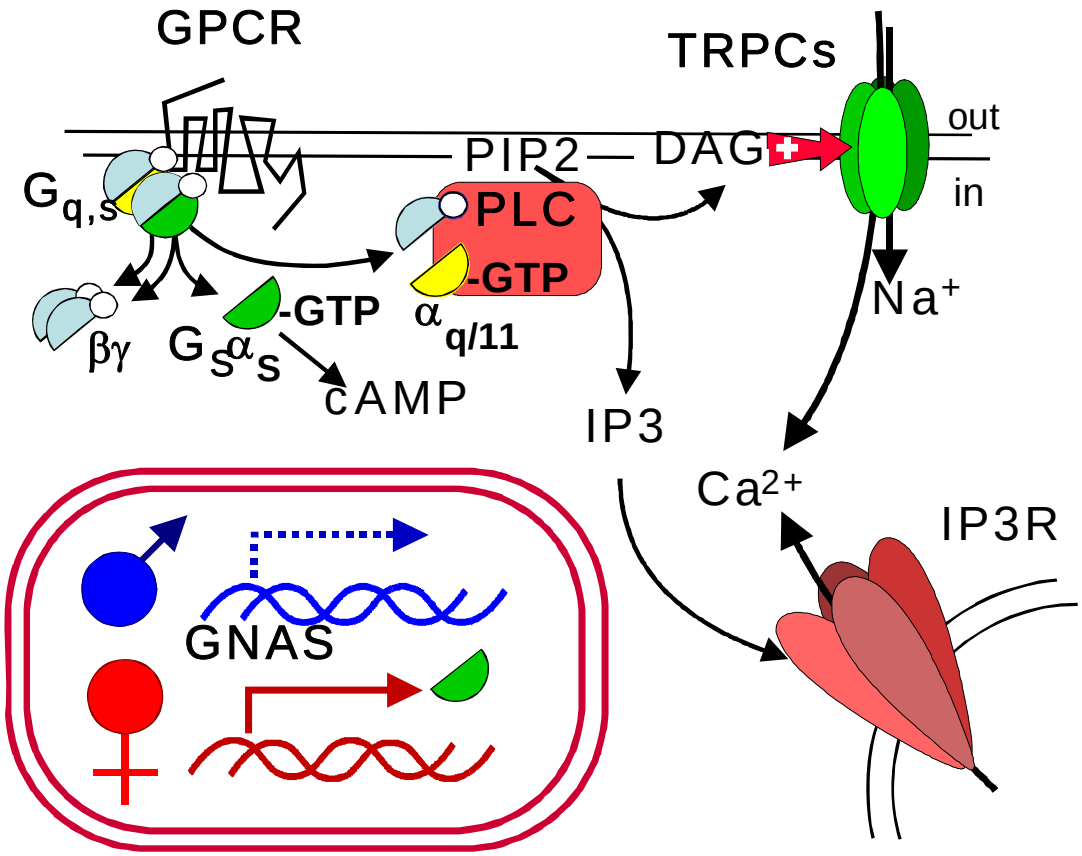
Transmembrane Signaling Group

Lutz Birnbaumer, Head



Project 4 Section of the April 2006 Progress Report from the Transmembrane Sigaling Group submitted to the NIEHS Board of Scientific Councilors.

Project 4: Epigenetic Regulation of Gene Expression: Imprinting at the GNAS (human)/Gnas (mouse) locus.

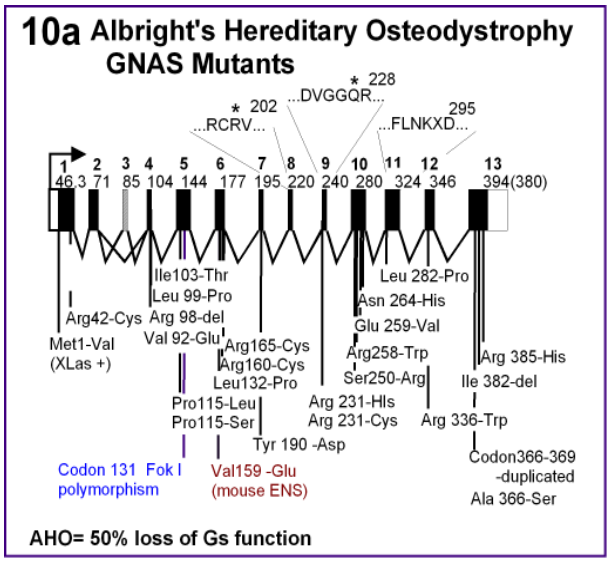
Epigenetics. Epigenetics has two meanings. To some it has a very broad meaning and is the study of chromatin states that are inherited at the cellular level from cell to cell and is responsible for permanent changes in gene expression, such as occur during development of tissues and organs. To others, epigenetics has a more restricted meaning to include only changes in chromatin structure that are inherited from one progeny to another without changes in the primary DNA sequence. Imprinting of genes is of the second type. Transgenerational effects caused by drugs or maternal behavior during the perinatal and early postnatal period of life also belong to the second "narrower" definition of epigentics.

