Transcript Isolation Techniques

Larry Brody
National Human Genome Research Institute
Bldg. 49 Rm. 3A14
Phone (301) 496-7824
FAX (301) 402-4929
email lbrody@helix.nih.gov

October 1997

Points to Consider

Advances in the end game during the last decade: The incremental and the quantum

How is/will the Genome project changing the process.

If you have a region of interest now, what is the best way to find "your" gene

The problem with complex traits vs. Mendelian

What size interval is feasible to tackle
- genetic vs. physical maps
- gene density considerations

Technical and person year requirements

Lecture Outline

Phases of a project that has gone "Post Linkage"

I. Contig building
Reagents
Methods

II. Gene Hunting

Physical methods
- functional cloning
-direct screening
- direct selection
- exon trapping
- sequencing

Virtual Methods
- database
- virtual evolution

Completing the gene file
- cDNA based
- genomic based

III. Candidate gene triage
-strategies
- informatics considerations

IV. Searching for mutations

- methods of choice circa 1997
- confirmation that this is the correct gene
- material and computational considerations

Transcript Isolation Techniques

FUNCTIONAL CLONING (complementation cloning)

- variation on old theme
- can use large insert DNA clones
  - YAC, PAC, BACS, P1
- requires robust assay
- not practical for many phenotypes
- false positives

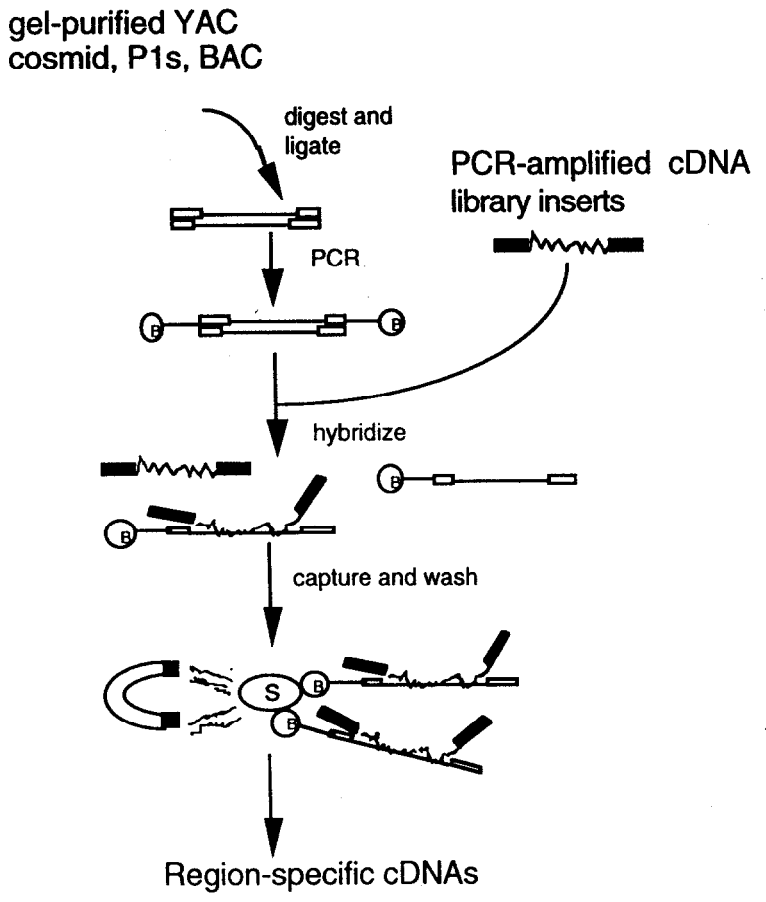
DIRECT SCREENING

Screen libraries directly with large insert clones
- cheap
- fast
- clone size a function of the library

- low yield
- requires repeat suppression

Direct Selection

- a.k.a. magnetic capture, cDNA selection
  - variation: Gene Trapper (LTI)
- utilizes cDNA libraries or primary RT-PCR product
- acts to normalize up to 10^5
- rare and absent messages missed
- significant background ~50%
- short fragments
- YACs, BACs, P1s, cosmids



Magnetic Bead Trapping of Region-Specific cDNAs.

