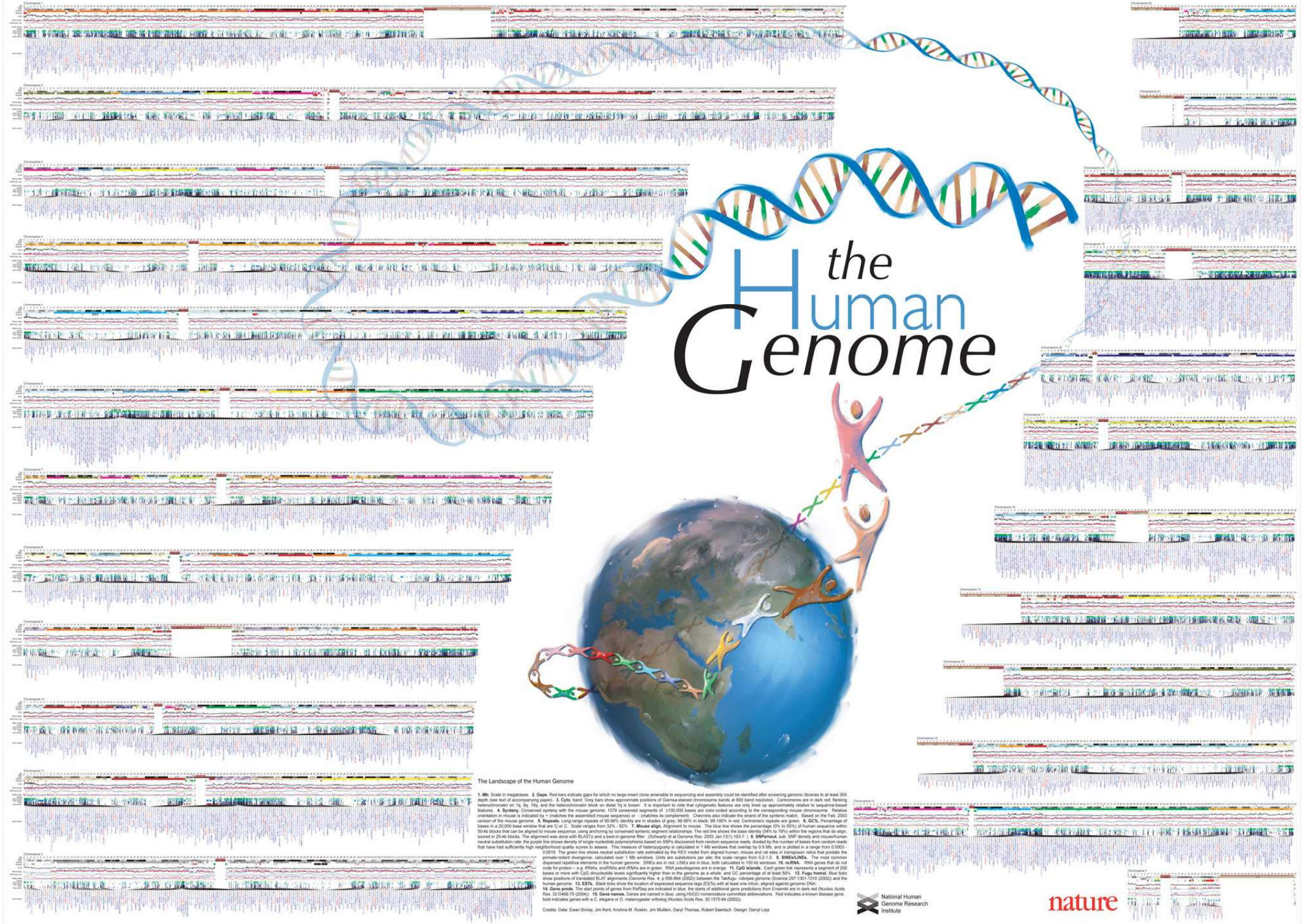
Biobanking

Meenakshi Sharma

Indian Council of Medical Research

New Delhi

Indo-Dutch Workshop
Trivandrum
Jan 21-23



The Landscape of the Human Genome

1. Mb. Scale in megabases. **2. Gaps.** Red bars indicate gaps for which no large-insert clone amenable to sequencing and assembly could be identified after screening genomic libraries to at least 30X depth (see text of accompanying paper). **3. Cyto.** band. Grey bars show approximate positions of Giemsa-stained chromosome bands at 800 band resolution. Centromeres are in dark red; flanking heterochromatin on 1q, 9q, 16q, and the heterochromatin block on distal Yq is brown. It is important to note that cytogenetic features are only lined up approximately relative to sequence-based features. **4. Synteny.** Conserved synteny with the mouse genome: 1079 conserved segments of ≥100,000 bases are color-coded according to the corresponding mouse chromosome. Relative orientation in mouse is indicated by + (matches the assembled mouse sequence) or - (matches its complement). Chemora also indicate the strand of the syntenic match. Based on the Feb. 2003 version of the mouse genome. **5. Repeats.** Long-range repeats of 90-98% identity are in shades of grey; 98-99% in black; 99-100% in red. Centromeric repeats are green. **6. GC%.** Percentage of bases in a 20,000 base window that are G or C. Scale ranges from 32% - 62%. **7. Mouse align.** Alignment to mouse: The blue line shows the percentage (0% to 85%) of human sequence within 50-kb blocks that can be aligned to mouse sequence, using anchoring by conserved syntenic segment relationships. The red line shows the base identity (94% to 79%) within the regions that do align, scored in 25-kb blocks. The alignment was done with BLASTz and a best-in-genome filter. (Schwartz et al Genome Res. 2003 Jan;13(1):103-7). **8. SNPs/neut.