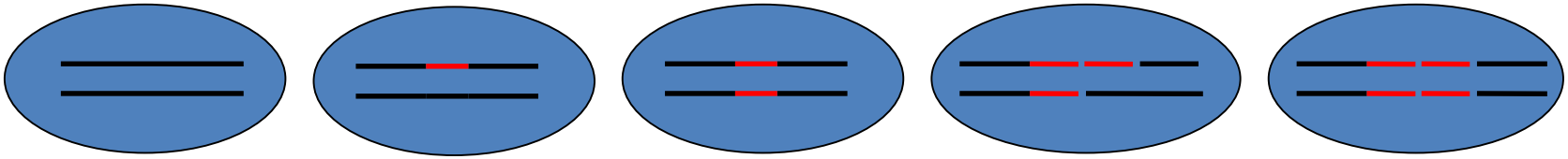
Non-allelic recombinations

……

Disease cells

deletion

duplication



CN=0 CN=1 CN=2 CN=3 CN=4

2) CNVs in Genome organization

Genomes used for CNVs analysis

1715 individuals involving:

- **38 normal Individuals from India - our lab**
- **270 HapMap samples**
  - **CEU (CEPH collection, Central Europe)**
  - **CHB (Han Chinese, Beijing, China)**
  - **JPT (Japanese Tokyo, Japan) and**
  - **YRI (Yoruba in Ibadan, Nigeria)**
- **31 Tibetan samples**
- **155 Chinese samples**
- **472 of Ashkenazi Jews replicate 1**
- **480 of Ashkenazi Jews replicate 2**
- **204 individuals from Taiwan**
- **55 from Australia and**
- **64 from New World population (Totonacs and Bolivians)**

Source of Genomes

- The raw, unprocessed data from Affymetrix Genome Wide SNP 6.0 Array for all the 11 populations obtained from the ArrayExpress Archive of the EBI.

Genotyping Platform

- Affymetrix Genome-wide Human SNP Array 6.0 chip
- Affymetrix CytoScan® High-Density (HD) Array having 1.8 million and 2.6 million combined SNP and CNV markers

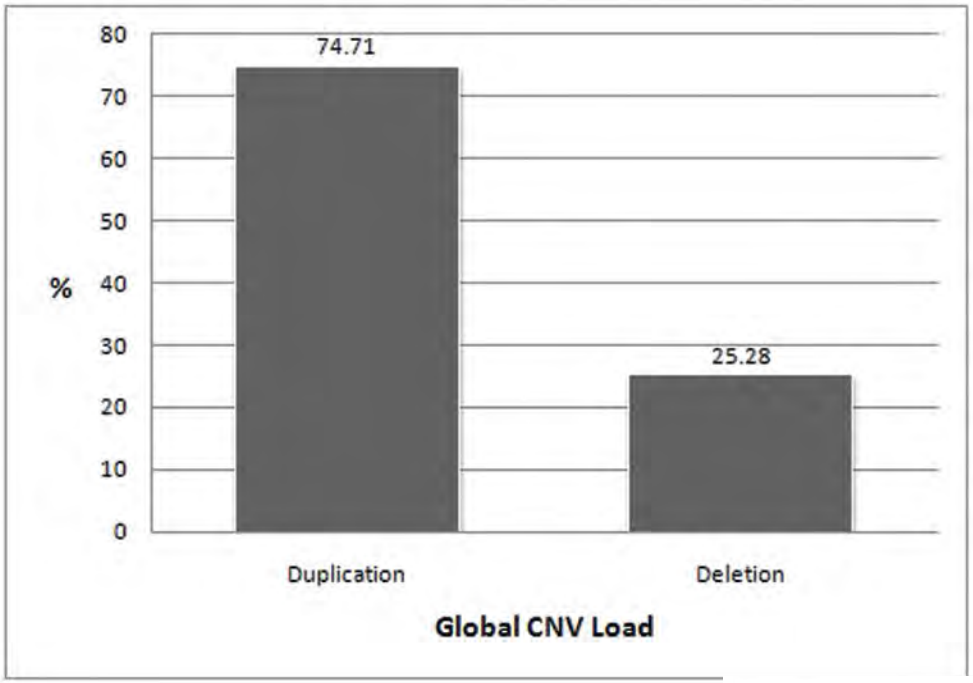
Algorithms for CNVs assessment

- BirdSuite
- Canary
- Genotyping Console
- CNVFinder
- SVS Golden Helix Ver. 7.2
- Web gestalt
- EnrichR
- Genome Decoration Page (GDP)
- IPA software
- HD-CNV

Distribution of CNVs

Populations	Individuals Assessed	Total Size (Mb)
HapMap-YRI-Africa	90	3.07±1.5
HapMap-CEU-Europe	90	5.95±3.0
Ashkenazi Jews I	464	7.83±3.9
Ashkenazi Jews II	480	7.32±3.6
HapMap-CHB-China	44	3.02±1.5
China	155	6.19±3.09
Tibet	31	5.5±2.7
India	38	8.9±2.2
HapMap-JPT-Japan	45	3.76±1.8
Australia	53	9.82±4.9
New World	41	13.05±6.5
Taiwan	184	10.24±5.1