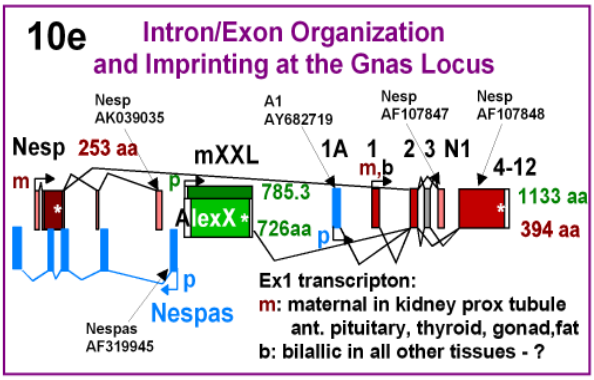
Two recent hallmark papers illustrate the relevance of epigenetic changes in chromatin structure for environmental health sciences and toxicology. Both studied the methylation status of DNA, i.e., methylation of CpG dinucleotides on position 5 of the cytidine to give ^{me}CpG dinucleotides. The first paper is from Meaney's group at McGill in Montreal, Canada, who showed that a genetically programmed intense licking and grooming (LG) by mothers altered the methylation status of the glucocorticoid receptor promoter in the hypothalamus and the capacity of the pups to cope with stress in their adult life. Switching pups from high LG mothers to low LG mothers, led to reversal of DNA methylation and switching pups from low LG to high LG mothers caused the opposite effect on promoter methylation and changed their stress responses at adulthood. Promoter methylation caused changes in histone acetylation and chromatin structure detectable in the adult (*Weaver et al. Nature Neuroscinces 2004*). The second paper is from Skinner's group in Seattle, and showed that a single dose of an endocrine disruptor (vinclozolin) given to pregnant mice during a critical period of gestation, lead to altered DNA methylation patterns and male infertility with diminished spermatogenesis, not only in the exposed fetuses, but also in their offspring and the offspring of their offspring, i.e. lead to structural chromatin alterations that were inherited without changes in DNA sequence (*Anway et al, Science 2005*).

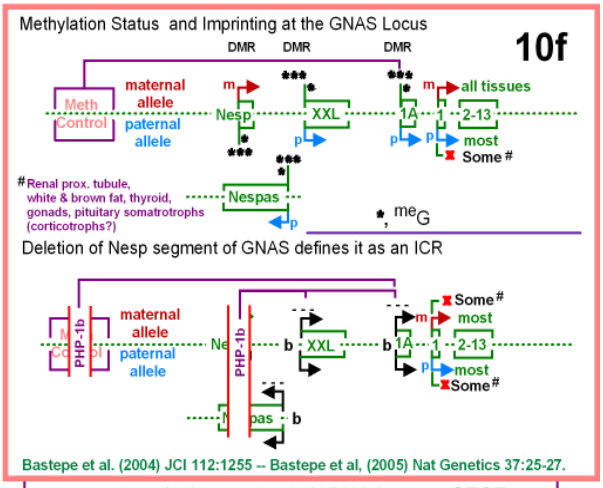
These papers have become the rationale for our studies in the specific aim that proposes to test the hypothesis that environmental exposures may leave footprints in the DNA methylation status of the bone marrow stem cell compartment that may serve as a biomarker of exposure, in a manner akin to hemoglobin A1c glycation is a biomarker of glycemic control in diabetes. The laboratory is thus moving into a direction that is committed to and totally in line with the stated mission of the institute to use environmental health sciences to inform about health and disease.

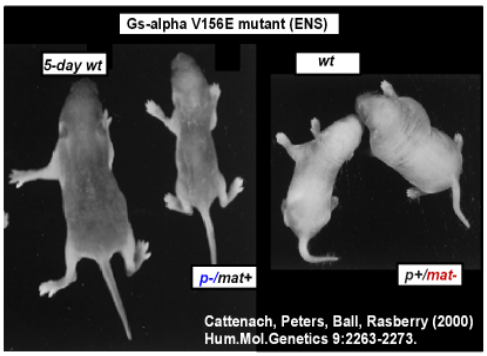
Imprinting. Total loss of Gs-*alpha* is embryonic lethal. Loss of one Gs-*alpha* allele results in phenotypes that vary with the parent of origin of the mutation. In the case of Gs-*alpha* +/- mice, they are lean, hyperactive and have increased urinary catecholamine metabolites if the mutation originates in the father, while they are hypoactive, obese, and have decreased urinary catecholamine metabolites if the mutation is transmitted by the mother (*Yu et al., 2000; Cattanach et al., 2000; Fig.10.a-b-c*). In humans, a maternal mutation results in pseudo-hypo-parathyroidism 1a, while a paternal mutation results in PPHP, pseudo-pseudo-hypo-parathyroidism. The explanation for these consequences of losing a maternal or paternal Gs-*alpha* allele lies in the fact that Gs-*alpha* is paternally imprinted (silenced) in five tissues: proximal tubule of the kidney, fat (white in humans and white and brown in mice), the anterior pituitary, the thyroid and granulosa cells of the ovarian follicle. As a consequence while we are diploid for Gs-*alpha* in most of our body, in these five selected tissues we are maternal haploids. Thus, the loss of the paternal allele due to a mutation, will make all tissues of the body haploid and be silent in the five tissues where the paternal allele is not expressed. The phenotype of a paternal mutation will therefore be whatever happens when from being haploid in just five tissues we become total haploids. Loss of the maternal allele, however, will cause the five issues to become null for Gs-*alpha*, while the rest of the body becomes identically haploid as is the case in loss of the paternal allele. The phenotype of a maternal mutation will be that seen in the loss of the paternal allele patients, plus whatever happens when these five tissues lack Gs-*alpha*. The general haploidy in Gs-*alpha* results in a syndrome called Albright's Hereditary Osteodystrophy, AHO (bradydactily, some neurologic abnormalities, low IQ, ossification problems, but normal response of the kidney to parathyroid hormone PTH, including excretion of cAMP). General haploidy plus loss of Gs-*alpha* function in kidney, fat, pituitary, thyroid and ovarian granulosa cells, results in AHO plus loss of the renal responsiveness to PTH, and multiple endocrine disturbances including short stature. This is referred to as "hormonal resistance." Of these the easiest to asses is the renal phenotype that gave the name to the disease, PHP-1a. These patients have a renal hypo-parathyroidism which is "pseudo" because circulating PTH levels are normal or slightly elevated.