Exon Trapping

- a.k.a. Exon Amplification
  - internal and 3' kits (BRL)
- utilizes genomic DNA
- independent of expression level
- ultimate form of normalization
- significant false positives
- short fragments, average exon 200-400 bp
- YACs, BACs, P1s, cosmids
  - intact cosmid vector avail.(van Ommen)

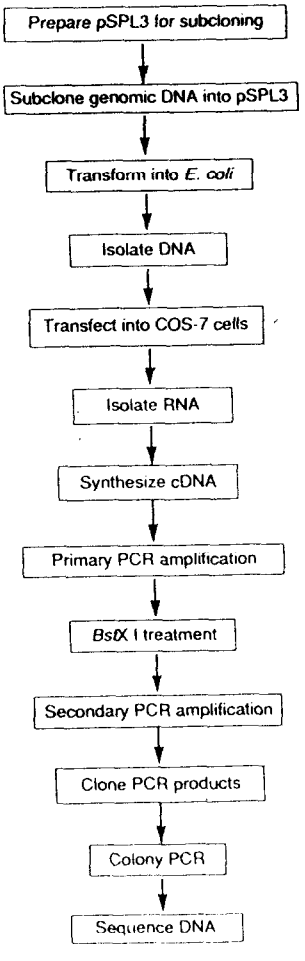


Figure 1. Schematic of the exon trapping procedure

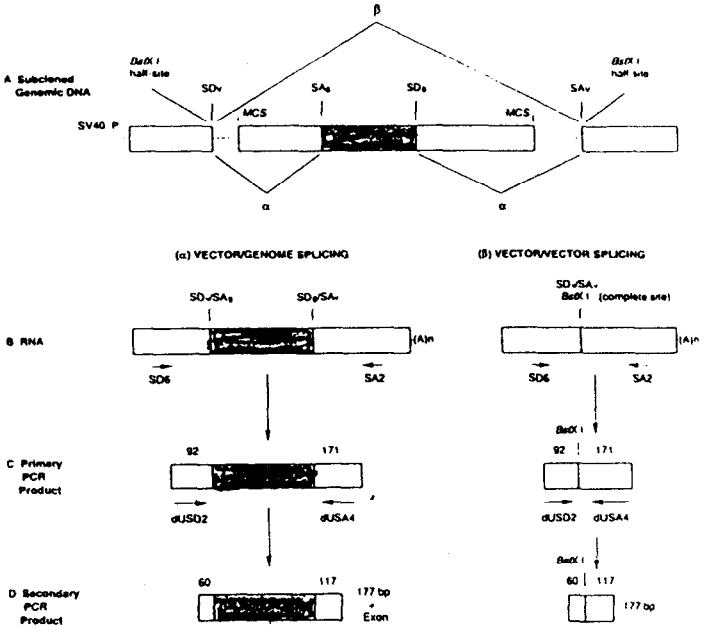


Figure 3. Potential splicing events in the exon trapping procedure. *Panel A*, Subcloned genomic DNA containing an exon. *Panel B*, RNA from two types of splicing events (α and β) are shown, and the products of these splicing events are depicted in the subsequent panels. *Panel C*, RNA from transfected COS-7 cells is isolated, converted to cDNA, and amplified in the primary PCR reaction. *Panel D*, The amplified DNA is reamplified by PCR. SD_v=vector splice donor, SD_g=genomic DNA splice donor, SA_v=vector splice acceptor, SA_g=genomic splice acceptor, MCS=multiple cloning site.