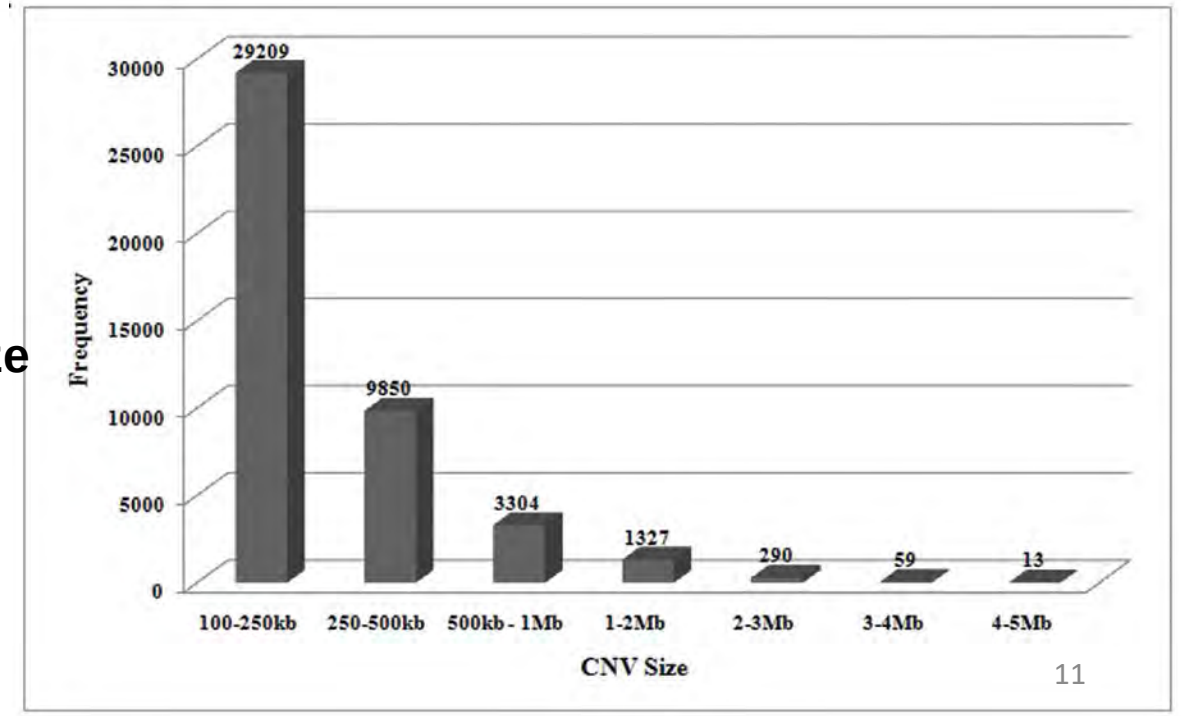
Veerappa et al., *PLOS ONE,* 2013



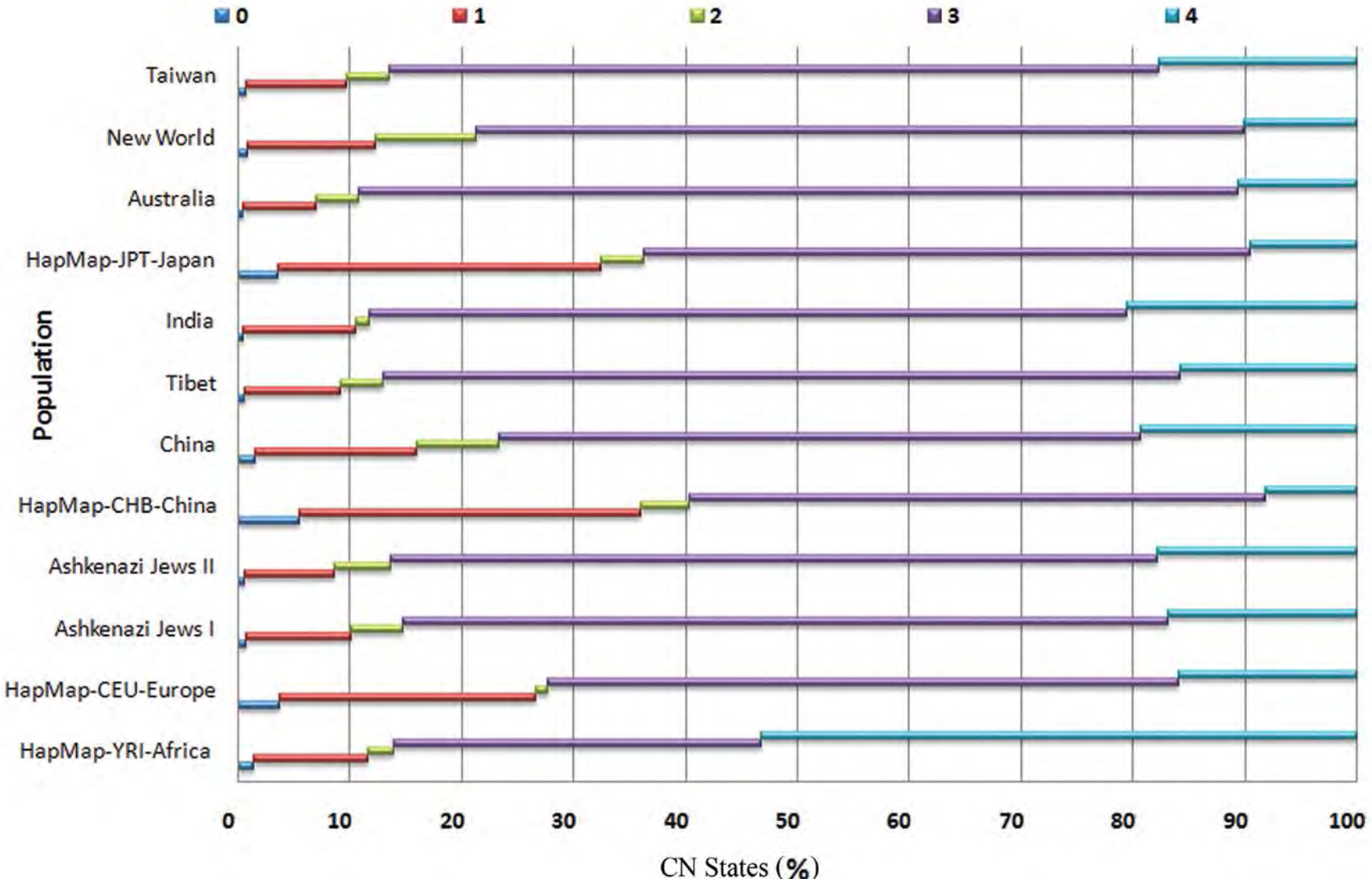
CNV Load

CNV Size

- More CNVs within 500kb
- Declined with increase in size



Allele state Variation



Veerappa et al., *PLOS ONE*, 2013

CNV Map

44,109 CNVs in
1,715 genomes of
12 populations

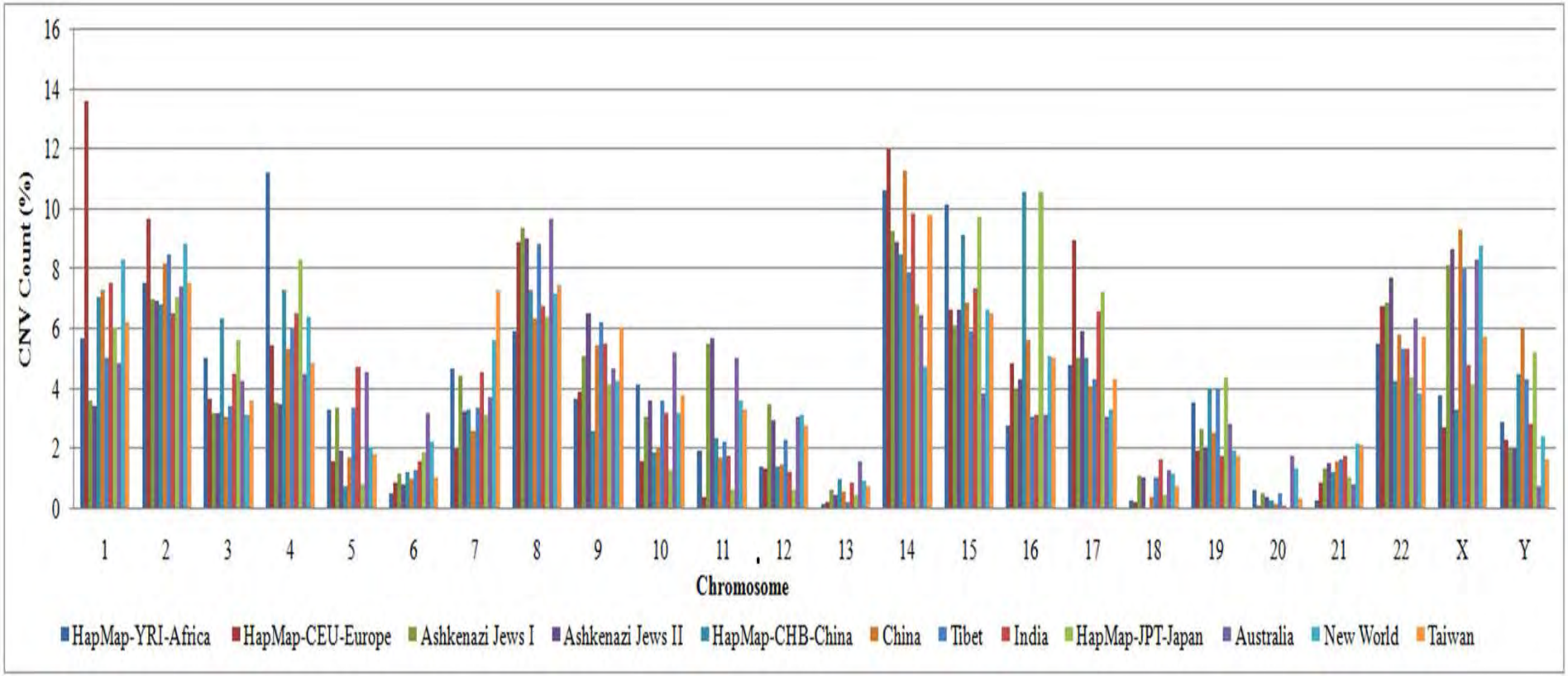
Mean size of CNVs 7±3Mb

Total - 1,26,190 genes

1,329 genes/ genome

Veerappa et al., *PLOS ONE*, 2013

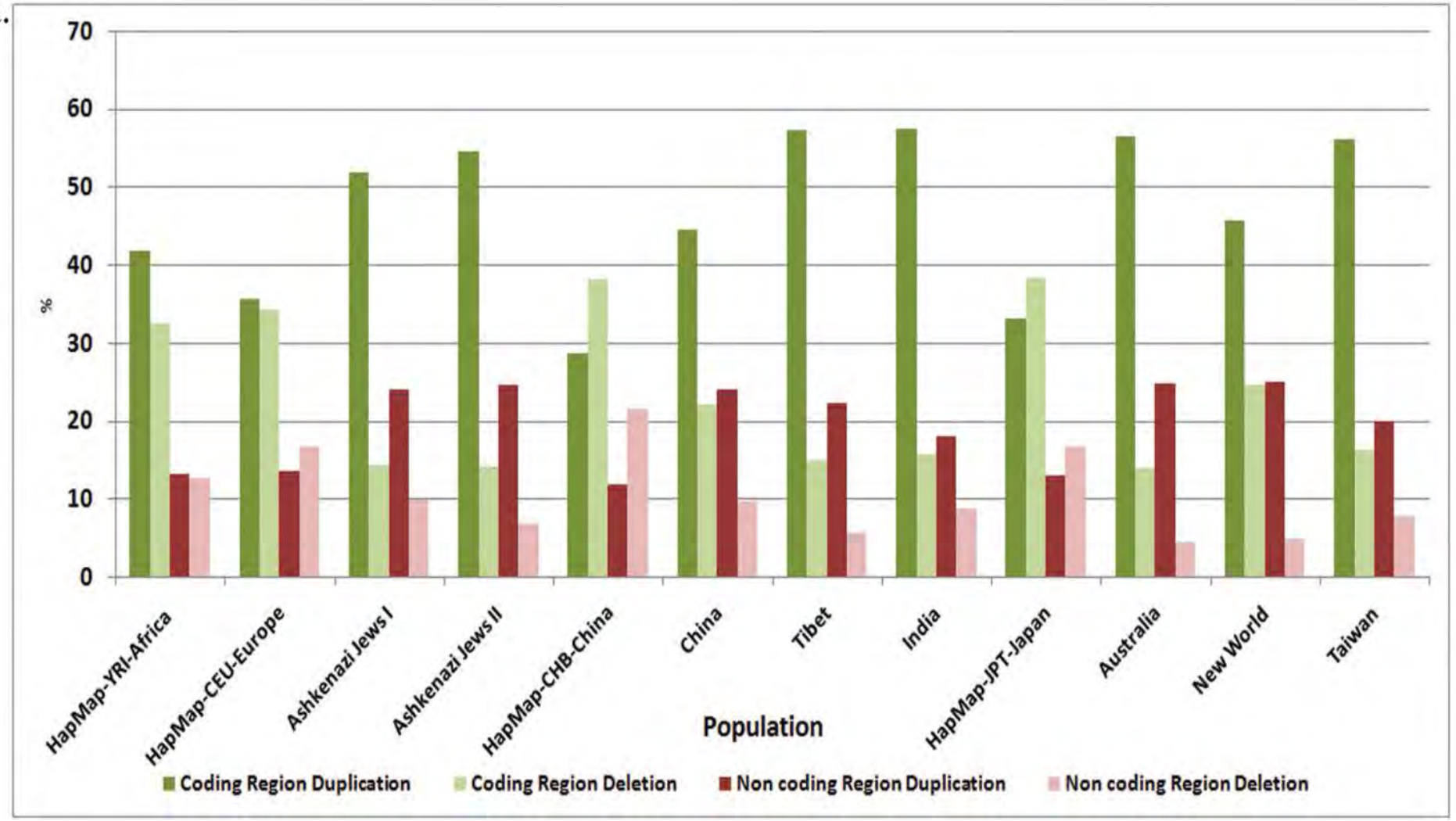
Chromosome-wise CNV Count



- More CNVs in chromosomes 14, 8, 2 and 15 (~8%)
- Less CNV in chromosomes 13, 20 and 18

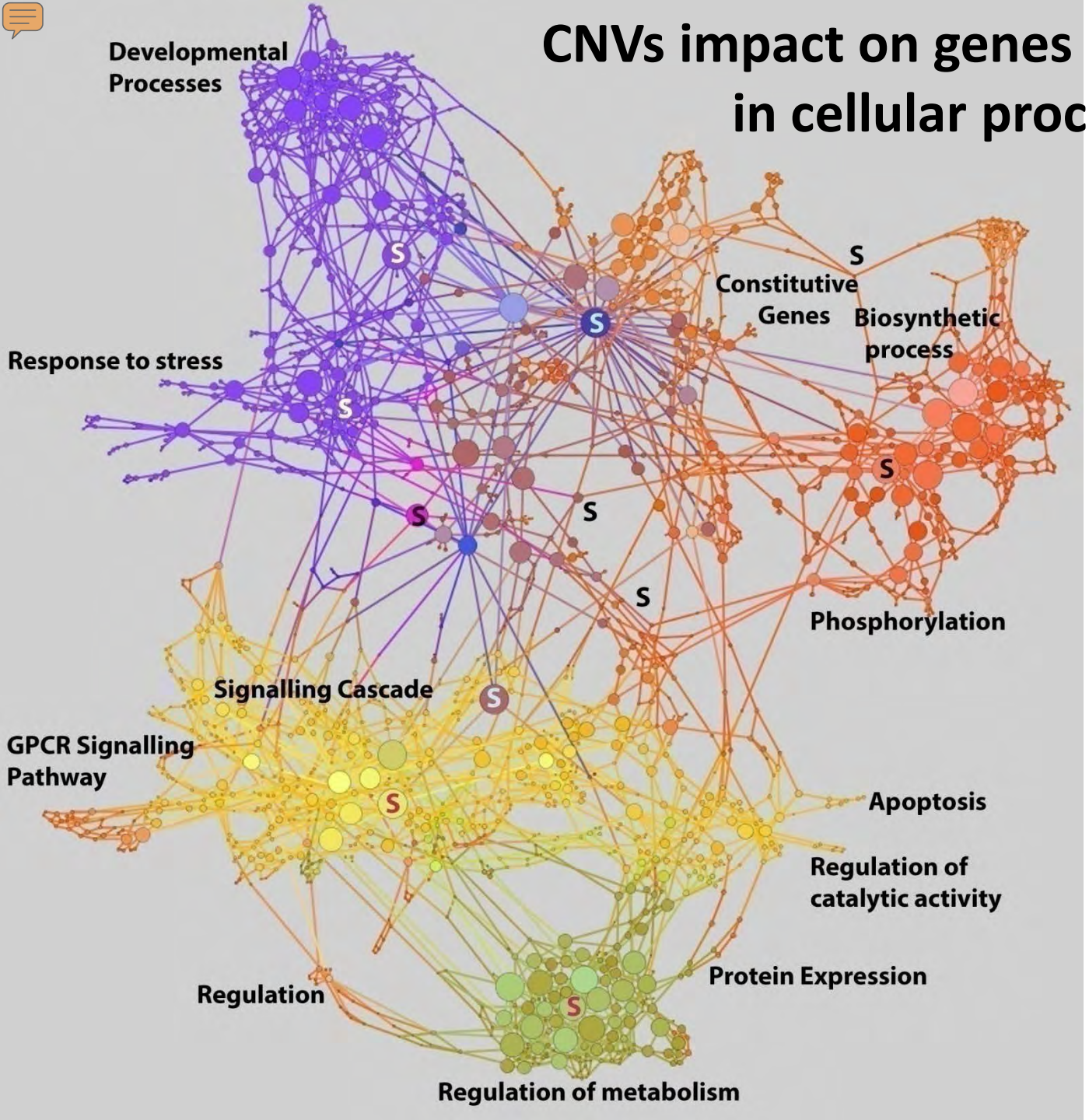
CNVs in Coding and Non-Coding Regions

a.