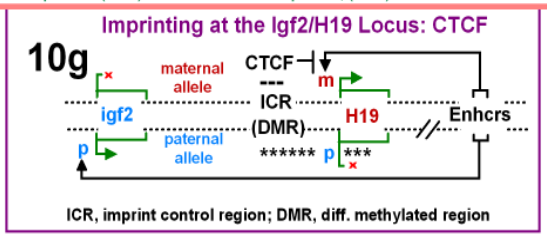
The GNAS/Gnas locus in man/mouse (Hs Chr 20q13/Mm Chr2 dist). The genomic intron/exon organization of the gene encoding the human *alpha* subunit of Gs was published by Kaziro's laboratory in 1988 (*Kozasa et al., 1988*) who reported the gene to have 13 exons transcribed as two splice forms that either included (long 395-codon mRNA) or excluded exon 3 (short 380-codon mRNA). In 1994, Huttner's laboratory reported the existence in rat of a second exon 1 encoding 499 aa instead of 47 and called it XL, extra-long (*Kehlenbach et al., 1994*). This was later revised to 367 aa. The human and mouse versions entered the GenBank database as coding for 551 and 422 aa, respectively. Exon 1 of XL *alpha*-s splices to exon 2 of Gs-alpha. Later, the same laboratory discovered Alex, a protein wholly encoded in the XL exon but in open reading frame (ORF) 3 instead of ORF 1 (*Klemke et al., 2001*). *Mus musculus* Alex was predicted to have 390 aa., its TAG stop codon being the last three nucleotides of the XL exon. Exons 2-13 of the Gs-*alpha* mRNA constitute the 3'-untranslated mRNA of the Alex cDNA. *Ischia et al. (1997)* reported the purification of the protein and the cloning of the cDNA of a chromogranin-like protein, Nesp55 or neuroendocrine secretory protein of 55 kDa which is

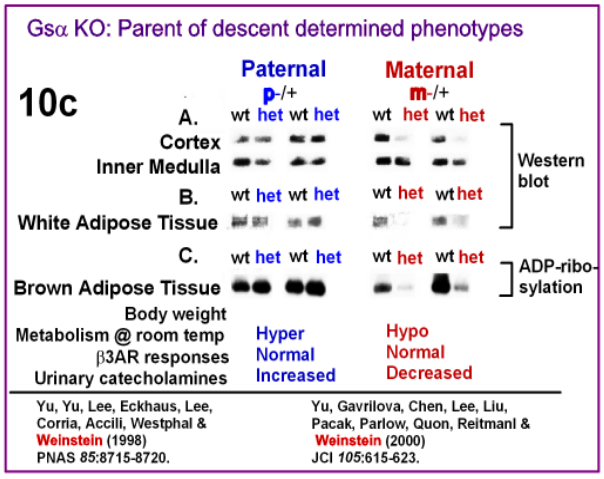

10a Albright's Hereditary Osteodystrophy
GNAS Mutants
...RCRV...
...DVGGQR...
...FLNKXD...
Met1-Val (XLas +)
Arg42-Cys
Ile103-Thr
Leu 99-Pro
Arg 98-del
Val 92-Glu
Arg165-Cys
Arg160-Cys
Leu132-Pro
Pro115-Leu
Pro115-Ser
Tyr 190 -Asp
Leu 282-Pro
Asn 264-His
Glu 259-Val
Arg258-Trp
Ser250-Arg
Arg 231-His
Arg 231-Cys
Arg 336-Trp
Arg 385-His
Ile 382-del
Codon366-369 -duplicated
Ala 366-Ser
Codon 131 Fok I polymorphism
Val159 -Glu (mouse ENS)
AHO= 50% loss of Gs function

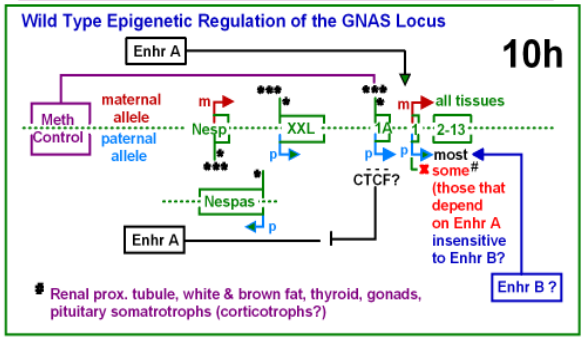

10e Intron/Exon Organization
and Imprinting at the Gnas Locus
Nesp AK039035
A1 AY682719
Nesp AF107847
Nesp AF107848
Nesp 253 aa
mXXL
1A 1 2 3 N1
4-12
785.3
726aa
1133 aa
394 aa
Nespas
Nespas AF319945
Ex1 transcripton:
m: maternal in kidney prox tubule
ant. pituitary, thyroid, gonad,fat
b: bilallic in all other tissues - ?