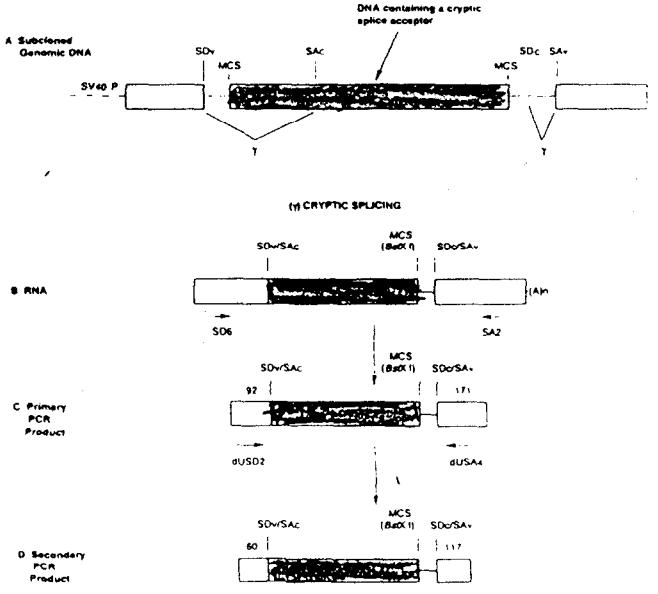


Figure 5. Cryptic splicing. A genomic DNA fragment containing a cryptic splice acceptor (SA_c) is subcloned into pSPL3. With cryptic splicing, part of the MCS containing a *BstX* I recognition site is retained, allowing digestion by that enzyme (not shown), which prevents PCR amplification. See figures 2 and 4 for a description of symbols.

SEQUENCING

"Complete" sequence not needed

"Single pass" sequence methods

- shotgun clone insert
- sequence a set number of clones

- cost considerations
  - 100 kb @ ~ $ 25/ run
    400 runs > $ 10,000

Locating genes in raw genomic sequence

- gene identification software
- dbest approaches

- data management and informatics

Transcript Isolation - References

* recomended reading

GENERAL REVIEW

Monaco AP
Isolation of genes from cloned DNA. Curr Opin Genet Dev 1994 Jun;4(3):360-5

DIRECT SCREENING

Elvin P, Slynn G, Black D, Graham A, Butler R, Riley J, Anand R, Markham AF. Isolation of cDNA clones using yeast artificial chromosome probes. Nuc Acids Res 1990;18:3913-3917

Friedman LS; Ostermeyer EA; Lynch ED; Welcsh P; Szabo CI; Meza JE; Anderson LA; Dowd P; Lee MK; Rowell SE; et al 22 genes from chromosome 17q21: cloning, sequencing, andcharacterization of mutationsin breast cancer families andtumors. Genomics 1995 Jan 1;25(1):256-63

van der Steege G; Draaijers TG; Grootscholten PM; Osinga J; Anzevino R; Velona I; Den Dunnen JT; Scheffer H; Brahe C; va A provisional transcript map of the spinal muscular atrophy (SMA) critical region. Eur J Hum Genet 1995;3(2):87-95

*Wallace MR; Marchuk DA; Andersen LB; Letcher R; Odeh HM; Saulino AM; Fountain JW; Brereton A; Nicholson J; Mitchell AL; Type 1 neurofibromatosis gene: identification of a large transcript disrupted in three NF1 patients [published erratum appears in Science 1990 Dec 21;250(4988):1749] Science 1990 Jul 13;249(4965):181-6

SEQUENCE BASED TRANSCRIPT ISOLATION

Fairhead C; Dujon B Transcript map of two regions from chromosome XI of Saccharomyces cerevisiae for interpretation of systematic sequencing results. Yeast 1994 Nov;10(11):1403-13

Thomas A, Skolnick MH A probalistic model fopr detecting coding regions in DNA sequences. J. Math. Appl. Med. Biol. 1994 11:149-160

Snyder EE, Stormo GD Identification of coding regions in genomic DNA sequences: an application of dynamic programming and neural networks. Nuc. Acids Res. 1993; 21:607-613

*Uberbacher EC, Mural RJ Locating protein coding regions in human DNA sequences by a multiple sensor-neural network approach. Proc. Natl. Acad. Sci USA 1991; 88:11261-11265

INTERNET RESOURCES

GRAIL
http://avalon.epm.ornl.gov/Grail-bin/

GENEFINDER
http://dot.imgen.bcm.tmc.edu:9331/gene-finder/gf.html

GENETICS LINK OF LINKS
HARVARD http://golgi.harvard.edu/genome.html
BAYLOR http://gc.bcm.tmc.edu:8088/home.html