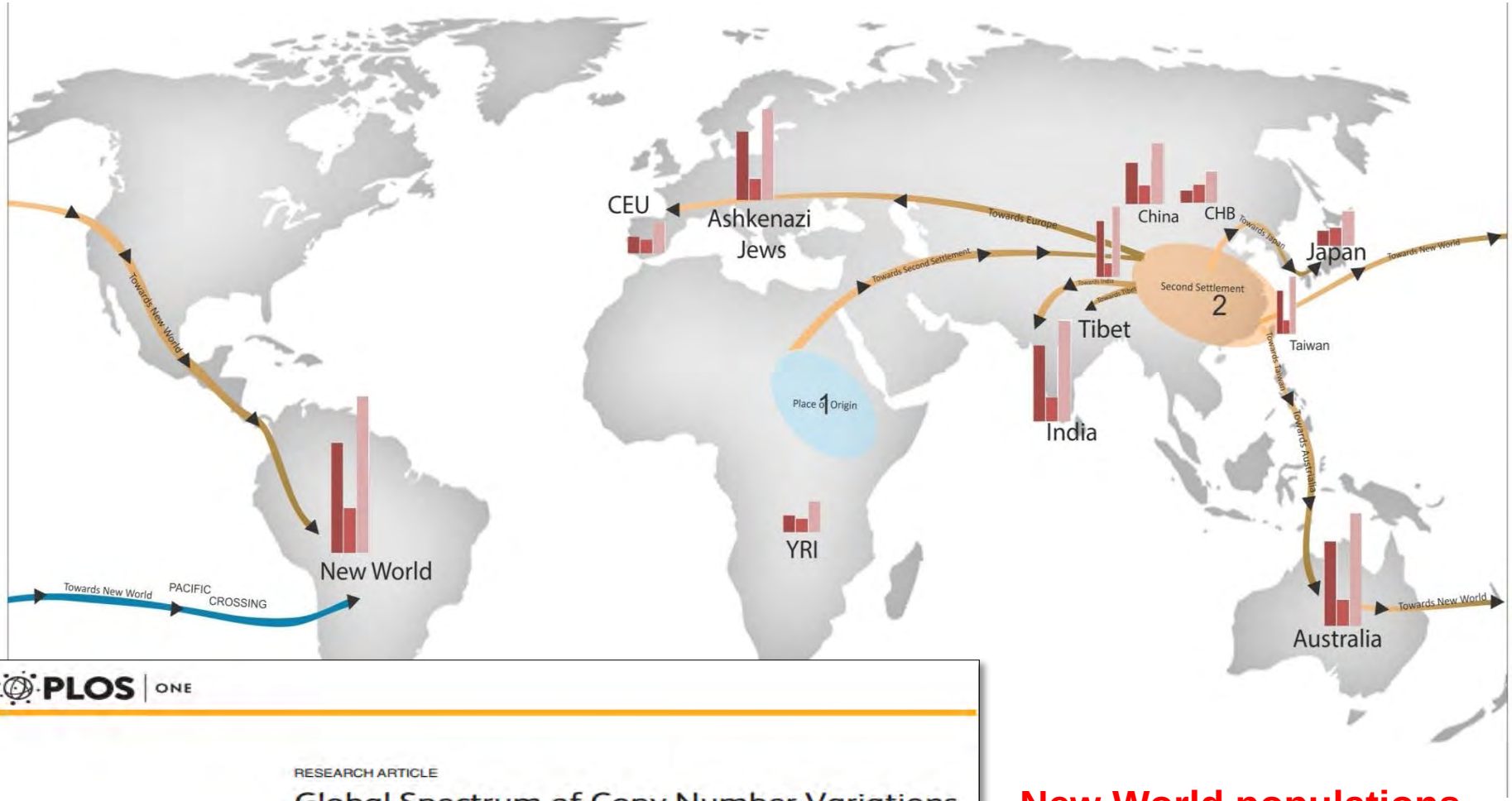


- More in Coding regions
- Duplications are more

CNVs impact on genes participating in cellular processes



Path of Human Migration based on CNVs



PLOS ONE

RESEARCH ARTICLE

Global Spectrum of Copy Number Variations Reveals Genome Organizational Plasticity and Proposes New Migration Routes

Avinash M. Veerappa[1], Sangeetha Vishweswaraiah[1☯], Kusuma Lingaiah[1☯], Megha Murthy[1☯], Raviraj V. Suresh[1☯], Dinesh S. Manjegowda[2], Nallur B. Ramachandra[1]*

1 Genetics and Genomics Lab, Department of Studies in Zoology, University of Mysore, Manasagangotri, Mysore-06, Karnataka, India, 2 NUCSER, KS Hegde Medical Academy, Nitte University, Mangalore-18, Karnataka, India

☯ These authors contributed equally to this work.
* nallurbr@gmail.com

New World populations – the highest number of CNVs

Summary

- **Identified 44109 CNVs from 1715 genomes with size of 7±3Mb**
- **Larger CNVs were few and seen in chromosomes 9, 21, X and Y**
- **CNV distribution is independent of chromosome size**
- **Population specific CNVs were observed**
- CNV map uncovered the unexplored genomic regions
- New World populations - highest number of CNVs which proposes new human migration routes in addition to the existing ones
- **CNVs in ~1,329 genes/ genome**

Our other Publications on CNVs

OPEN ACCESS Freely available online

PLOS ONE

Copy Number Variations Burden on miRNA Genes Reveals Layers of Complexities Involved in the Regulation of Pathways and Phenotypic Expression

Avinash M. Veerappa[1], Megha Murthy N[1☯], Sangeetha Vishweswaraiah[1☯], Kusuma Lingaiah[1☯], Raviraj V. Suresh[1☯], Somanna Ajjamada Nachappa[1], Nelchi Prashali[1], Sangeetha Nuggehalli Yadav[1], Manjula Arsikere Srikanta[1], Dinesh S. Manjegowda[2,4], Keshava B. Seshachalam[3], Nallur B. Ramachandra[1*]

1 Genetics and Genomics Lab, Department of Studies in Zoology, University of Mysore, Manasagangotri, Mysore, Karnataka, India, 2 Department of Anatomy, Yenepoya Medical College, Yenepoya University, Mangalore, Karnataka, India, 3 Department of Neurology, JSS Hospital, Mysore, Karnataka, India, 4 Nitte University Centre for Science Education & Research, K S Hegde Medical Academy, Nitte University, Deralakatte, Mangalore, Karnataka, India

Genet. Res., Camb. (2014), vol. 96, e12. © Cambridge University Press 2014
doi:10.1017/S0016672314000159

Genome-wide copy number scan identifies *IRF6* involvement in Van der Woude syndrome in an Indian family

DINESH S. MANJEGOWDA[1,2*], MANU PRASAD[3], AVINASH M. VEERAPPA[4] AND NALLUR B. RAMACHANDRA[4]

[1]*Genetics Unit, Department of Anatomy, Yenepoya Medical College, Mangalore-575018, Karnataka, India*
[2]*NUCSER, K. S. Hegde Medical Academy, Nitte University, Mangalore-575018, Karnataka, India*
[3]*Centre for Craniofacial Anomaly, Yenepoya Dental College and Hospital, Mangalore-575018, Karnataka, India*
[4]*Genetics and Genomics Lab, Department of Studies in Zoology, University of Mysore, Manasagangotri, Mysore-570006, Karnataka, India*

(*Received 18 February 2014; revised 13 July 2014; accepted 12 September 2014*)

RESEARCH ARTICLE

2013, 9999:1–9

AMERICAN JOURNAL OF medical genetics PART B
Neuropsychiatric Genetics

Genome-Wide Copy Number Scan Identifies Disruption of *PCDH11X* in Developmental Dyslexia

Avinash M. Veerappa,[1] Marita Saldanha,[2] Prakash Padakannaya,[2] and Nallur B. Ramachandra[1*]

[1]Genomics Laboratory, Department of Studies in Zoology, University of Mysore, Manasagangotri, Mysore, Karnataka, India
[2]Department of Studies in Psychology, University of Mysore, Manasagangotri, Mysore, Karnataka, India

Manuscript Received: 30 April 2013; Manuscript Accepted: 13 August 2013

PLOS ONE | Published: July 3, 2013

Unravelling the complexity of human olfactory receptor repertoire by copy number analysis across population using high resolution arrays.

Veerappa AM[1], Vishweswaraiah S, Lingaiah K, Murthy M, Manjegowda DS, Nayaka R, Ramachandra NB.

RESEARCH ARTICLE

AMERICAN JOURNAL OF medical genetics PART B
Neuropsychiatric Genetics

2014, 165(7): 572-580

Family Based Genome-Wide Copy Number Scan Identifies Complex Rearrangements at 17q21.31 in Dyslexics

Avinash M. Veerappa,[1] Marita Saldanha,[2] Prakash Padakannaya,[2] and Nallur B. Ramachandra[1*]

[1]Genetics and Genomics Laboratory, Department of Studies in Zoology, University of Mysore, Manasagangotri, Mysore
[2]Department of Studies in Psychology, University of Mysore, Manasagangotri, Mysore

Manuscript Received: 31 January 2014; Manuscript Accepted: 26 June 2014

Genet. Res., Camb. (2014), vol. 96, e17. © Cambridge University Press 2014
doi:10.1017/S0016672314000202

Impact of copy number variations burden on coding genome in humans using integrated high resolution arrays

AVINASH M. VEERAPPA[1], KUSUMA LINGAIAH[1†], SANGEETHA VISHWESWARAIAH[1†], MEGHA N. MURTHY[1†], RAVIRAJ V. SURESH[1†], DINESH S. MANJEGOWDA[2] AND NALLUR B. RAMACHANDRA[1*]

[1]*Genetics and Genomics Lab, Department of Studies in Zoology, University of Mysore, Manasagangotri, Mysore-06, Karnataka, India*
[2]*NUCSER, K. S. Hegde Medical Academy, Nitte University, Deralakatte, Mangalore-575 018, Karnataka, India*

(*Received 29 September 2014; revised 12 November 2014; accepted 13 November 2014*)

Genet. Res., Camb. (2015), vol. 97, e18. © Cambridge University Press 2015
doi:10.1017/S0016672315000191

Global patterns of large copy number variations in the human genome reveal complexity in chromosome organization.

Veerappa AM[1], Suresh RV[1], Vishweswaraiah S[1], Lingaiah K[1], Murthy M[1], Manjegowda DS[2], Padakannaya P[3], Ramachandra NB[1]

Hindawi Publishing Corporation
Journal of Nucleic Acids
Volume 2016, Article ID 1648527, 7 pages
http://dx.doi.org/10.1155/2016/1648527

Hindawi

Research Article

Copy Number Variation of UGT 2B Genes in Indian Families Using Whole Genome Scans

Avinash M. Veerappa,[1] Prakash Padakannaya,[2] and Nallur B. Ramachandra[1]

[1]*Genetics and Genomics Lab, Department of Studies in Genetics and Genomics, University of Mysore, Manasagangotri, Mysore 570 006, India*
[2]*Department of Studies in Psychology, University of Mysore, Manasagangotri, Mysore 570 006, India*