Methylation Status and Imprinting at the GNAS Locus
10f
DMR DMR DMR
maternal allele
paternal allele
Meth Control
Nesp
XXL
1A
2-13
all tissues
most
Some #
Renal prox. tubule, white & brown fat, thyroid, gonads, pituitary somatotrophs (corticotrophs?)
Nespas
Deletion of Nesp segment of GNAS defines it as an ICR
PHP-1b
Bastepe et al. (2004) JCI 112:1255 -- Bastepe et al, (2005) Nat Genetics 37:25-27.


10b
Gs-alpha V156E mutant (ENS)
5-day wt
wt
p-/mat+
p+/mat-
Cattenach, Peters, Ball, Rasberry (2000)
Hum.Mol.Genetics 9:2263-2273.


Imprinting at the Igf2/H19 Locus: CTCF
10g
maternal allele
paternal allele
CTCF
ICR
(DMR)
Igf2
H19
Enhcrs
ICR, imprint control region; DMR, diff. methylated region


Gsα KO: Parent of descent determined phenotypes
10c
Paternal p-/+
Maternal m-/+
wt het wt het
A. Cortex
Inner Medulla
Western blot
B. White Adipose Tissue
C. Brown Adipose Tissue
ADP-ribo-sylation
Body weight
Metabolism @ room temp
β3AR responses
Urinary catecholamines
Hyper Normal Increased
Hypo Normal Decreased
Yu, Yu, Lee, Eckhaus, Lee, Corria, Accili, Westphal & Weinstein (1998) PNAS 85:8715-8720.
Yu, Gavrilova, Chen, Lee, Liu, Pacak, Parlow, Quon, Reitman & Weinstein (2000) JCI 105:615-623.


Wild Type Epigenetic Regulation of the GNAS Locus
10h
Enhr A
maternal allele
paternal allele
Meth Control
Nesp
XXL
1A
2-13
all tissues
most
some#
(those that depend on Enhr A insensitive to Enhr B?
CTCF?
Nespas
Enhr A
Enhr B ?
Renal prox. tubule, white & brown fat, thyroid, gonads, pituitary somatotrophs (corticotrophs?)

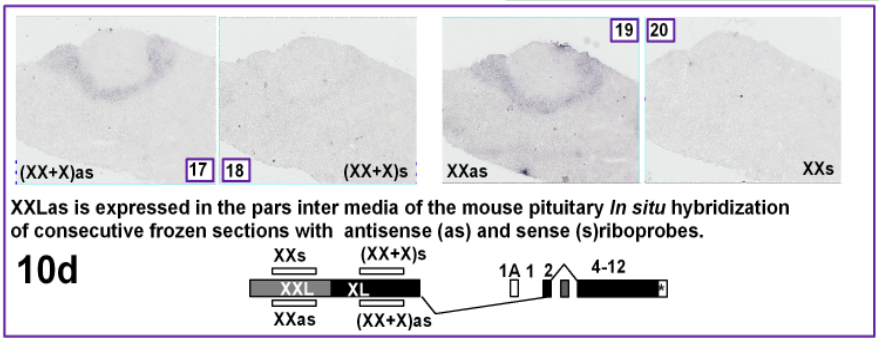

(XX+X)as
17
18
(XX+X)s
XXas
19
20
XXs
XXLas is expressed in the pars inter media of the mouse pituitary In situ hybridization of consecutive frozen sections with antisense (as) and sense (s)riboprobes.
10d
XXs
(XX+X)s
XXL
XL
XXas
(XX+X)as
1A 1 2
4-12

Fig. 10 (previous page): Imprinting at the GNAS Locus. (**a**) Summary of loss of function mutation in human Gs-alpha. (**b**) and (**c**) Differing phenotypes of mice inheriting a loss of function mutation from the father or the mother. (**d**) *In situ* hybridization showing presence of extended XXL*alpha*-s mRNA in the mouse *pars intermedia*. (**e**) Summary of transcription units and intron-exon organization of the mouse Gnas locus. (**f**) Paternal and maternal expression of GNAS transcription units (**f** top) and effect of PHP-1b causing deletions on allelic expression pattern (**f**, bottom). (**g**) Model of imprint control at the IGF2/H19 locus with insulator role for CTCF. (**h**) Summary of hypothesis on the mechanism of imprinting at the GNAS/Gnas locus if it were mediated by a CTCF insulator-type mechanism.

---

encoded in a 254-codon ORF. The same group reported in 2000 that the 3'-untrans-lated Nesp mRNA is derived from exons 2-13 of the gene coding for Gs-*alpha* (*Weiss et al., 2000*). Exon 1 of Nesp is therefore the third exon 1 of the locus. There are three splice variants of Nesp55 (GenBank numbers AF107848, with exons 2-13 from the Gs-*alpha* gene, AF107547, and AK039035). In 1998, Weinstein and colleagues reported the inactivation of the mouse Gnas gene (the Gs-alpha gene) and its imprinting (*vide supra, Yu et al., 1998; 2000*).