** sub. SNP density and mouse/human neutral substitution rate: the purple line shows density of single nucleotide polymorphisms based on SNPs discovered from random sequence reads, divided by the number of bases from random reads that have had sufficiently high neighborhood quality scores to assess. This measure of heterozygosity is calculated in 1 Mb windows that overlap by 0.9 Mb, and is plotted in a range from 0.0003 - 0.0018. The green line shows neutral substitution rate estimated by the REV model from aligned human, mouse and rat sites in transposon relics that predate the primate-rodent divergence, calculated over 1 Mb windows. Units are substitutions per site; the scale ranges from 0.2-1.0. **9. SINEs/LINEs.** The most common dispersed repetitive elements in the human genome. SINEs are in red, LINEs are in blue, both calculated in 100 kb windows. **10. ncRNA.** RNA genes that do not code for protein – e.g. tRNAs, snoRNAs and rRNAs are in green. RNA pseudogenes are in orange. **11. CpG islands.** Each green tick represents a segment of 200 bases or more with CpG dinucleotide levels significantly higher than in the genome as a whole, and GC percentage of at least 50%. **12. Fugu homol.** Blue ticks show positions of translated BLAT alignments (Genome Res. 4, p 656-664 (2002)) between the Takifugu rubripes genome (Science 297:1301-1310 (2002)) and the human genome. **13. ESTs.** Black ticks show the location of expressed sequence tags (ESTs) with at least one intron, aligned against genomic DNA. **14. Gene preds.** The start points of genes from RefSeq are indicated in blue, the starts of additional gene predictions from Ensembl are in dark red (Nucleic Acids Res. 32:D468-70 (2004)). **15. Gene names.** Genes are named in blue, using HUGO nomenclature committee abbreviations. Red indicates a known disease gene. bold indicates genes with a C. elegans or D. melanogaster ortholog (Nucleic Acids Res. 30:1575-84 (2002)).

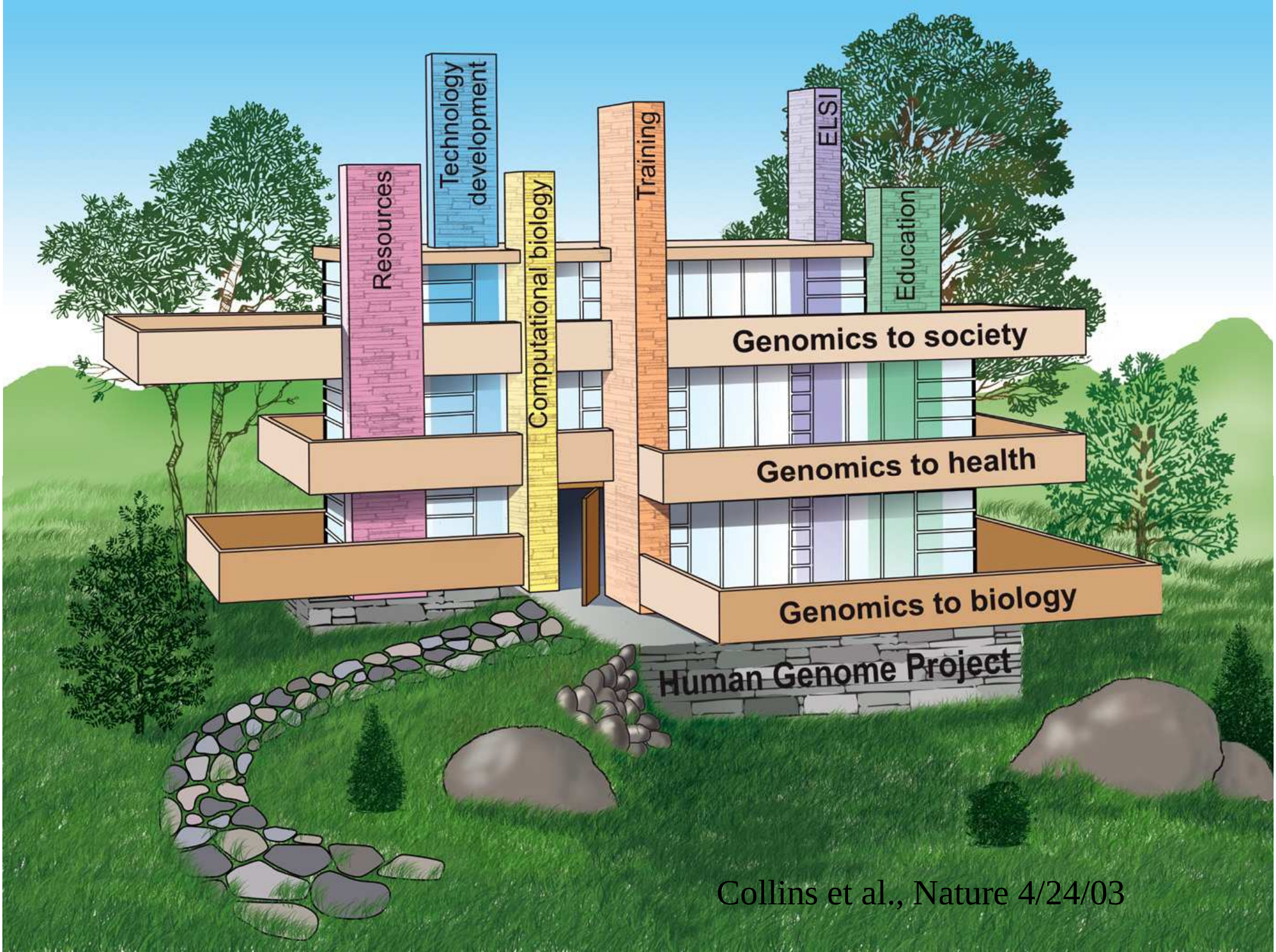
Credits: Data: Ewan Birney, Jim Kent, Krishna M. Roskin, Jim Mullikin, Daryl Thomas, Robert Baertsch. Design: Darryl Leja