3) CNVs in disease risk genes in normal cohorts

(Type 2 Diabetes, Parkinson Disease and Asthma)

a) Type 2 Diabetes: A Raging Global Epidemic

India is a CAPITAL of diabetes

Consequences

- Atherosclerosis
- Retinopathy
- Neuropathy
- Foot problems
- Nephropathy

Causes

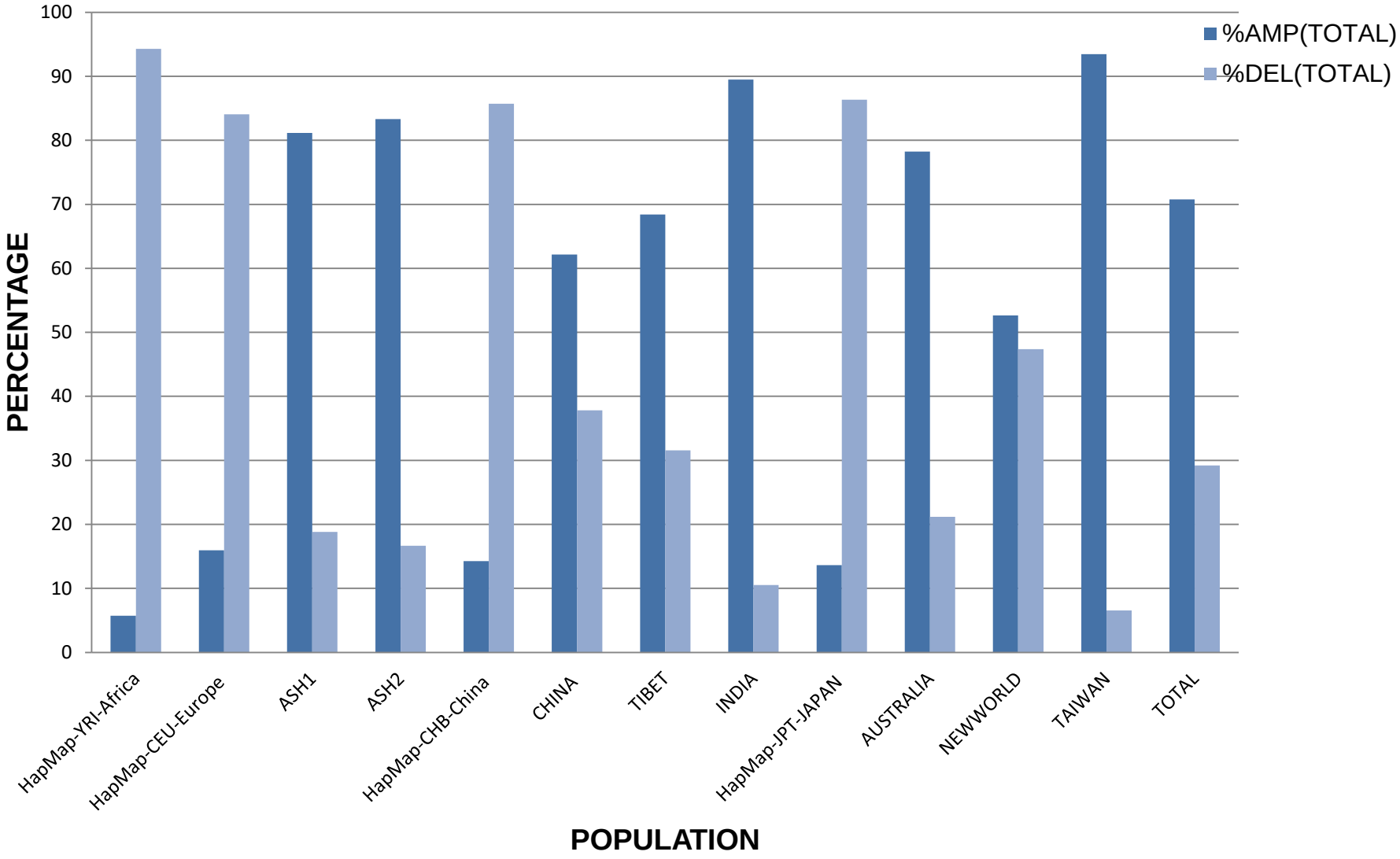
Defect in the genes involved in:

- pancreatic β cell function,
- insulin action/glucose metabolism /
- other metabolic conditions