In 2000, Cattanach, Peters and colleagues at the MRC Mammalian Genetics Unit in Harwell, Oxfordshire, UK, isolated from a pool of ethyl-nitroso-urea (ENU)-mutagenized mice a mutant phenotype identical to that of mice with uniparental disomy of distal chromosome 2, described by *Cattanach and Kirk in 1985* to be an imprinted locus, prior to knowing the underlying coding gene or genes, and found it to be a Val to Glu mutation in Gs-*alpha* at position 156 (V156E). Thus, Weinstein's group at the NIH and the team at the MRC Mammalian Genetics Unit in the UK converged in concluding that the Gs-*alpha* locus is imprinted. Between 1999 and 2003 it became apparent that Nesp (*Weiss et al., 2000*) and XL*alpha*-s (*Peters et al, 1999; Hayward, et al., 1999*) were, respectively, paternally and maternally imprinted (silenced) in all tissues and that there also existed two maternally imprinted untranslated RNA transcripts termed 1A (or A/B), with a transcription initiation site ~2 kb upstream of Gs-*alpha*'s exon 1, transcribed in the sense direction and spliced to exon 2-13 of Gs-*alpha*, and Nespas, for Nesp antisense, with a transcription initiation site ca. 7kb upstream of the exon XL translation start site, having 5 exons (*Wroe et al., 2000; Williamson, et al., 2002; Fig. 10e*) of which the last exon is located 20.1 kb 5' of exon 1 of Nesp. The imprinting pattern and the number of transcripts found in man are recapitulated in the mouse. The mouse Gnas gene differs from the human GNAS gene in that the mouse has 12 exons instead of 13, due to fusion of exons 9 and 10 into a single exon (*Abramowitz et al, 2004*).

In 2003, we analyzed the mouse and human GNAS loci with the idea of cloning and expressing XL*alpha*-s. By inspection of the genomic sequences (GenBank) we came upon the realization that the XL and Alex ORFs were not "closed" at their 5' ends by in-frame Stop codons, but remained open for hundreds of nucleotides. By RT-PCR analysis of mouse RNA, we determined that the XL-*alpha*-s is instead an XXL-*alpha*-s (extra-extra-long *alpha*-s) with an additional 422 codons) and its companion to be AlexX (Alex-extralong) with an additional 360 codons, both in the 5' direction. Moreover, the XXL and AlexX ORFs are conserved for humans, mice and rats (*Abramowitz et al., 2004*). *In situ* hybridization confirmed the existence of the XXL/AlexX mRNA (*Fig. 10d*). We are currently investigating if XXL-*alpha*-s is competent in activating adenylyl cyclase and whether it is a cholera-toxin substrate, which requires association with G-*beta.gamma*. The "new" transcription initiation site of XXL is preceded by a TATA box and is separated from the transcription initiation site of Nespas by a 2kb CpG island that is methylated in the maternal allele (*Fig. 10f, 10h*).

Analysis of the methylation status of the five GNAS (human) and Gnas (mouse) promoters has

shown that, while the Gs-*alpha* promoter (promoter 1) is unmethylated in both alleles, the XXL, Nespas and 1A promoters are maternal methylated and paternal unmethylated, while the Nesp promoter is methylated in the paternal and unmethylated in the maternal allele. Thus, for Nesp, Nespas, XXL and 1A, promoter methylation correlates with imprinting (silencing) while absence of methylation correlates with expression (*Fig. 10f*).

The epigenetic control of promoter 1 expression differs from those of Nesp, Nesas, XXL and 1A, in that it is not methylated and its activity is "dictated" by the methylation status of 1A (*Fig. 10f bottom*). This was revealed by studies on the molecular basis of the human PHP-1b syndrome, largely by the group of Juppner at MGH.

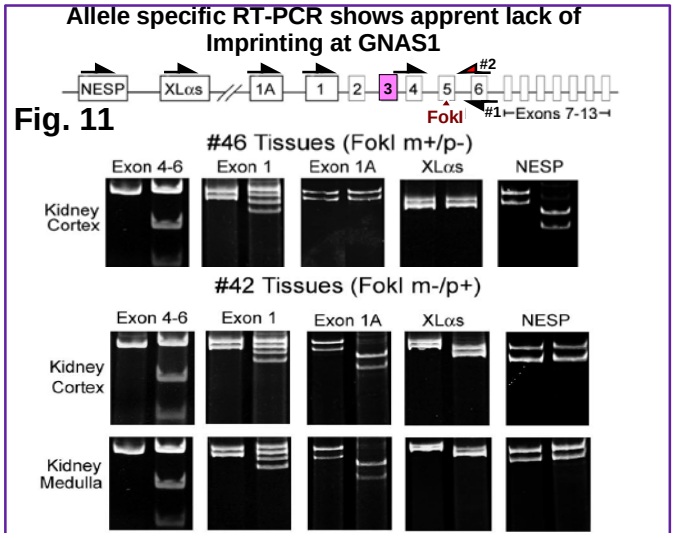
Table 2: Phenotype, Genotype and Epigenotype of Heterozygous Human Gs-*alpha* Disruptions

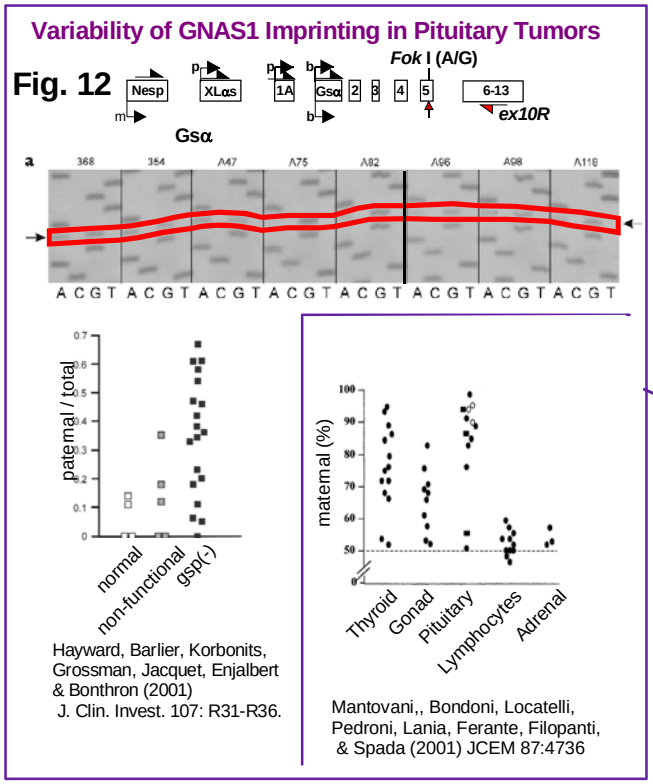
	Phenotype			
Syndrome	HR Hormone Resistance	AHO Albright's Hereditary Osteodystrophy	Gs-alpha Genotype	Tissue Specific Paternal Imprinting
PHP-1a	HR	AHO	m-/p+	yes
PPHP	--	AHO	m+/p-	yes
PHP-1b	HR	--	+/+[1,2]	pat & mat