National Human Genome Research Institute

nature

All of the original goals of the Human Genome Project have been accomplished

What's next?



Collins et al., Nature 4/24/03

Personalized Medicine?

What is a biobank?

- ***BANKING*** – The process of storing material or specimens for future use
- An entity that receives, stores, processes and/or disseminates specimens, as needed. It encompasses the physical location as well as the full range of activities associated with its operation. It may also be referred to as a BIOREPOSITORY.

Published in 2008, Best Practices for Repositories, Cell Preservation Technology, *Collection, Storage, Retrieval and Distribution of Biological Materials for Research* in International Society for Biological and Environmental Repositories (ISBER).

Advantage India

Population >1 Billion

- World Class Clinical facilities: Public and Private
- Manpower: Chemists,GCP trained clinicians, scientists, IT
- Industry: Strong pharma, software
- Largest English speaking country in South Asia
- Low labour costs
- World class GMP facility

Population >1 Billion

- DBT, DST, CSIR, ICMR, UGC, ICAR
- Institutions

Challenges and opportunities for Biobanking in India

- **Need to assess the available resources in the country and how important these are to be considered as biobanking resources:**

EXAMPLE

1999 RAND study in the USA:

- ‘conservatively estimated that there were more than 307 million human biospecimens from more than 178 million cases stored in the United States, accumulating at a rate of more than 20 million specimens per year.