➢ **Targeted 83 genes for CNVs study**

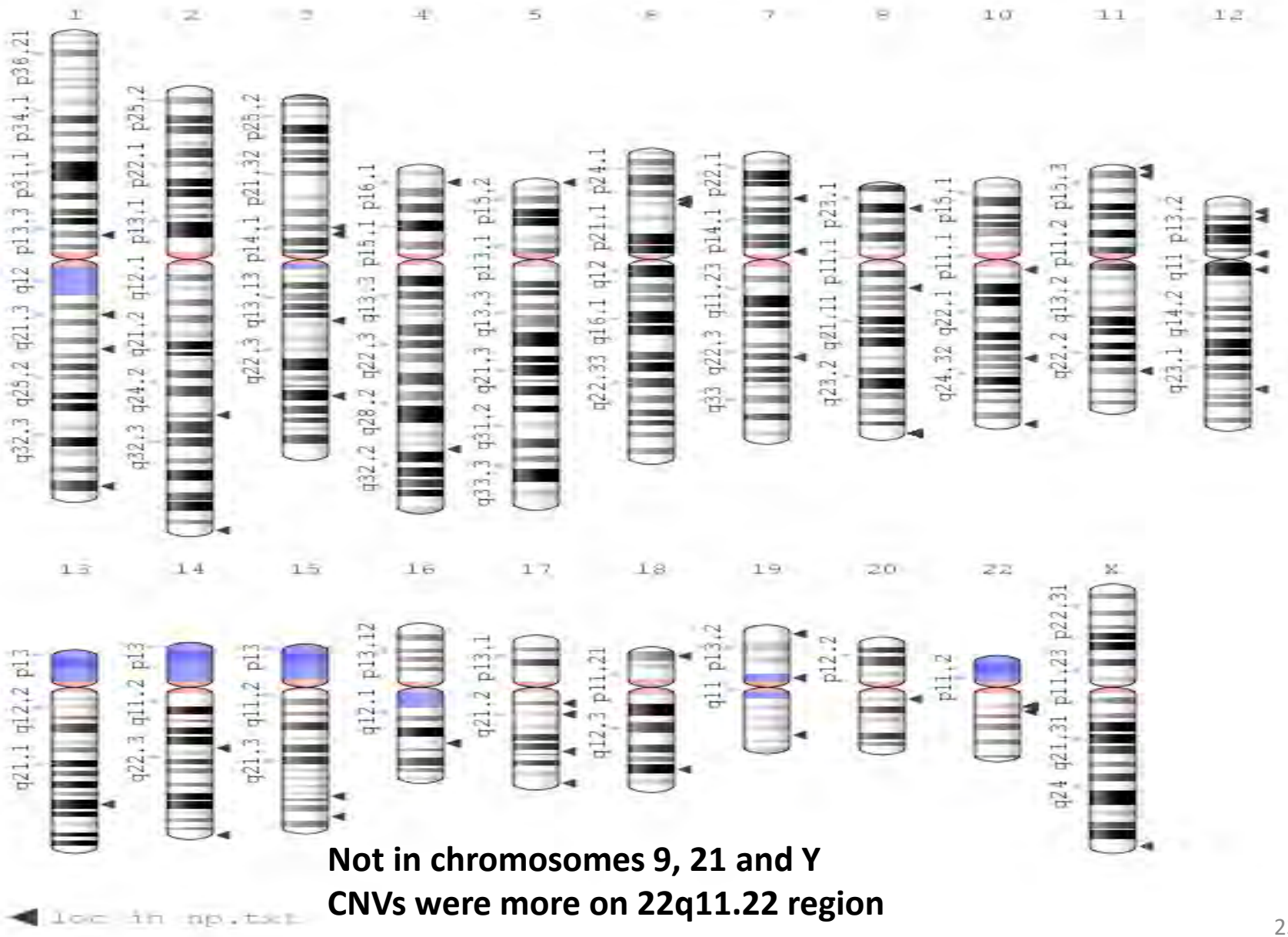
WHO, 2017

Duplication and Deletion T2DM-CNVs in 12 study populations



Prabhanjan et al., 2016, 113, Diabetes Research and Clinical Practice, 160-170

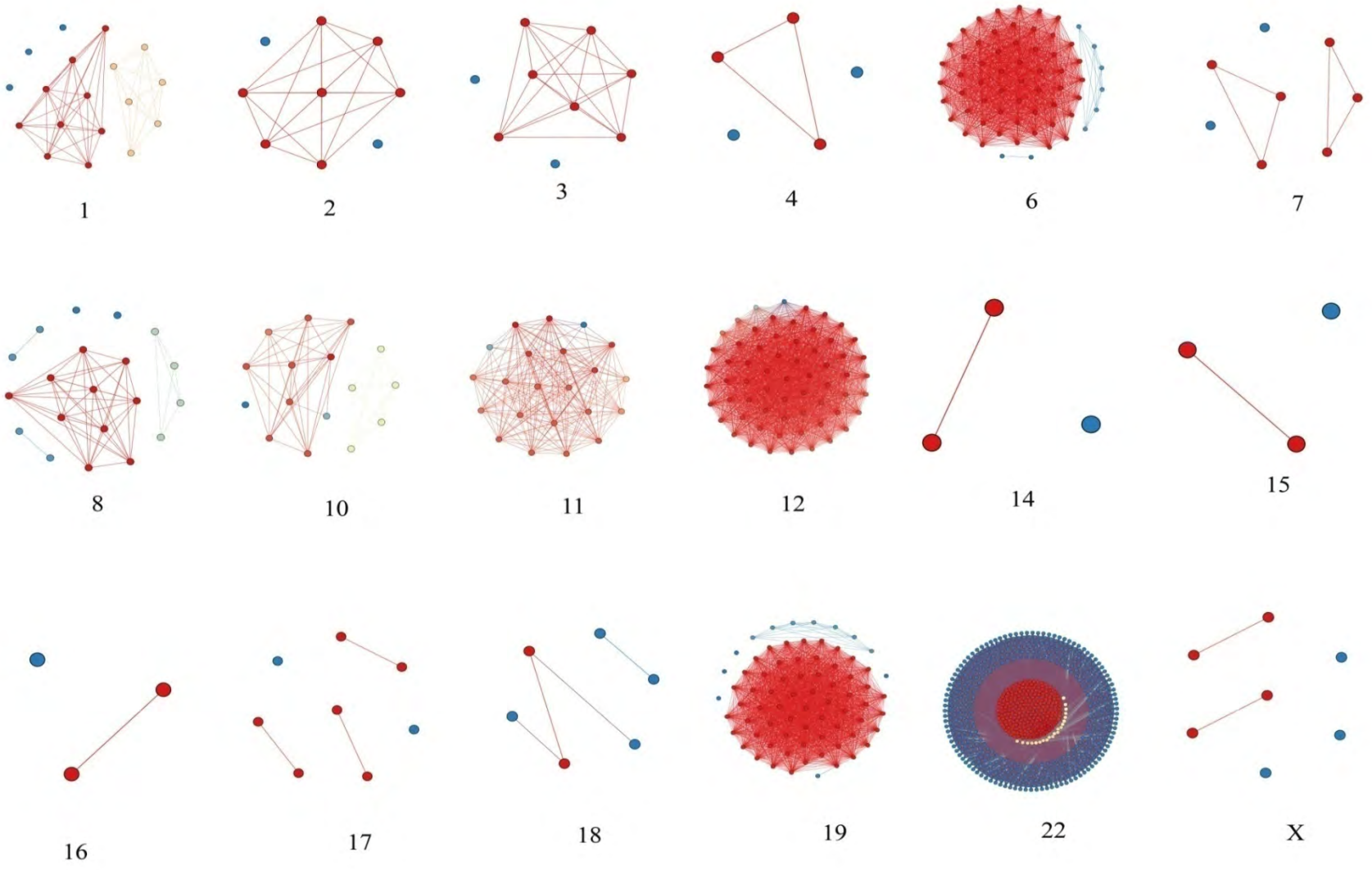
Chromosomes maps of T2DM-CNVs



Not in chromosomes 9, 21 and Y

CNVs were more on 22q11.22 region

Hot spot detection of T2DM-CNVs in 12 populations

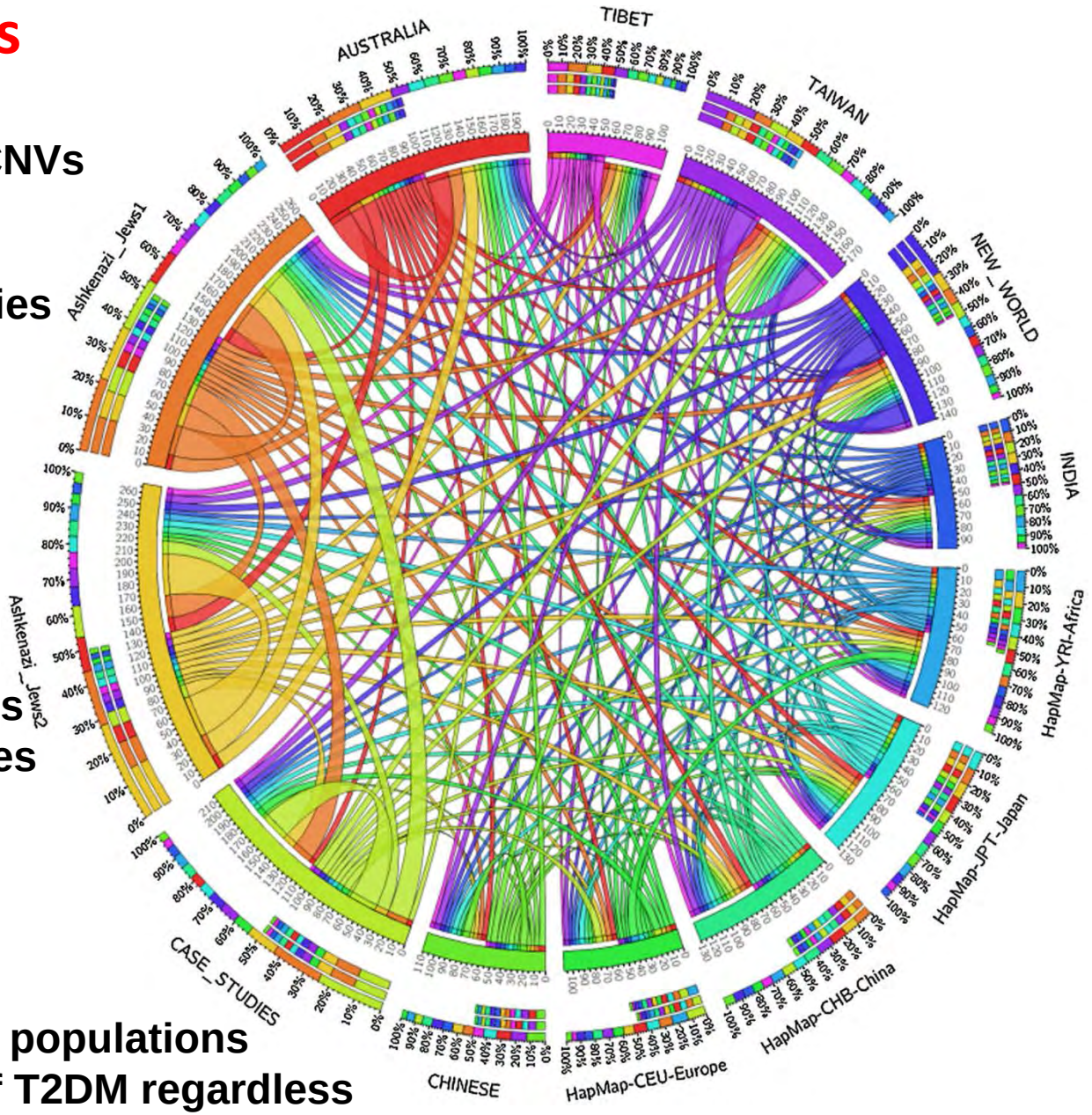


Shared CNVs -Circos

Outer ring = % of Shared CNVs

Innermost ring =
12 populations + case studies

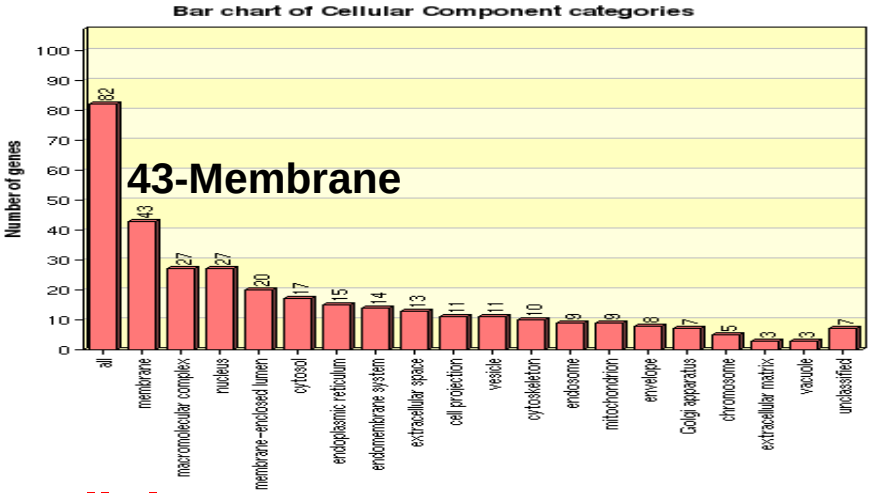
Shared map of 83 genes under CNVs



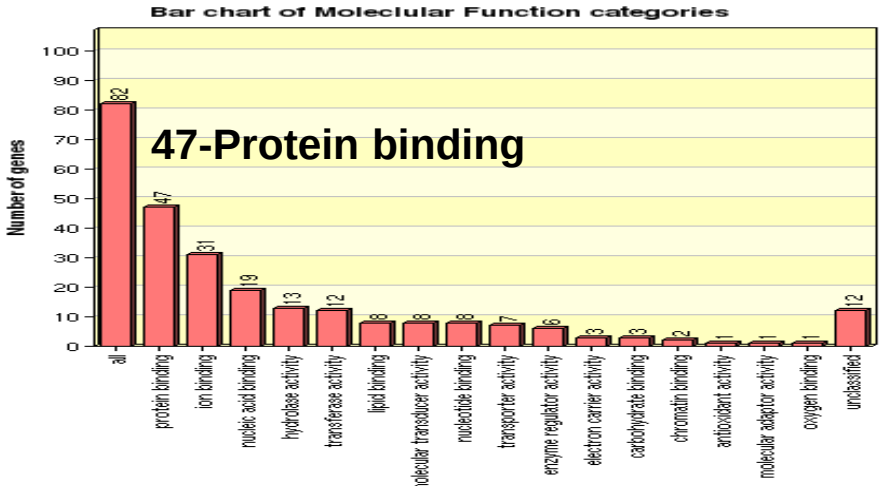
- **24 T2DM-CNV gene regions shared between case studies & Ashkenazi Jews 1**

- **Common occurrence in all populations**
- **Affect in predisposition of T2DM regardless to the ethnicity**

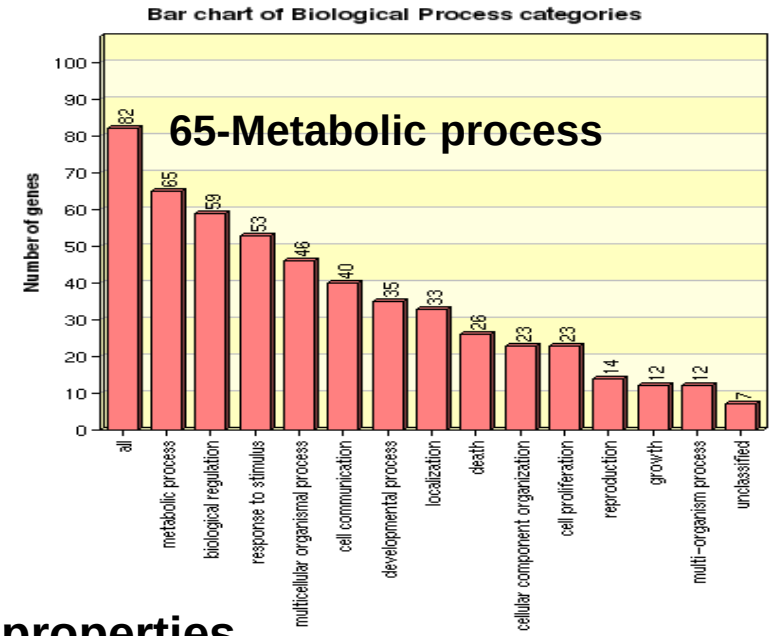
Enrichment of T2DM–CNV genes



Cellular Process



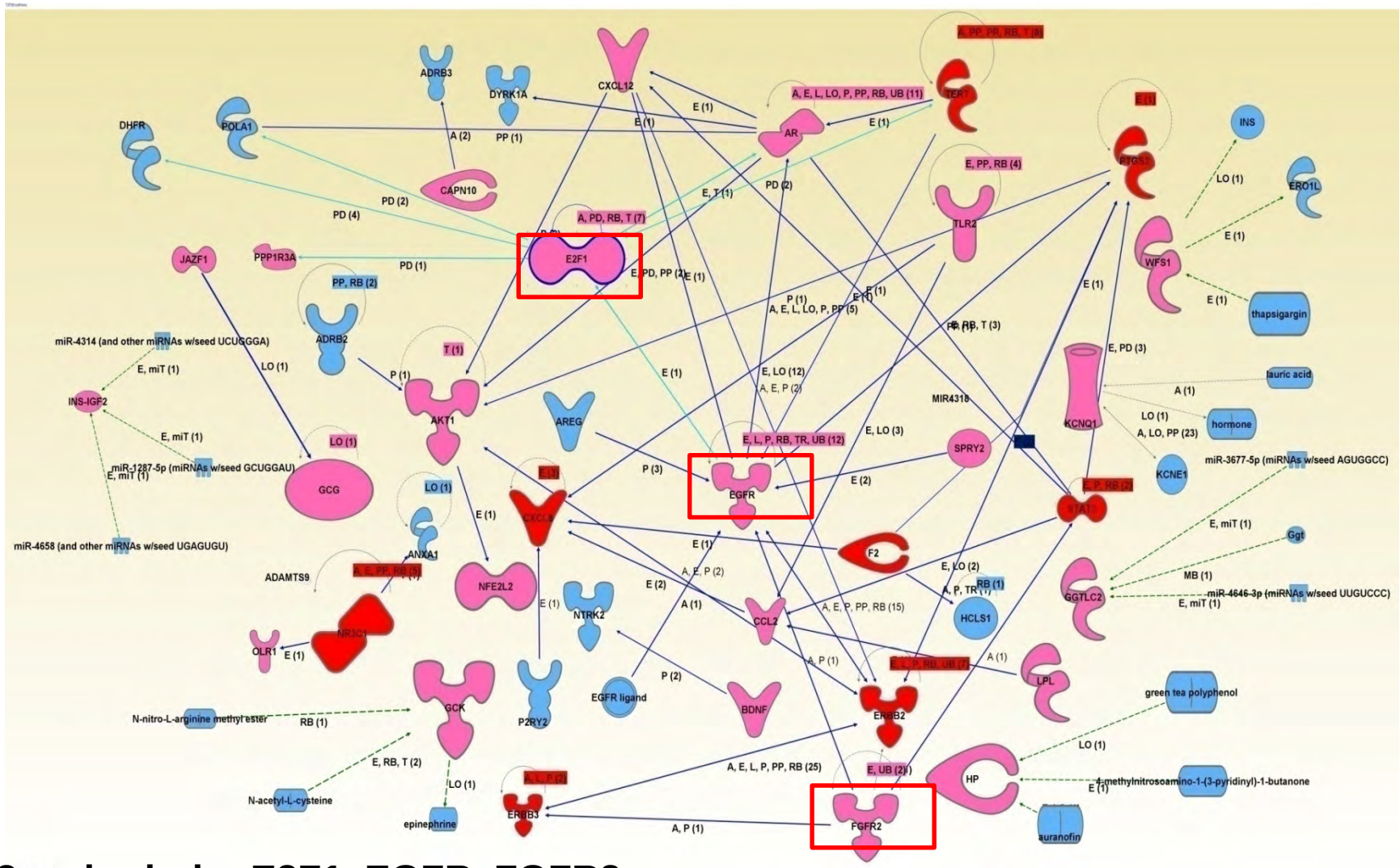
Molecular function



Biological process

Affect the functional properties of the above processes

Network of T2DM-CNV genes obtained through IPA



3 major hubs-E2F1, EGFR, FGFR2 interaction leads to change in the expression level - severity of the disease

Pink – T2DM genes affected by CNVs
Red – Primary interactors
Blue – T2DM genes NOT under CNV burden

DIABETES RESEARCH AND CLINICAL PRACTICE 113 (2016) 160–170

Contents available at ScienceDirect

Diabetes Research and Clinical Practice

journal homepage: www.elsevier.com/locate/diabres

ELSEVIER

International Diabetes Federation

Type 2 diabetes mellitus disease risk genes identified by genome wide copy number variation scan in normal populations