DIRECT SELECTION

Chiannilkulchai N; Pasturaud P; Richard I; Auffray C; Beckmann JS A primary expression map of the chromosome 15q15 region containing the recessive form of limb-girdle muscular dystrophy (LGMD2A) gene. Hum Mol Genet 1995 Apr;4(4):717-25

Ellis NA; Groden J; Ye TZ; Straughen J; Lennon DJ; Ciocci S; Proytcheva M; German J The Bloom's syndrome gene product is homologous to RecQ helicases. Cell 1995 Nov 17;83(4):655-66

*Hattier T, Bell R, Shaffer D, Stone S, Phelps RS, Tavtigian SV, Skolnick MH, Shattuck-Eidens D, Kamb A. Monitoring the efficacy of hybrid selection during positional cloning: the search for BRCA1. Mammalian Genome 1995 6:873-879

Jacob A; Kandpal G; Patanjali SR; Kandpal RP Molecular cloning and expression pattern of genes from a 470 Kb region near BRCA1 locus on chromosome 17q21. Oncogene 1995 Sep 7;11(5):981-6

Kurahashi H; Akagi K; Inazawa J; Ohta T; Niikawa N; Kayatani F; Sano T; Okada S; Nishisho I Isolation and characterization of a novel gene deleted in DiGeorge syndrome. Hum Mol Genet 1995 Apr;4(4):541-9

*Lovett M. Fishing for complements: finding genes by direct selection. Trends Genet 1994 Oct;10(10):352-7

Osborne-Lawrence S; Welcsh PL; Spillman M; Chandrasekharappa SC; Gallardo TD; Lovett M; Bowcock AM. Direct selection of expressed sequences within a 1-Mb region flanking BRCA1 on human chromosome 17q21.Genomics 1995 25, 2488-255

*Parimoo S. Patanjali SR, Shukla H, Chaplin DD, Weissman SM. CDNA selection: efficient PCR approach for the selection of cDNAs encoded in large chromosomal DNA fragments. Proc. Natl Acad. Sci. USA 1991 88:9623-9627

Peterson A; Patil N; Robbins C; Wang L; Cox DR; Myers RM A transcript map of the Down syndrome critical region on chromosome 21. Hum Mol Genet 1994 Oct;3(10):1735-42

Rommens JM; Lin B; Hutchinson GB; Andrew SE; Goldberg YP; Glaves ML; Graham R; Lai V; McArthur J; Nasir J; et al A transcription map of the region containing the Huntington disease gene [published erratum appears in Hum Mol Genet 1993 Sep;2(9):1524] Hum Mol Genet 1993 Jul;2(7):901-7

Rouquier S; Trask BJ; Taviaux S; van den Engh G; Diriong S; Lennon GG; Giorgi D Direct selection of cDNAs using whole chromosomes. Nucleic Acids Res 1995 Nov 11;23(21):4415-20

Simmons AD; Goodart SA; Gallardo TD; Overhauser J; Lovett M Five novel genes from the cri-du-chat critical region isolated by direct selection. Hum Mol Genet 1995 Feb;4(2):295-302

*Tagle DA; Swaroop M; Lovett M; Collins FS Magnetic bead capture of expressed sequences encoded within large genomic segments. Nature 1993 Feb 25;361(6414):751-3

Yamakawa K; Mitchell S; Hubert R; Chen XN; Colbern S; Huo YK; Gadomski C; Kim UJ; Korenberg JR Isolation and characterization of a candidate gene forprogressive myoclonus epilepsy on 21q22.3. Hum Mol Genet 1995 Apr;4(4):709-16

EXON TRAPPING/AMPLIFICATION

Bar I; Lambert De Rouvroit C; Royaux I; Krizman DB; Dernoncourt C; Ruelle D; Beckers MC; Goffinet AM A YAC contig containing the reeler locus with preliminary characterization of candidate gene fragments. Genomics 1995 Apr 10;26(3):543-9

*Buckler AJ; Chang DD; Graw SL; Brook JD; Haber DA; Sharp PA; Housman DE Exon amplification: a strategy to isolate mammalian genes based on RNA splicing. Proc Natl Acad Sci U S A 1991 May 1;88(9):4005-9