Challenges

- **Involving Indian Pharma without a backleash**

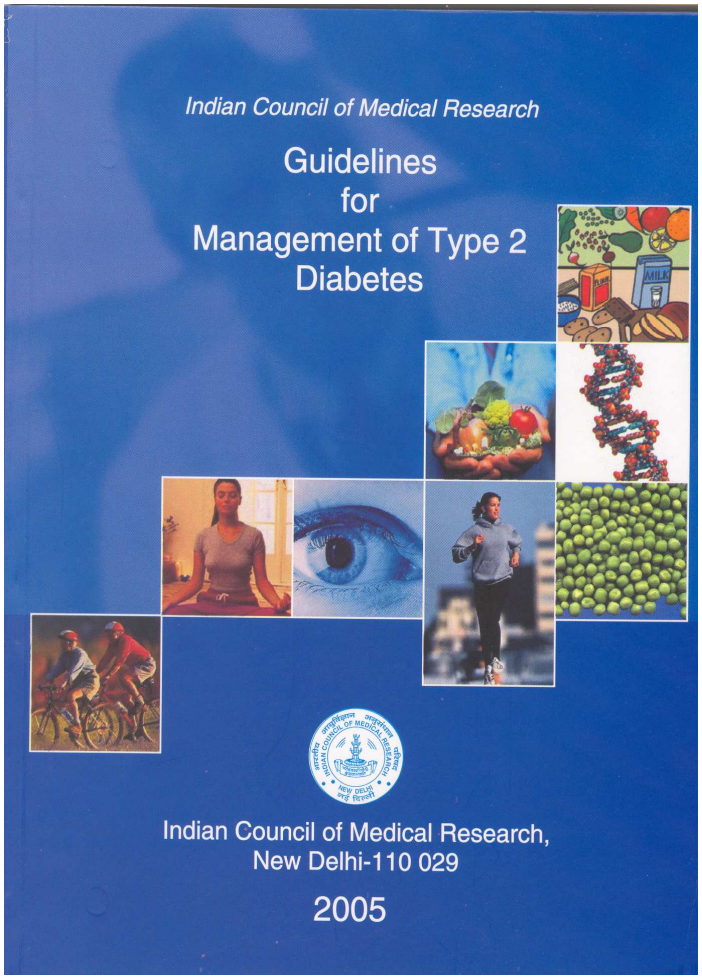
- **Need for regulations:**

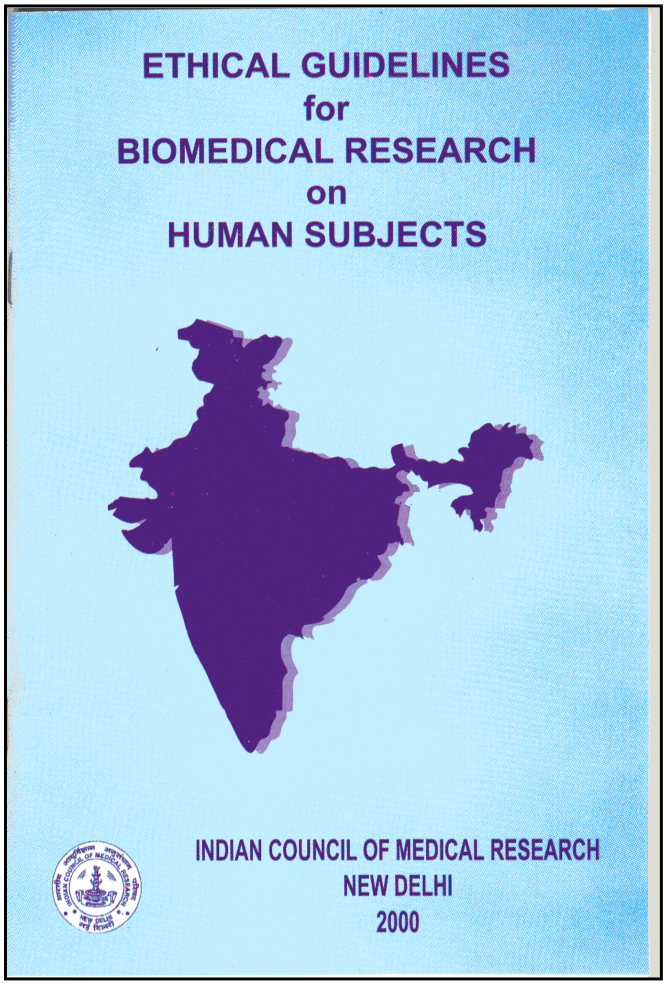
Ethical Issues Related to Biobanks

- Anonymity
- Informed Consent

Strengths

Guidelines prepared



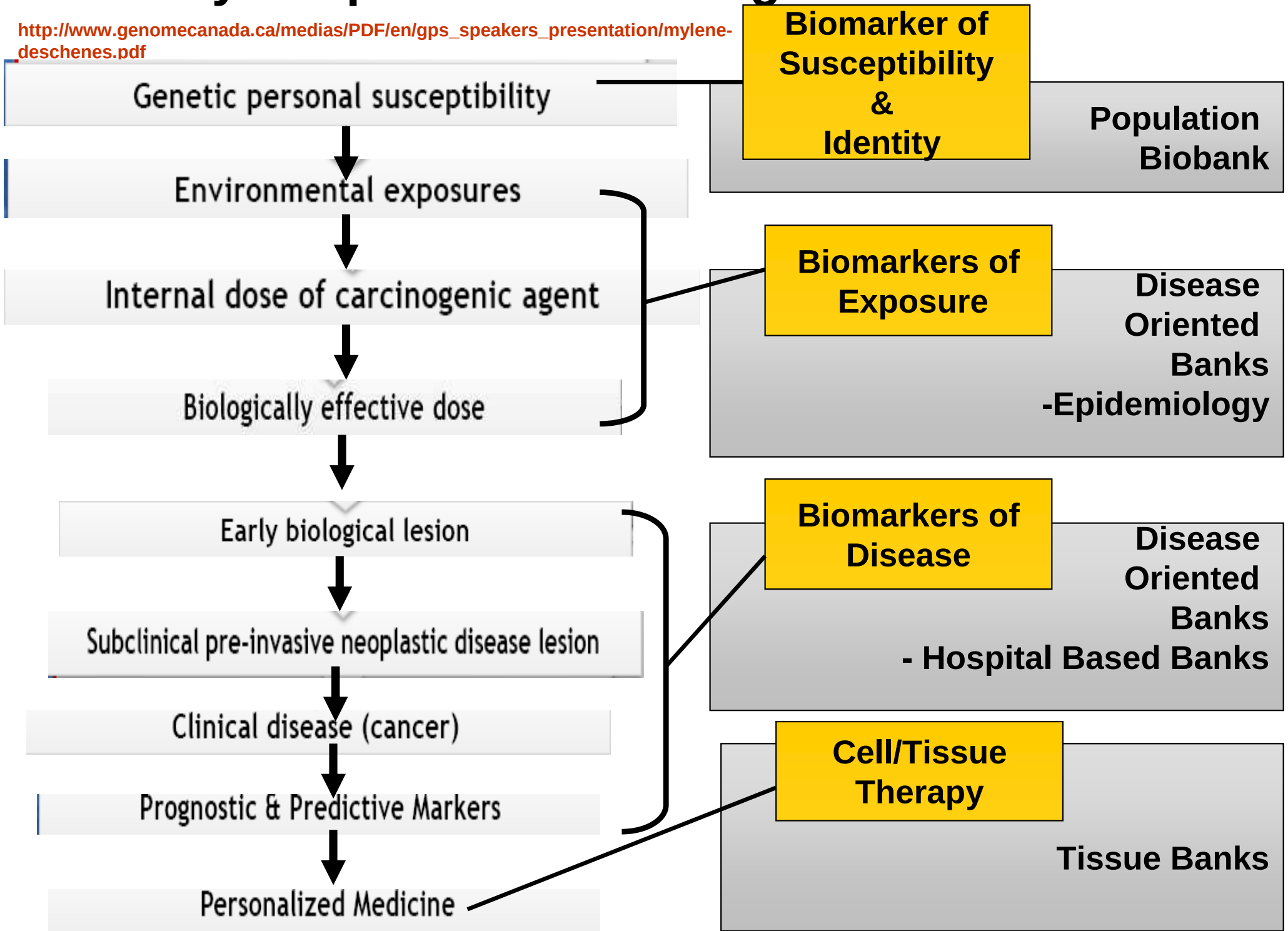