Manasa Prabhanjan [a,1,2], *Raviraj V. Suresh* [b,1,3], *Megha N. Murthy* [b,1,4], *Nallur B. Ramachandra* [b,*]

[a] *Department of Studies in Genetics and Genomics, University of Mysore, Manasagangotri, Mysore 570006, Karnataka, India*

[b] *Genetics and Genomics Lab, Department of Studies in Genetics and Genomics, University of Mysore, Manasagangotri, Mysore 570006, Karnataka, India*

- **83 disease causal and associated genes identified in 34.4% of the individuals under study**
- **24 T2DM-CNV gene regions shared between case studies & Ashkenazi Jews 1**
- **Hotspots identified on chromosomes 22, 12, 6, 19 and 11**

b) Parkinson Disease

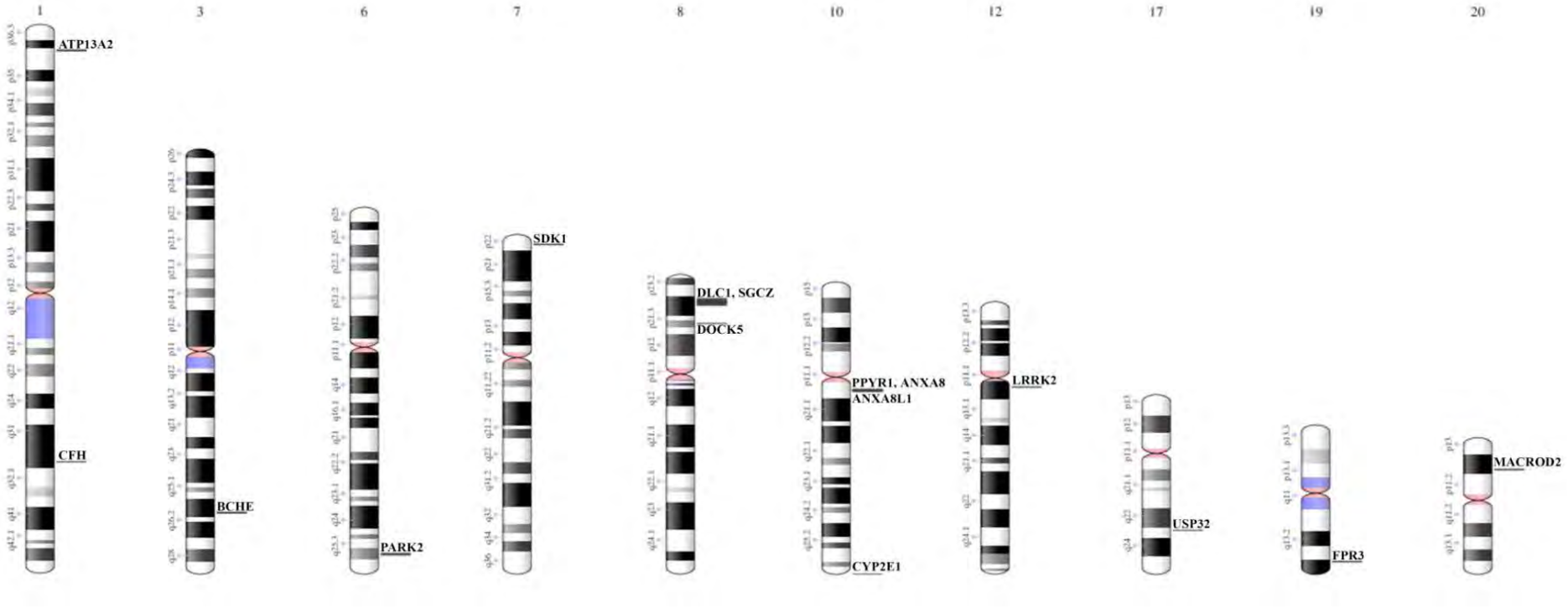
- Progressive neurodegenerative movement disorder
- Characterized by loss of dopaminergic neurons in the *substantia nigra (mid brain)*
- Tetrad of clinical motor symptoms (TRAP)
  - **T**remor at rest
  - **R**igidity
  - **A**kinesia (slowness of movements)
  - **P**ostural instability
- Other non motor symptoms
  - Sleep disorders
  - Cognitive disabilities – dementia
  - Depression

- **First described by James Parkinson in 1817**
- **>60 years - Prevalence of 1%**

➢ **Targeted 220 genes for CNVs study**

Perlmutter, 2009; Lee and Gilbert, 2016

Karyogram - location of 16 CNV- PD genes

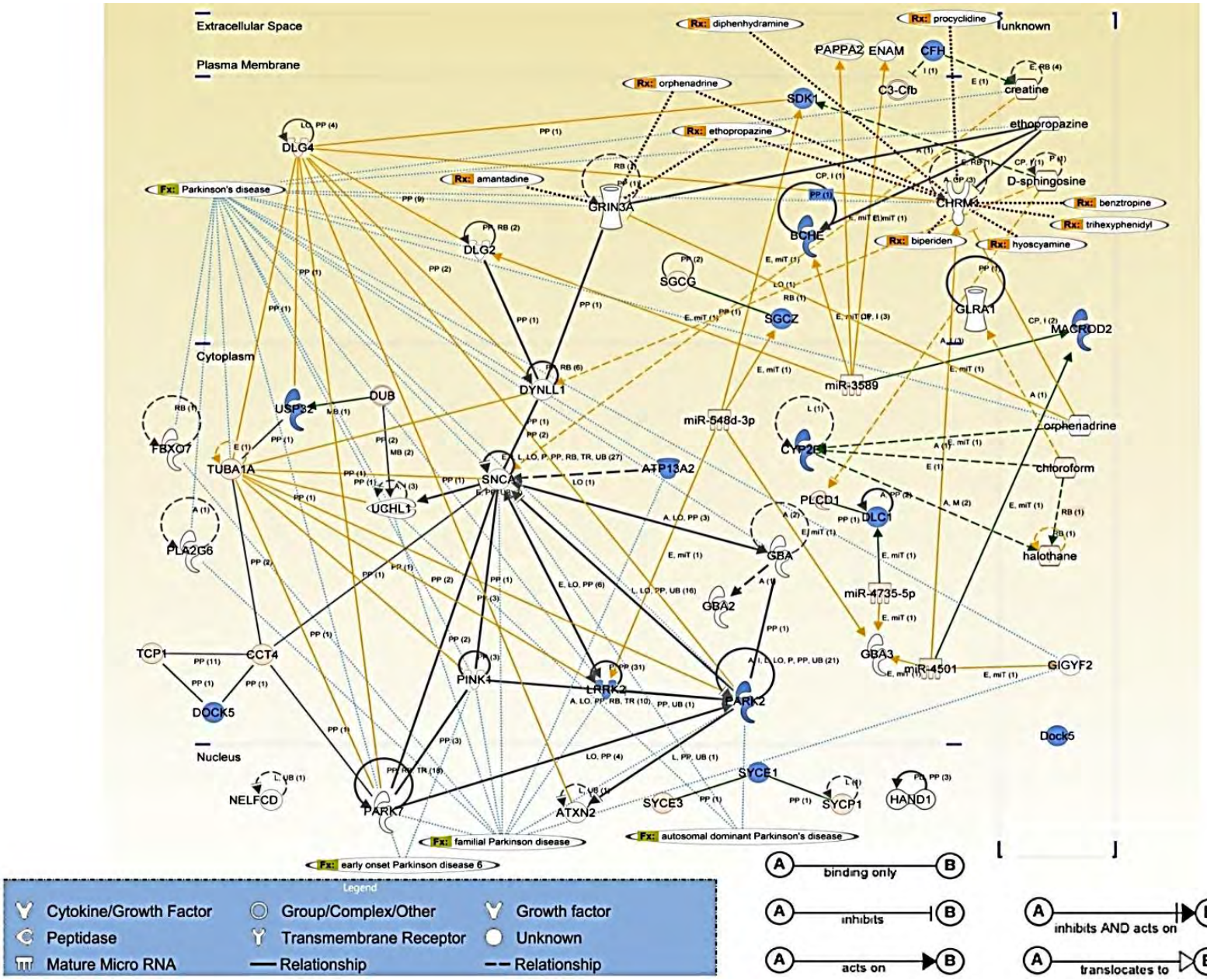


Located on 10 chromosome regions

Genes identified under CNV burden

Gene	Chr Location	Size (Mb)	Markers	Breakpoint	Start-End Marker	CN State	Gene Function	Gain /Loss	Intact/ Partial
LRRK2	12q12	138	108	40552750-40690922	CN_604968-SNP_A-8603327	1	Role in the phosphorylation of proteins central to Parkinson disease.	Loss	Partial
ATP13A2	1p36.13	484	182	16830808-17314702	CN_444434-SNP_A-2090089	3	May play a role in intracellular cation homeostasis and the maintenance of neuronal integrity	Gain	Intact
USP32	17q23.2	102	47	58400794-58502312	CN_748298-SNP_A-4286683	3	Unknown	Gain	Partial
DOCK5	8p21.2	110	83	24978480-25088714	CN_1274461-CN_1276612	3	Guanine nucleotide exchange factor (GEF) for Rho and Rac.	Gain	Partial
PARK2	6q26	475	393	162569201-163043760	SNP_A-2069197-CN_1168439	3	Functions within E3 ubiquitin ligase complex	Gain	Partial
MACROD2	20p12.1	329	328	14756011-15085365	CN_887478-SNP_A-2258435	1	Deacetylates O-acetyl-ADP ribose	Loss	Intact
BCHE	3q26.1	349	199	165518759-165867791	CN_1009363-CN_1011514	3	Inactivates the neurotransmitter acetylcholine	Gain	Partial
SDK1	7p22.2	239	218	4273289-4512294	SNP_A-8545829-CN_1227020	3	Guides axonal terminals to specific synapses in developing neurons	Gain	Intact
DLC1	8p22	1303	1333	13355424-14658080	SNP_A-8315185-CN_1287618	4	Critical role in biological processes such as cell migration and proliferation	Gain	Intact

Protein Interaction Network analysis



Blue - we identified

LRRK2
PARK2
ATP13A2
DOCK5

Orange- Primary interactors

White- Secondary interactors

Taylor & Francis

Neurological Research

A Journal of Progress in Neurosurgery, Neurology and Neuro Sciences

Volume 38, 2016-Issue 9

ISSN: 0161-6412 (Print) 1743-1328 (Online) Journal homepage: http://www.tandfonline.com/loi/yner20

High-resolution arrays reveal burden of copy number variations on Parkinson disease genes associated with increased disease risk in random cohorts

Megha N. Murthy, Avinash M. Veerappa, Keshava B. Seshachalam & Nallur B. Ramachandra

- **16 CNV-PD genes, 3 known to be causal and 13 associated, were found in 18.9% of the individuals under study.**
- **PARK2, was under heavy burden with ~1% of the population containing CNV in the exonic region.**
- **Novel genes in PD pathway + overlaps study in PD cases + normal cohorts CNVs + exome data + transcriptome data - biomarker for PD**

c) Asthma

- Asthma is a chronic disease characterized by recurrent attacks of breathlessness and wheezing, which vary in severity and frequency from person to person

- Asthma is caused by interplay of genes and environment

- **The global prevalence of asthma is increasing, affects 339 million people**

- **~ 1000 people dying each day from asthma, and is in the top 20 causes of years of life lived with disability'**