1, A 3kb deletion at -150 kb; *Bastepe et al., 2001; 2004; Linglart et al., 2005*; 2, Spontaneous deletion of the Exon 2 of Nesp (with Nesp DMR) including exons 3 and 4 of Nespas; *Bastepe et al., 2005*.

Table 2 summarizes the human syndromes associated with loss of function mutations in Gs-*alpha* and their relation to genotype and epigenotype of the affected individuals. While PHP-1a and PPHP are due to loss of function mutations in the Gs-*alpha* coding sequence, analysis of several kindreds with PHP-1b showed it to map to either a locus ~150 kb upstream of Nesp or to Nesp. Direct sequencing showed one PHP-1b to arise from a micro-deletion of ca.3kb at the STX16 (syntaxin 16) gene, while the other is due to a deletion of the differentially methylated region of Nesp, that removes exons 1 and 2 of Nesp and exons 3 and 4 of Nespas (*Fig. 10f bottom*). The deletion at STX16 affected methylation at 1A (lost on maternal allele) causing the maternal allele of GNAS exon 1, to be silenced in the same way and the same tissues as the paternal (1A unmethylated) allele and the tissues to become null for Gs-*alpha*. Deletion of the Nesp DMR (differentially methylated region) had the same effect as the deletion at STX16, but in addition also caused loss of methylation at the XXL DMR, i.e., the locus has lost all differentially methylated domains. Associated with loss of methylation at exon 1A, was therefore a double imprinting (silencing) of GNAS exon 1 so that previously mono allelically imprinted tissues became null for Gs-*alpha*, as is the case with the deletion at STX16, hence the PHP-1b syndrome. While this deletion of Nesp led to imprinting at exon 1, the associated loss of methylation at what had been DMRs caused all non-exon 1 transcription units to lose their imprinted quality and to be expressed from both alleles (*Fig. 10f bottom*). The region encompassing exons 1 and 2 of Nesp is therefore the imprint control region or ICR of the GNAS locus. While in humans there is an additional

methylation control region at STX16, it is not clear at this time whether this is also true for mice.





Variability in the penetrance of the GNAS1 imprint. The maternal imprint of GNAS1 was analyzed for its penetrance by allele specific PCR based on a human single nucleotide polymorphism (SNP) that creates/obliterates a *Fok*I restriction site in exon 5 and found to be surprisingly variable (*Zheng et al., 2001; Hayward et al., 2001; Mantovani et al., 2001*).

The study by *Zheng et al. (2001)* analyzed GNAS1 transcripts in tissues from human fetuses and, instead of finding GNAS1 transcripts to be of maternal origin in renal proximal tubules, found them to be of biallelic origin (*Fig. 11*). Nesp, Nespas, XL, and 1A, on the other hand were found to be imprinted as expected from studies with mice.

The study of *Hayward et al. (2001)* examined pituitary tumors from acromegalics, and found GNAS1 expression to vary between 100% maternal to 30%:70% maternal:paternal (*Fig. 12*). *Mantovani et al. (2001)*, analyzed also human biopsies and found GNAS1 to be variably imprinted in the pituitary, thyroid and granulosa cells from ovarian Graafin follicles, but with a clear bias towards paternal silencing, without finding specimens with a paternal bias (*Fig. 12 inset*).

These results are surprising and puzzling, especially the lack of GNAS1 imprinting seen by Zheng et al. The biallelic expression in kidney proximal tubules detected by PCR analysis stands in contrast with the maternal imprinting that had been deduced on the basis of the clinical syndrome (PHP-1a). Likewise there is no good explanation as to why some of the pituitaries of acromegalics in the study of Hayward et al. should show partial paternal preference while this was not observed in specimens analyzed by Mantovani and collaborators.

Future Experiments in Epigenetic Control of Gene Expression

There are four areas relating to epigenetic regulation of gene expression we wish to investigate:

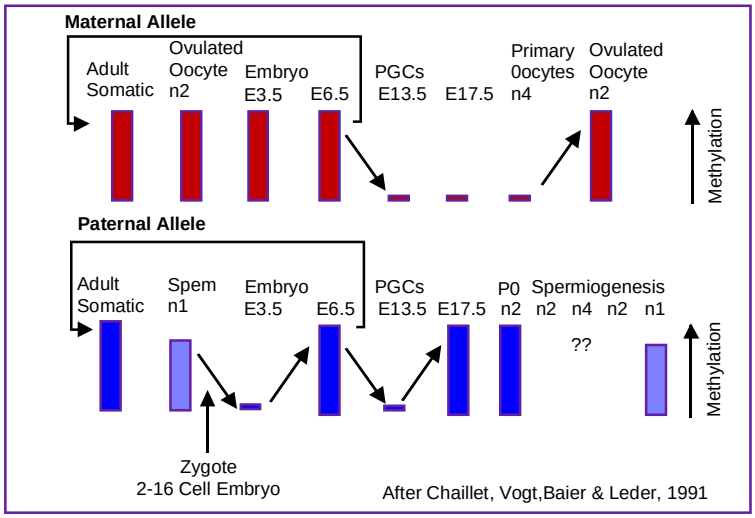
A. *<u>Imprinting of Gnas1</u>*

- Is the imprinting of Gnas1 as variable as that of GNAS1?
- Are kidney, pituitary, thyroid, fat and female granulosa cells indeed the only imprinted tissues?
- Is variability coordinated for all tissues?
- With what does variability correlate?