Challenges

Globality of Biobanks: Major Investments around the world

Diversity : Importance of Design

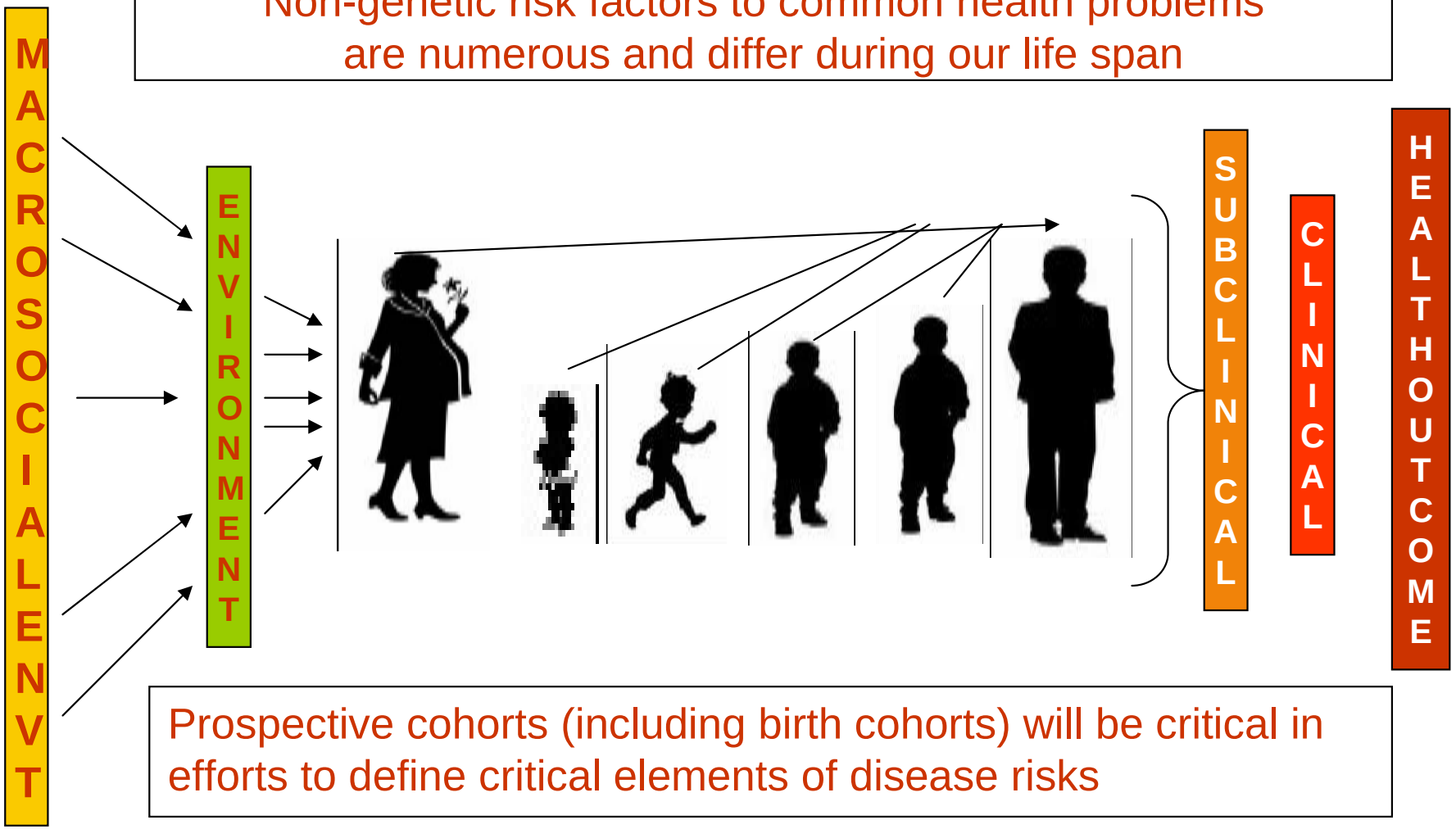
http://www.genomecanada.ca/medias/PDF/en/gps_speakers_presentation/mylene-deschenes.pdf



Challenges for Biobanking in India

PROSPECTIVE LARGE COHORTS ???