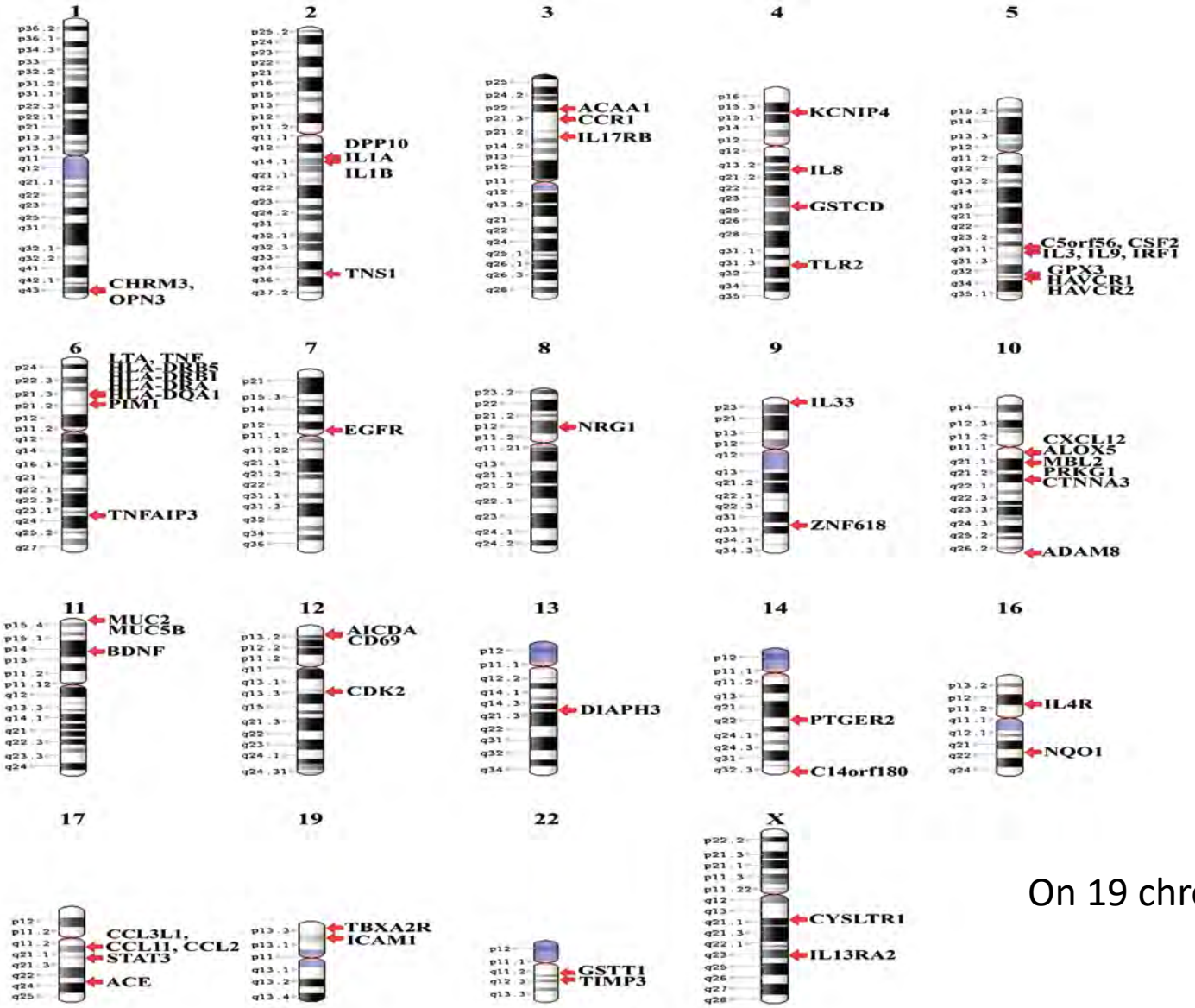
- **Targeted 300 genes for CNVs study**

http://globalasthmareport.org/Global%20Asthma%20Report%202018.pdf

CNVs in Asthma susceptible genes

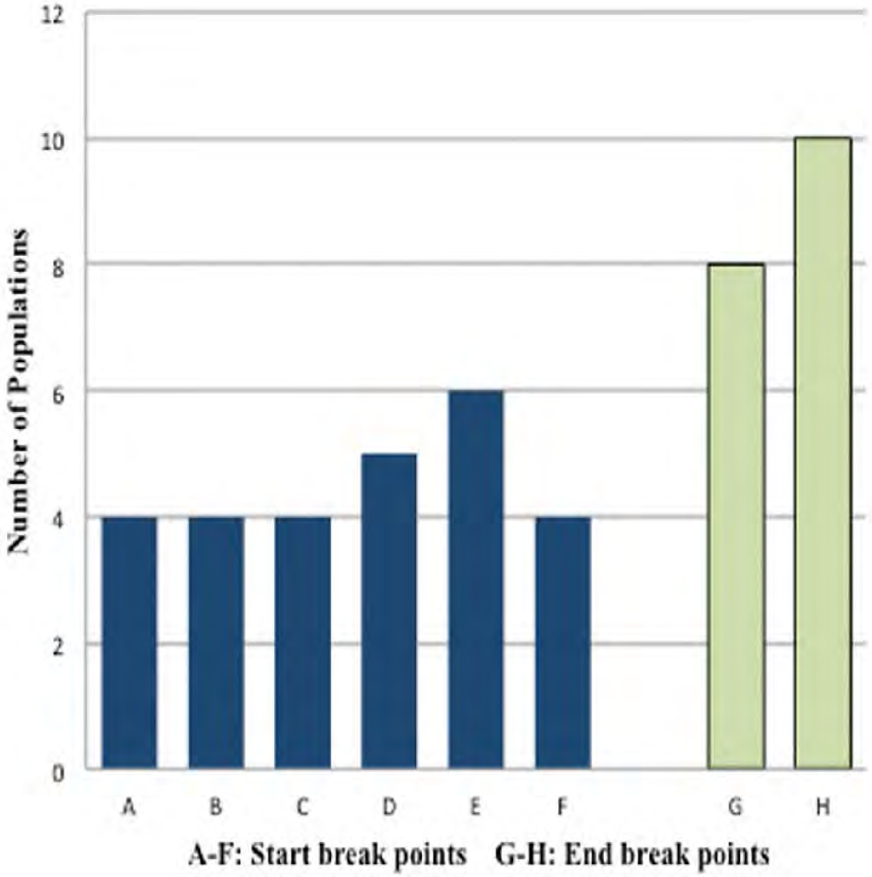
Population	Individuals	Asthma CNVs (%)
HapMap-YRI-Africa	90	30
HapMap-CEU-Europe	90	**3**
Ashkenazi Jews I	464	20
Ashkenazi Jews II	480	28
HapMap-CHB-China	44	16
China	155	15
Tibet	31	65
India	38	**8**
HapMap-JPT-Japan	45	22
Australia	53	42
New World	41	**68**
Taiwan	184	16

Karyogram - location of 61 asthma-CNV associated genes



On 19 chromosomes

Conserved asthma-CNV genes break points



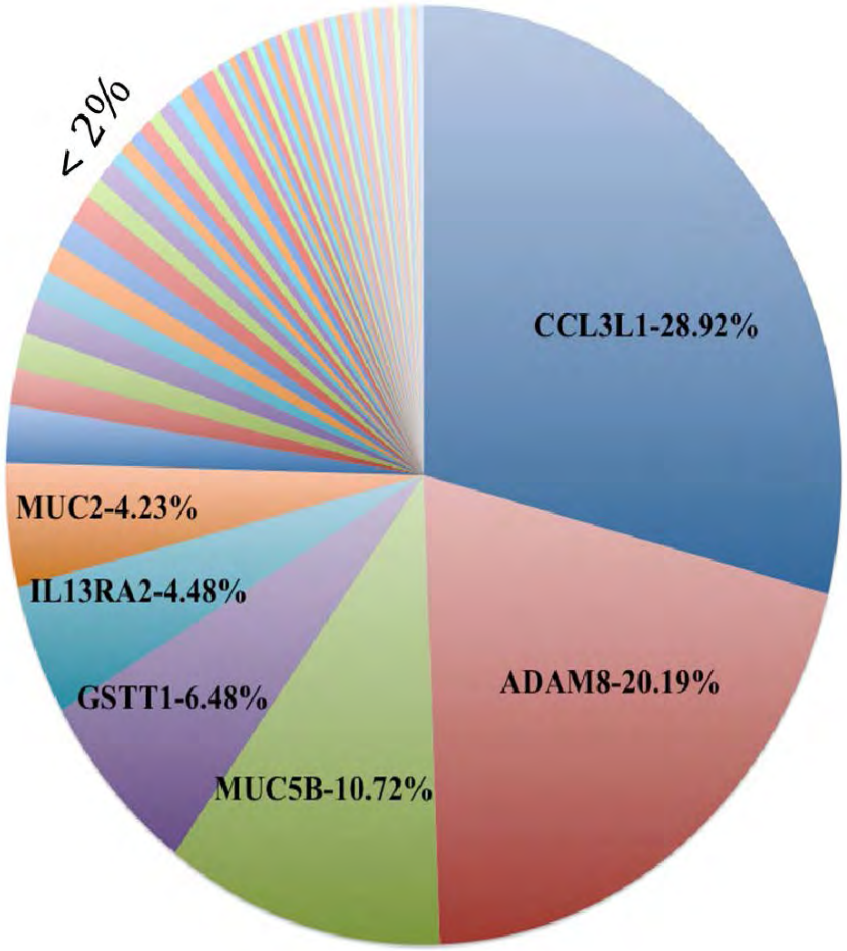
A- 34435512, B- 34439966, C- 34443800, D- 34437116 and E- 34528113 encompasses *CCL3L1*

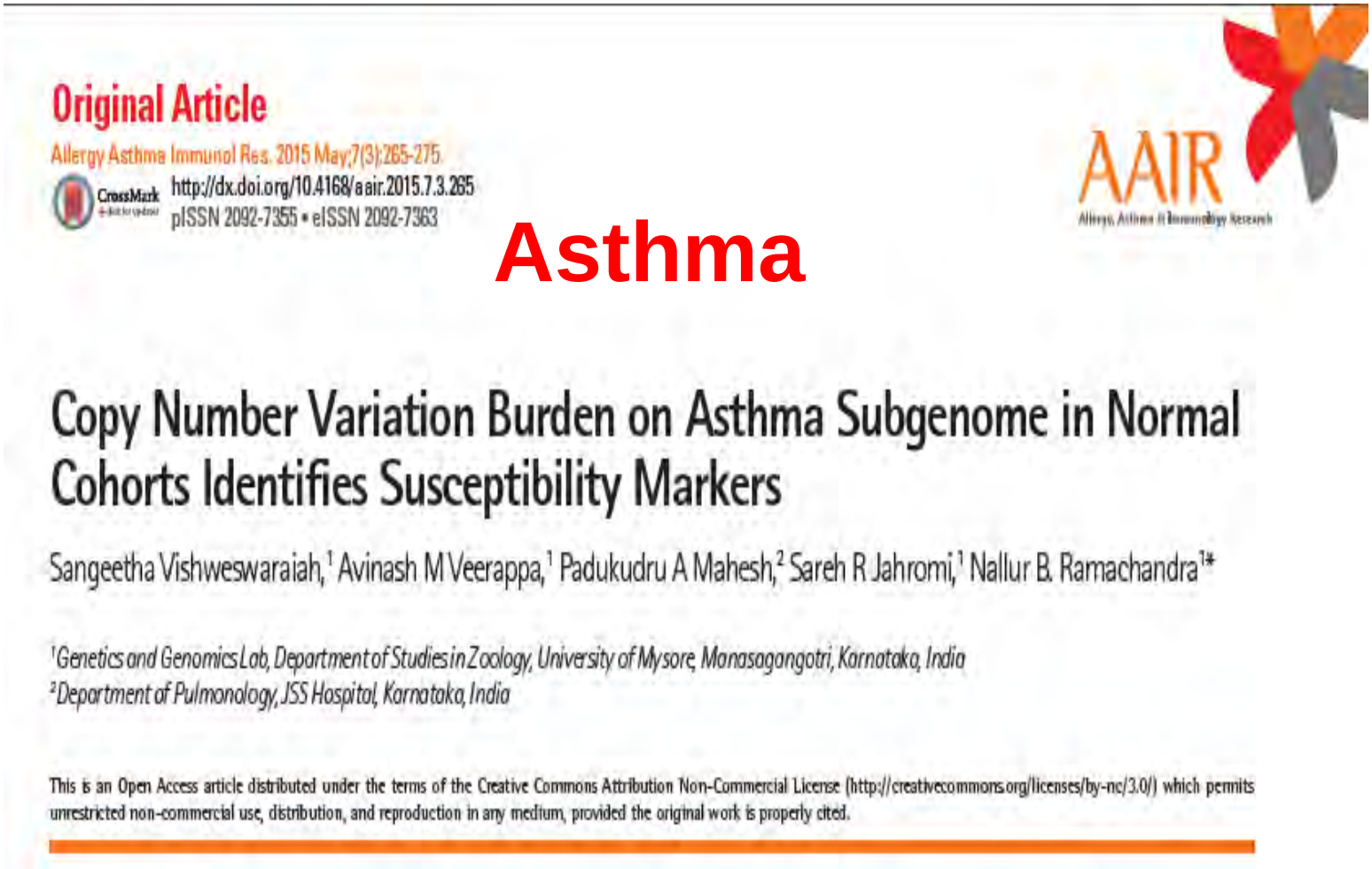
F- 24283004 encompasses *GSTT1*

G- 24396802 encompasses *GSTT1*

H- 34629684 encompasses *CCL3L1*

Frequency of asthma genes under CNV burden





Original Article
Allergy Asthma Immunol Res. 2015 May;7(3):265-275
http://dx.doi.org/10.4168/aair.2015.7.3.265
pISSN 2092-7355 • eISSN 2092-7363