Burfoot MS; Campbell RD Improved method of gene detection using exon amplification. Nucleic Acids Res 1994 Dec 11;22(24):5510-1

Church DM; Banks LT; Rogers AC; Graw SL; Housman DE; Gusella JF; Buckler AJ. Identification of human chromosome 9 specific genes using exon amplification. Hum Mol Genet 1993 Nov;2(11):1915-20

*Church DM; Stotler CJ; Rutter JL; Murrell JR; Trofatter JA; Buckler AJ Isolation of genes from complex sources of mammalian genomic DNA sing exon amplification. Nat Genet 1994 Jan;6(1):98-105

Datson NA; Duyk GM; Van Ommen JB; Den Dunnen JT
Specific isolation of 3'-terminal exons of human genes by exon trapping. Nucleic Acids Res 1994 Oct 11;22(20):4148-53

*Duyk GM; Kim SW; Myers RM; Cox DR
Exon trapping: a genetic screen to identify candidate transcribed sequences in cloned mammalian genomic DNA. Proc Natl Acad Sci U S A 1990 Nov;87(22):8995-9

*Gibson F; Lehrach H; Buckler AJ; Brown SD; North MA
Isolation of conserved sequences from yeast artificial chromosomes by exon amplification. Biotechniques 1994 Mar;16(3):453-8

Koyama K; Sudo K; Nakamura Y
Isolation of 115 human chromosome 8-specific expressed-sequence tags by exon amplification. Genomics 1995 Mar 20;26(2):245-53

*Krizman DB; Berget SM
Efficient selection of 3'-terminal exons from vertebrate DNA. Nucleic Acids Res 1993 Nov 11;21(22):5198-202

Krizman DB; Hofmann TA; DeSilva U; Green ED; Meltzer PS; Trent JM
Identification of 3'-terminal exons from yeast artificial chromosomes. PCR Methods Appl 1995 Jun;4(6):322-6

Lerner TJ; D'Arigo KL; Haines JL; Doggett NA; Taschner PE; de Vos N; Buckler AJ
Isolation of genes from the Batten candidate region using exon amplification. Batten Disease Consortium. Am J Med Genet 1995 Jun 5;57(2):320-3

Lucente D; Chen HM; Shea D; Samec SN; Rutter M; Chrast R; Rossier C; Buckler A; Antonarakis SE; McCormick MK.
Localization of 102 exons to a 2.5 Mb region involved in Down syndrome. Hum Mol Genet 1995 Aug;4(8):1305-11

Nehls M; Pfeifer D; Boehm T
Exon amplification from complete libraries of genomic DNA using a novel phage vector with automatic plasmid excision facility: application to the mouse neurofibromatosis-1 locus. Oncogene 1994 Aug;9(8):2169-75

North MA; Sanseau P; Buckler AJ; Church D; Jackson A; Patel K; Trowsdale J; Lehrach H
Efficiency and specificity of gene isolation by exon amplification. Mamm Genome 1993 Sep;4(9):466-74

Walker AP; Muscatelli F; Monaco AP
Isolation of the human Xp21 glycerol kinase gene by positional cloning. Hum Mol Genet 1993 Feb;2(2):107-14 Comment in: Hum Mol Genet 1993 Feb;2(2):95-6

Yaspo ML; Gellen L; Mott R; Korn B; Nizetic D; Poustka AM; Lehrach H
Model for a transcript map of human chromosome 21: isolation of new coding sequences from exon and enriched cDNA libraries. Hum Mol Genet 1995 Aug;4(8):1291-304

Yaspo ML; North MA; Lehrach H
Exon-enriched probe derived from a human chromosome 21 YAC by exon-amplification. Nucleic Acids Res 1993 May 11;21(9):2271-2

INTEGRATED TRANSCRIPT MAPPING EFFORTS

Brody LC; Abel KJ; Castilla LH; Couch FJ; McKinley DR; Yin G; Ho PP; Merajver S; Chandrasekharappa SC; Xu J; et al Construction of a transcription map surrounding the BRCA1 locus of human chromosome 17. Genomics 1995 Jan 1;25(1):238-47