***B**. <u>When are the methylation marks that characterize and differentiate the parental GNAS alleles laid down?</u>*

The answer to this question is central to the imprinting phenomenon as a whole and has been addressed by several groups during the last 20 years or so. Yet, the final answer is not in. One of the original studies on allelic imprinting was done by Leder's group at Harvard analyzing a reporter transgene, as transgenes had been found to recapitulate the imprinting process, if they were large enough to contain all the necessary regulatory elements (*Swaine et al., 1987, Sapienza et al, 1987, Reik et al., 1987*). Working with mice, Leder's group showed that methylation marks are reset during gametogenesis and embryogenesis with a timing pattern that differs markedly for the male and female genomes. Moreover, it was also shown that marks are erased and laid down more than once, especially in the case of the male allele. In the 1991 paper, Chaillet et al, showed that the male mark was immature in the mature sperm, fully erased in the 3.5 day pre-implantation embryo, and established in mature from in the 6.5 day implanted embryo. After this, somatic cells retain these marks for life and are recognized in the adult as the imprinted paternal allele, while in primordial germ cells the imprint of the paternal allele is erased (complete by E13.5), remethylated (complete by E17.5) and not touched until gametogenesis. In the case of the female gamete, the methylation mark is mature at the time of fertilization, and remains unchanged in the zygote and throughout embryonic life giving adults whose somatic cells have the maternal allele marked with the prefertilization imprint. In contrast to the paternal imprint that was erased in the pre-implantation embryo, the imprint of the maternal allele is reset only in the primordial germ cells, in which the maternal imprint is lost and re-established only during resumption of meiosis - unfertilized eggs (ovulated oocytes) have a female, soon to be maternal imprint. These paternal/maternal methylation fate differences are summarized in *Fig. 13*.

Fig 13. Parental allele specific changes in DNA methylation during embryo- and gametogenesis as inferred from the studies of Chaillet et al.(1991)



The analysis of the GNAS locus DRMs has shown them to be of different hierarchies. We therefore wish to determine when the different maternal and paternal methylation marks are laid down. XXL, Nespas and 1A promoters are differentially methylated (present on the maternal allele, absent on the paternal allele), while Nesp is the only methylated promoter on the paternal allele, and, as was discovered in two independent kindreds with the PHP-1b syndrome (*Bastepe al., 2005*), this Nesp DMR is the imprint control region of the locus. This region will therefore receive special attention in our sequencing efforts.

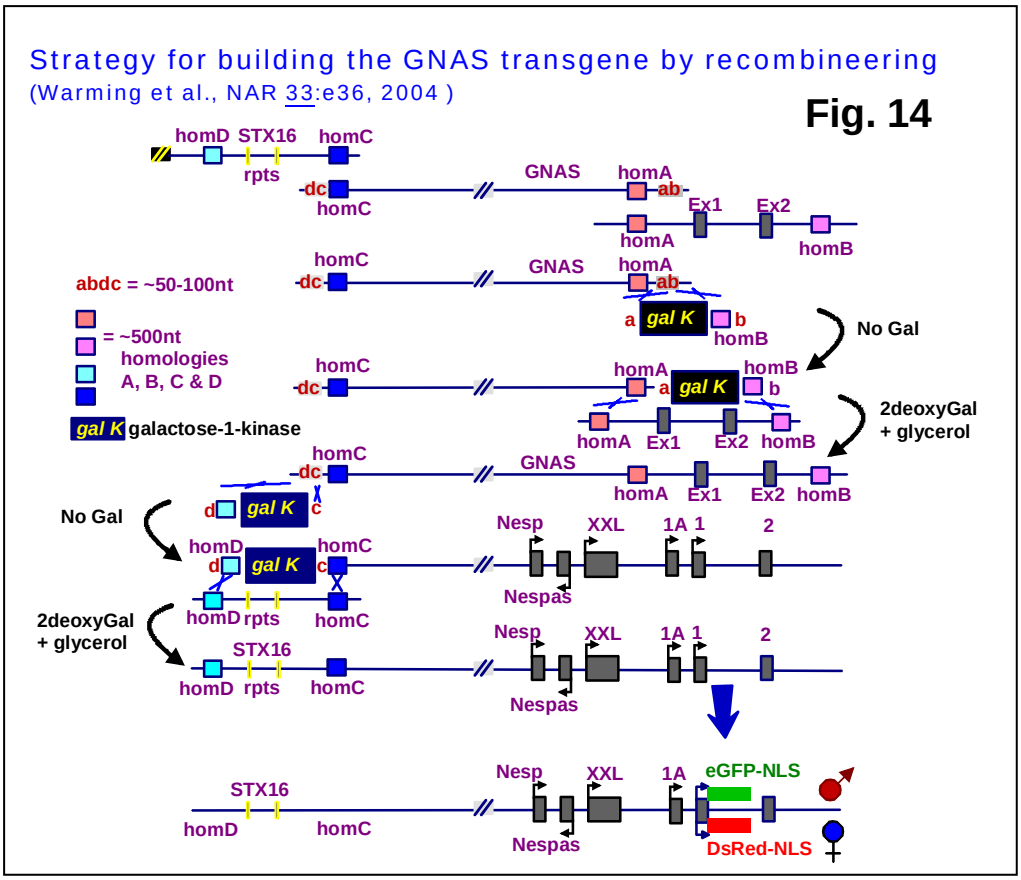
The discussion so far has dealt exclusively with methylation marks that relate to imprinting. These marks need to be differentiated from general DNA methylation unrelated to heritable imprints but nevertheless important for regulating gene expression. General changes in methylation of CpG dinucleotides at fertilization was followed by Meyer et al. (2000) using a 5meC specific antibody. Their studies showed that concomitant with the reconstitution process of the male genome (loss protamines and restitution of histone-based nucleosomes), the male genome is massively demethylated by the oocyte's enzymes, while at the same time methylation of the female genome remains unaffected. It is said that this demethylation spares imprints. Prior to fertilization-demethylation, the male genome also undergoes massive demethylation during spermiogenesis, between the pachytene N4 and the spermatid N1 stage (*Geyer et al., 2004*). Also here it is not clear whether this demethylation affects imprints or not. Using the in-agarose bead deamination procedure (see below) followed by PCR amplification and sequencing, we are in the process of comparing general demethylation as seen by studying methylation of the Pgk2 gene to imprint demethylation (if it occurs) as seen by studying the various Gnas DMR's