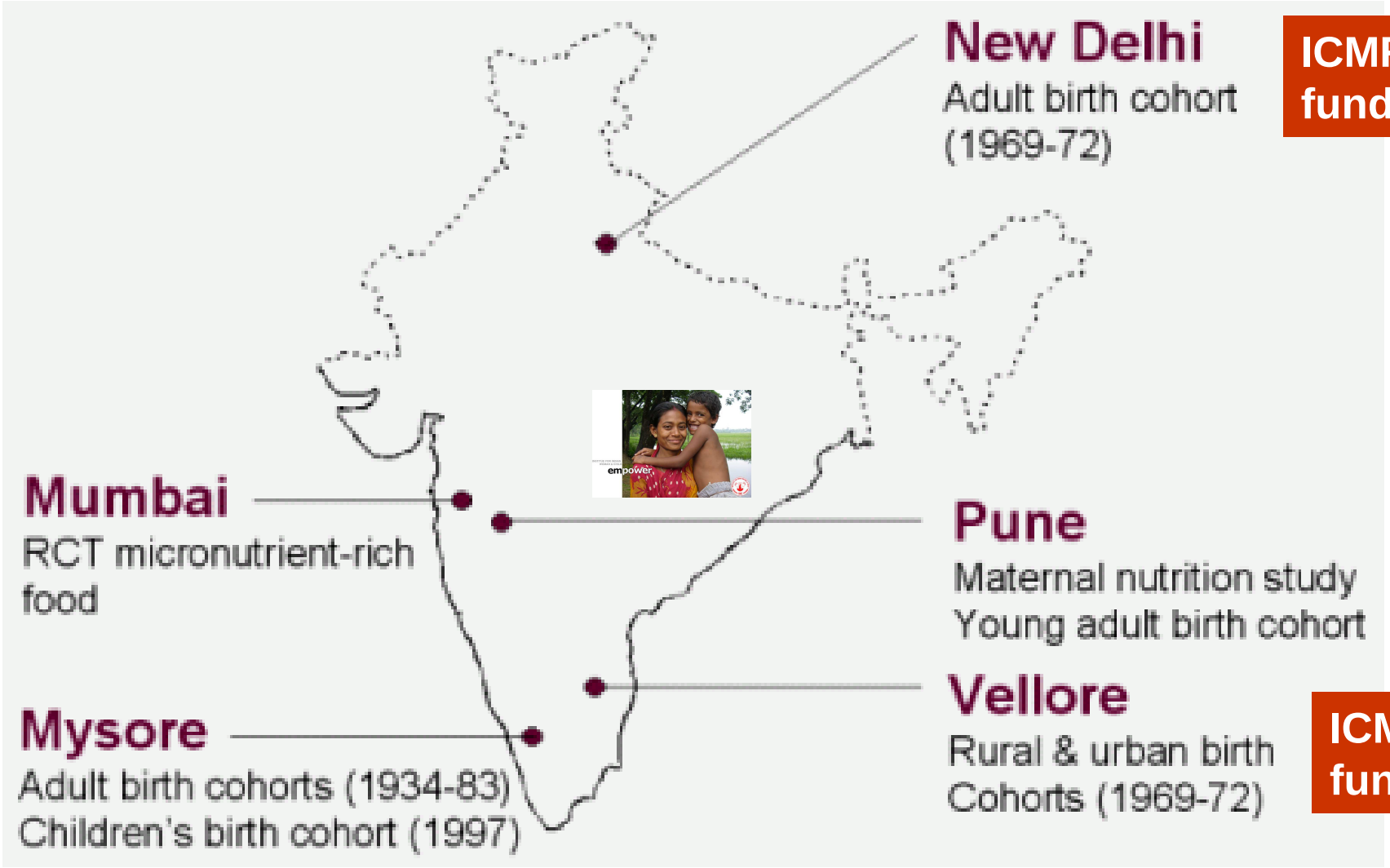
Non-genetic risk factors to common health problems are numerous and differ during our life span



Prospective cohorts (including birth cohorts) will be critical in efforts to define critical elements of disease risks

Opportunities

India: Birth Cohorts



India: Birth cohorts

Cohort	Year	Sample Size
Delhi Birth Cohort	1969-73	7119 babies, Adults (1998-2002) = 2584; consented: 1583
Vellore Birth cohort	1969-73	5753 babies, Adults (1998-2002) = 2218
Pune Maternal Nutrition Study	1994-96	631 babies born in six rural villages 338 babies in Southampton
Mumbai Cohort Study	2006	randomized controlled trial of a food-based micronutrient-rich supplement
Mysore Birth Records Studies **Mysore Parthenon Study**	1934-83 1997	816 adults 832 babies