Gecz J; Villard L; Lossi AM; Millasseau P; Djabali M; Fontes M
Physical and transcriptional mapping of DXS56-PGK1 1 Mb region: identification of three new transcripts. Hum Mol Genet 1993 Sep;2(9):1389-96

Harshman K; Bell R; Rosenthal J; Katcher H; Miki Y; Swenson J; Gholami Z; Frye C; Ding W; Dayananth P; et al
Comparison of the positional cloning methods used to isolate the BRCA1 gene. Hum Mol Genet 1995 Aug;4(8):1259-66

*Segre JA; Nemhauser JL; Taylor BA; Nadeau JH; Lander ES
Positional cloning of the nude locus: genetic, physical, and transcription maps of the region and mutations in the mouse and rat. Genomics 1995 Aug 10;28(3):549-59

Yaspo ML; Gellen L; Mott R; Korn B; Nizetic D; Poustka AM; Lehrach H
Model for a transcript map of human chromosome 21: isolation of new coding sequences from exon and enriched cDNA libraries. Hum Mol Genet 1995 Aug;4(8):1291-304

CpG ISLAND BASED METHODS

*Antequera F; Bird A
Number of CpG islands and genes in human and mouse. Proc Natl Acad Sci U S A 1993 Dec 15;90(24):11995-9

John RM; Robbins CA; Myers RM
Identification of genes within CpG-enriched DNA from human chromosome 4p16.3. Hum Mol Genet 1994 Sep;3(9):1611-6

Valdes JM; Tagle DA; Collins FS
Island rescue PCR: a rapid and efficient method for isolatingtranscribed sequences from yeast artificial chromosomes andcosmids. Proc Natl Acad Sci U S A 1994 Jun 7;91(12):5377-81

MUTATION DETECTION METHODS

REVIEWS

Cotton RG.
Current methods of mutation detection. Mutation Research 1993 285:125-144

Grompe M,
The rapid detection of unknown mutations in nucleic acids. Nature Genetics 1993 5:111-117

CHEMICAL MIS-MATCH CLEAVAGE

Cotton RGH, Dahl HHM, Forrest S, Howells DN, Ramus SJ, Bishop RE, Dianzi I, Saleeba JA, Palombo E, Anderson MJ, Milner CM, Campbell RD.
Analysis of sequence contexts flanking T.G. mismatches leadts to predictions about reactivity of the mismatched T to osmium tetroxide. DNA Cell Biol. 12:945-949

Cotton RGH, Rodrigues NR, Campbell RD.
Reactivity of cytosine and thymine in single-base-pair mismatch with hydroxylamine and osmium tetroxide and its application to the study of mutations. Proc. Natl. Acad. Sci. USA 1988 85:4397-4401

*Rowley G, Saad S, Giannelli F, Green PM.
Ultrarapid mutation detection by multiplex, solid-phase chemical cleavage. Genomics 1995 30:574-582

SSCP (Single -strand conformation polymorphism analysis)

Bottema CD; Sarkar G; Cassady JD; Ii S; Dutton CM; Sommer SS
Polymerase chain reaction amplification of specific alleles: a general method of detection of mutations, polymorphisms, and haplotypes. Methods Enzymol 1993;218:388-402

Bottema CD; Sommer SS
PCR amplification of specific alleles: rapid detection of known mutations and polymorphisms. Mutat Res 1993 Jul;288(1):93-102

*Blaszyk H; Hartmann A; Schroeder JJ; McGovern RM; Sommer SS; Kovach JS
Rapid and efficient screening for p53 gene mutations by dideoxy fingerprinting. Biotechniques 1995 Feb;18(2):256-60

Liu Q; Sommer SS
Parameters affecting the sensitivities of dideoxy fingerprinting and SSCP. PCR Methods Appl 1994 Oct;4(2):97-108

*Liu Q; Sommer SS
Restriction endonuclease fingerprinting (REF): a sensitive method for screening mutations in long, contiguous segments of DNA. Biotechniques 1995 Mar;18(3):470-7

*Orita MY, Suzuki Y, Sekiya T, Hatashi K.
Detection of polymorphism of human DNA by gel electrophoresis as single-strand conformation polymorphism. Proc. Natl. Acad. Sci. USA 1992 89:2331-2335