AAIR

Asthma

Copy Number Variation Burden on Asthma Subgenome in Normal Cohorts Identifies Susceptibility Markers

Sangeetha Vishweswaraiah,[1] Avinash M Veerappa,[1] Padukudru A Mahesh,[2] Sareh R Jahromi,[1] Nallur B. Ramachandra[1]*

[1]Genetics and Genomics Lab, Department of Studies in Zoology, University of Mysore, Manasagangotri, Karnataka, India
[2]Department of Pulmonology, JSS Hospital, Karnataka, India

- **61 asthma –CNV associated genes observed in 17% of the populations under study**
- **CN state effects the protein level that may lead to the susceptibility to asthma**

4) CNVs involvement in Evolution -

Y-chromosome XTR as pseudoautosomal region 3

Evolution of Y chromosome

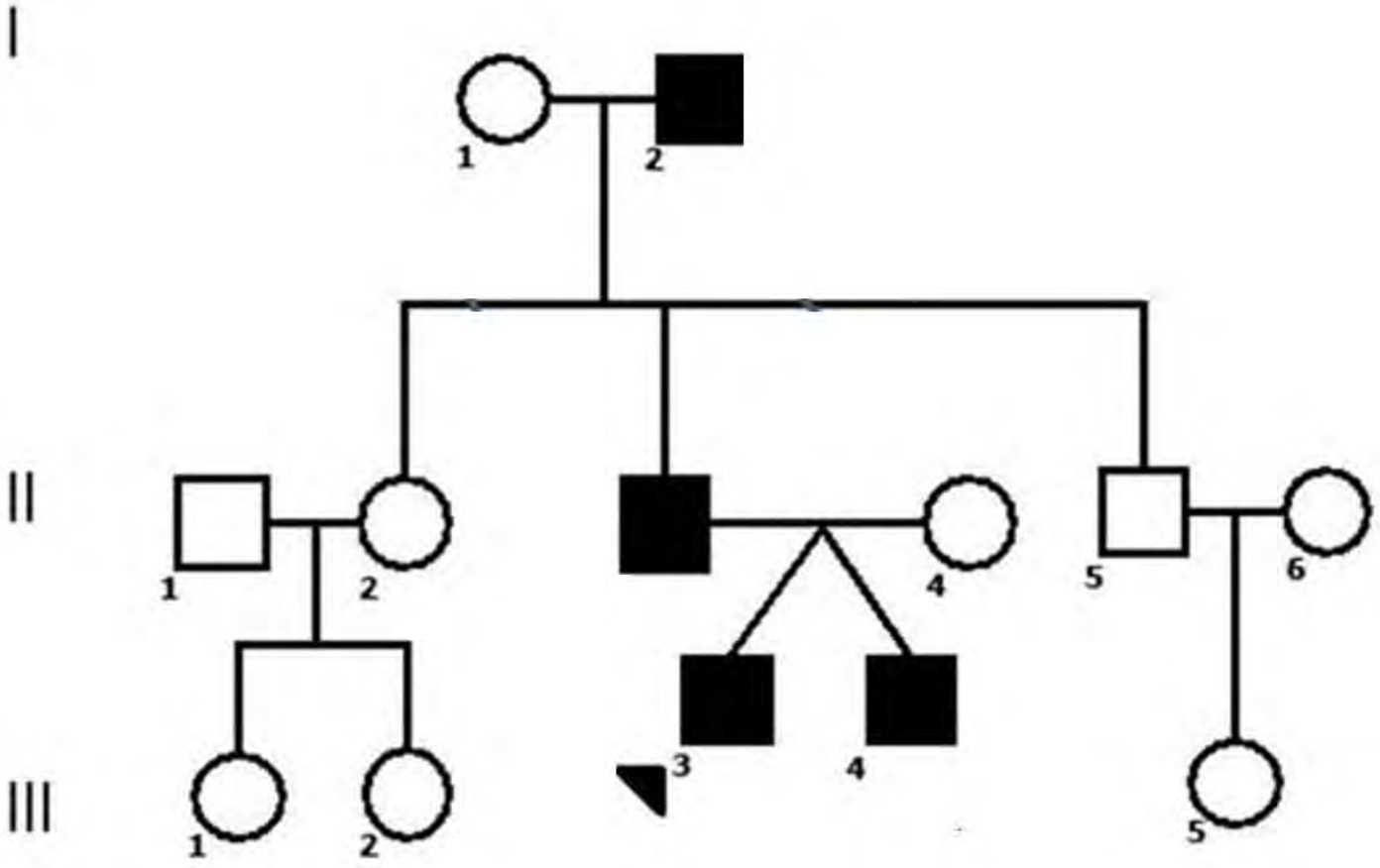
- **A 3.5-Mb region of the X chromosome underwent duplication and transposition to the Y chromosome ~5–6 Mya.**

- **This X-transposed-region (XTR) originated at Xq21.3 and was inserted at Yp11.2.**

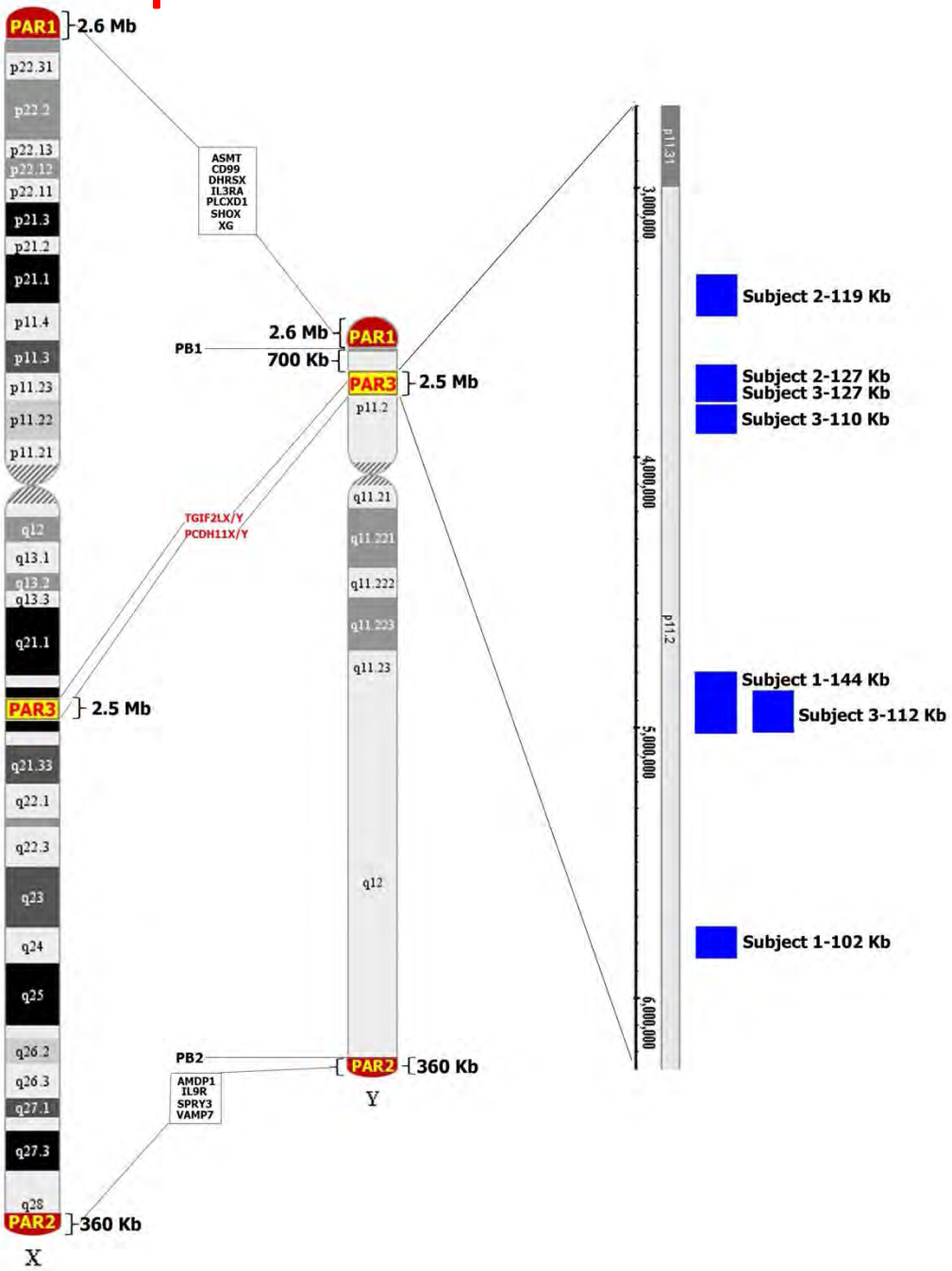
(Mangs and Morris 2007)

- **We identified >102 kb of the Yp11.2 region transposed to Xq21.3 through allelic unequal recombination in three females from three different families.**

A Dyslexic family



Yp11.2 blocks recombined with Xq21.3 region



Displays the proposed PAR3 blocks on both the X and Y chromosomes indicating the size, genes, location and the pseudoautosomal boundaries together with the already identified PAR1 and PAR2 features.

Shared features of the Pseudoautosomal regions 1 and 2 with PAR3 (XTR)

Shared Features	PAR1	PAR2	PAR3 (XTR)
Sequence Homology	> 98 %	> 98 %	> 98 %
Size	2.6 Mb	320 kb	~ 2.5 Mb
Has allelic homologues on both X and Y	Yes	Yes	Yes
Formed due to duplication	No	Yes	Yes
Genes escape inactivation	Yes	Yes	Yes
Recombination	Yes	Yes	Yes
Recombination Frequency	Obligatory per meiosis	1 in 40 times	1 in 40 times
Genes	24	5	3

Funct Integr Genomics (2013) 13:285–293
DOI 10.1007/s10142-013-0323-6

ORIGINAL PAPER

Copy number variation-based polymorphism in a new pseudoautosomal region 3 (PAR3) of a human X-chromosome-transposed region (XTR) in the Y chromosome

Avinash M. Veerappa · Prakash Padakannaya · Nallur B. Ramachandra

- **Observed in 500 females 2% of recombination in PAR3 region between X and Y chromosomes**

5) Conclusions

- **CNVs - the most powerful tools for genomes analysis**
- **Conserved - > 90% sequence identity - marker**
- **Differences in genome size contribute to our uniqueness**
- **Role in human disease causing**
- **Size and location of variations- impact on phenotype**
- **Helps in better understanding of human genome evolution**








Acknowledgements

Funding – DST and Univ. of Mysore

THANK YOU