Our Strengths: Jai Vigyan Mission Mode Project Control of RF/ RHD in India

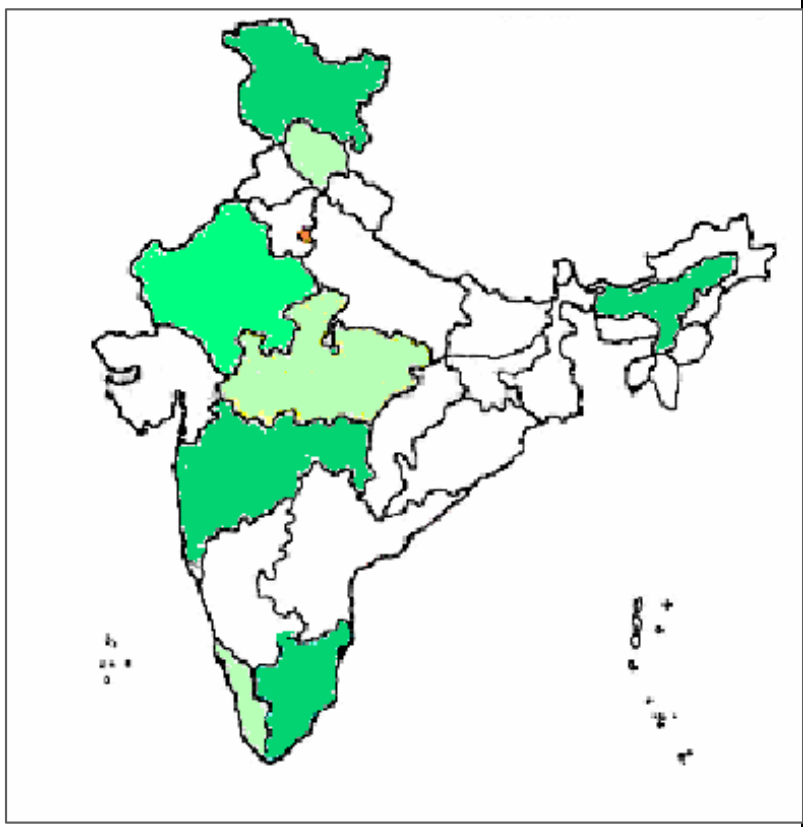
Under Prime Minister's initiative to boost R & D in Science

Duration: Started in Year 2000 ONGOING

Coordinating Centre: ICMR Hqrs, New Delhi

Centres/ Registries

1. PGIMER, Chandigarh
2. CMC, Vellore, Tamilnadu
3. Amrita Institute Medical Sciences, Kochi, Kerala
4. M.G.M.Medical College,Indore, MP
5. Govt Medical college, Jammu, J & K
6. Indira Gandhi Medical College , Shimla, HP
7. DMRC, ICMR Institute, Jodhpur, Rajasthan
8. KEM Hospital, Mumbai, Maharashtra
9. RMRC, ICMR Institute, Dibrugarh, Assam
10. AIMS, Wayanad, Kerala



Our Strengths: ICMR-WHO NCD Surveillance Centers for NCD Risk Factors in India

Coordinated at Div. of NCD, ICMR (2003- 2006)

Centres

IBHAS, Delhi,	MDRF, Chennai	GMC, Nagpur
AIIMS, Delhi	SCTIMS,Trivandrum	RMRC, Dibrugarh

NCD Risk Factors

- Alcohol consumption
- Tobacco consumption
- Intake of fruits and vegetables
- Physical inactivity
- Obesity (BMI, Waist circumference)
- Blood pressure
- Biochemical Risk Factors- glucose,cholesterol, HDLc, triglycerides

Our Strengths

Biobanks: Indian Scenario

Brain Biobank

ICMR, DST, DBT funded

NIMHANS, Bangalore

Cancer Biobank

ACTREC, Mumbai

The **Advanced Centre for Treatment, Research and Education in Cancer**

Organ Retrieval Banking Organization (ORBO), AIIMS, New Delhi

National Repository for Cell Lines / Hybridomas, NCCS, Pune

Mycobacterial Repository, JALMA, Agra

Challenges for Biobanking in India

- **Cost: Need for technological innovations:**

Example: Sample collection

Dried Blood Spots (DBS) instead serum samples

Benefits

- Do not require to be kept in cold continuously
- Less Refusals: Only 5% refusals to give dried blood samples, compared with nearly 40 % for serum samples.
- Less costly: US$10 per person, compared to US$240 per person for the serum-based UK Biobank.

S. K. Sgaier et al. Science 2007

A Recent ICMR Initiative

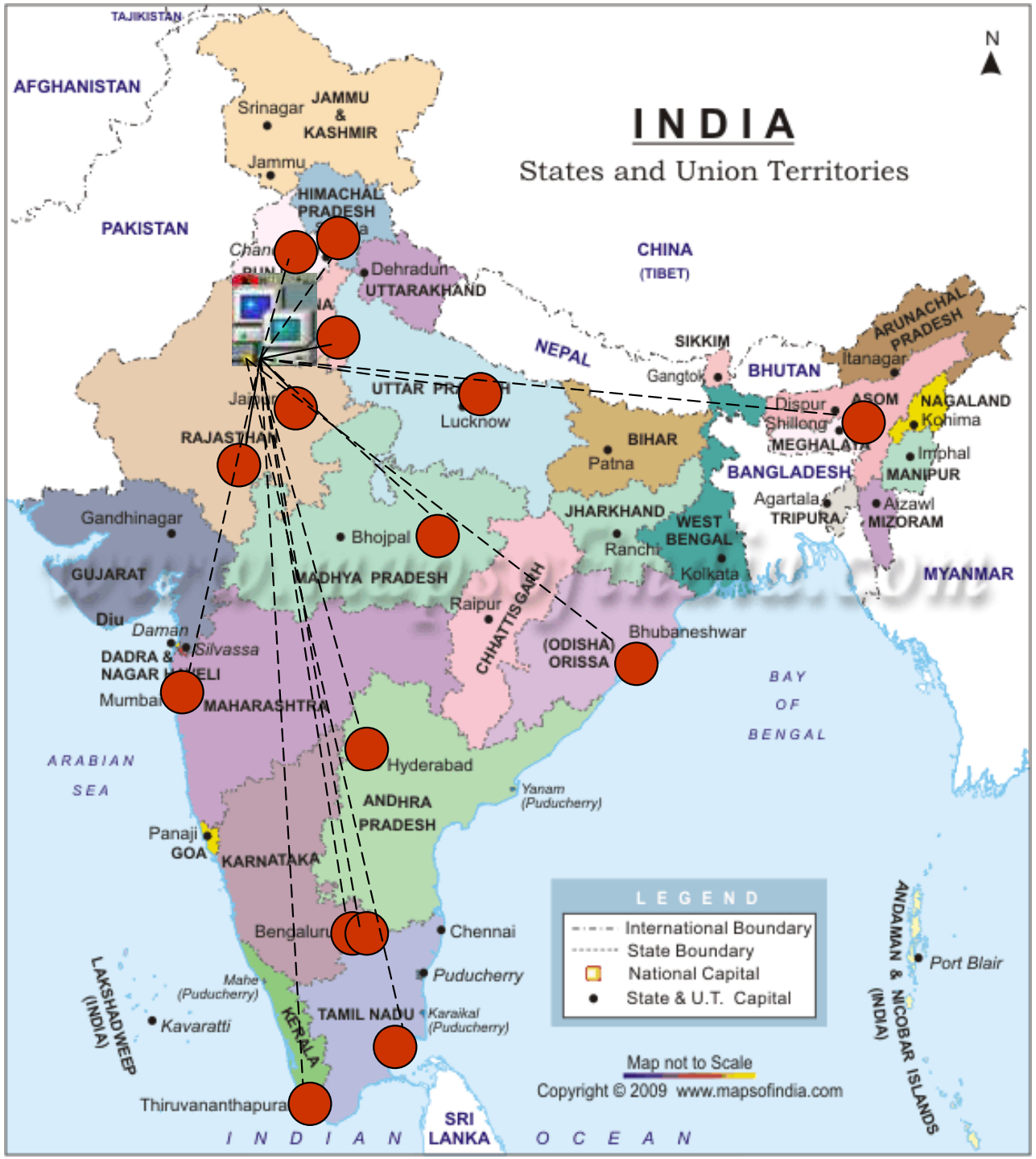
Management of Acute Coronary Event (MACE) Registries

NODAL Centre:

Div of NCD, ICMR Hqrs.

Clinical Coordinating Centre:

St Johns, Bangalore



Management of Acute Coronary Event (MACE) Registries

Feasibility study: 10-15 studies
Pilot Study: 100 Centres
Main study; As many centres as possible

- Demographics
- Medical History
- Management Practices
- Follow Ups

OBJECTIVES

Objectives

To ascertain the characteristics of hospitalized ACS patients

To study the current trends in management of ACS patients, use of guidelines in Indian context and assess their impact on outcomes in ACS patients

To study the relationship of routine invasive strategy and adverse outcomes at 6 months.

Designing a biobank for biological samples collected from a megastudy on acute coronary events in Indians (ICMR-INSERM (Indo French) collaboration)

- **Objectives:**

PHASE I (Current study)

- To develop a protocol for sample collection, handling, transport and storage
- To pilot the protocol in three Indian centers

PHASE II

- To expand the pilot study so as to develop a network of collaborative centres across the country for long term prospective collection of biological samples in the field of CVDs.

Biobank facility at ICPO

Centres : PGIMER, Chandigarh, AIIMS, New Delhi and St John's, Bangalore

Development of Biomarkers for CVDs using Cohorts, Registries and Biobanks

- Requires A Large Sample Size
  Well Annotated Biological Samples
  Requires Synergies & Connections to be Created Between Different Stages of Scientific Endeavor & Clinical Practice

A network linking Cohorts, BIOBANK and ACE/CVD registries

Linking Biobank with Quality Assured Registries Enables Population Representative Studies With Minimum Case Ascertainment Bias

THANK YOU