Methods in Epigenetic Studies.

Comments will be provided only on those techniques that are new in our laboratory.

Construction of the fluorescent GNAS transgene. We will use a human GNAS transgenes in which green (eGFP) or red (DsRedExpress) fluorescent proteins are fused onto Exon 1. The expectation is that by two color fluorescence microscopy we will be able to evaluate the transcriptional activity of the GNAS promoter 1 at the single cell level and also to easily detect imprinting in many tissues and during development.

Given that GNAS has two methylation control regions, we are constructing a 250 kb transgene assembled from three BACs using recombineering (recombinant engineering) in *E.coli* SM102 cells developed by Warming et al. (2005). SM102 cells carry the *lambda* Red recombination genes in a stably integrated *lambda* prophage under the control of *cI857* inducible *p*L promoter. The *lamda*Red recombination locus encodes *exo*, *bet* and *gam*, where *exo* is a 5'→3 'exonuclease; *bet* is a pairing protein [binds to 3'-overhang and promotes annealing and homologous recombination with homologous target DNA sequence]; and *gam* is an inhibitor of *E. coli* RecBCD exonuclease, protecting the overhangs created by *exo*. The strain also carries the complete galactose operon from which *galK* (galactose-1-kinase gene) has been precisely deleted. *Fig. 14* depicts the strategy we are using. Insertion of *galK* into the target locus (BAC) is done by electroporating a PCR construct of *galK* with the homologies fused to its 5' and 3' ends into SM101 cells carrying the BAC one wants to engineer. DNA fragments intended for insertion into the BAC are then introduced into the cells with the galK-BAC in the form of linearized restriction fragments by electroporation. At the time of this writing, we have assembled the genomic DNA from the three BACs into one and are in the process of inserting the fluorescent proteins to obtain the final transgene.



High sensitivity bisulfite sequencing. At this time the only reliable method to determine location of methyl groups is based on chemical deamination of C to T, under conditions that don't affect 5-methyl C, followed by dideoxy sequencing with standard procedures. Chemical deamination is produced by a 4-8 hr treatment of the DNA at pH 5.0 with 5 M NaHS03 in the presence of hydroquinone, followed by alkalinization. The DNA is then recovered by "standard" methods. This works, but the yield of the deaminated DNA is low, requiring ca. 100 ng of starting material. We wish to analyze a few cells 20-40 at 7 pg of DNA per cell. The yield is dramatically improved if the DNA is physically trapped in an agarose bead (*Olek et al., 1996*). This has allowed us to analyze the H19 DMR in 50 pachytene spermatocytes (ca. 700 pg or 200 haploid genomes). The pachytene spermatocytes were manually picked from seminiferous tubule squashes and deaminated DNA was amplified by nested PCR.

Enzymatic deamination. The activation induced cytidine deaminase (AID) has a 10-fold preference for cytidine over 5-methylcytidine (*Larijani et al., 2005*). Based on sequence alignments and a model of AID based on the crystal structure of a yeast deaminase (*Xie et al., 2004*) and sequence homology to the RNA ediciting APOBEC and the help of Tom Darden (Laboratory of Structural Biology, NIEHS), we have identified amino acid residues not likely to be involved in catalysis that could be changed to more bulky amino acids so as to fully occupy the cytidine binding pocket and impede binding of methylcytidine. The aim is to increase the selectivity for cytidine over deoxycytidine to 100,000 or more and to use enzymatic deamination followed by direct PCR increasing the sensitivity to enable us to asses location of methylcytidines in DNA samples.

C. *Does CTCF play a role in the imprinting process that regulates Gs-alpha expression?*

One well studied imprinted locus is the Igf2/H19 locus, in which mesodermal enhancers down stream of H19 activate the Igf2 promoter located upstream of H19 promoter if a small differentially methylated imprint control region between Igf2 and H19 is methylated. These enhancers cannot act on the Igf2 promoter on the un-methylated allele of the DMR, due to the binding of the insulator protein CTCF (summarized in *Figs. 10g and 15)*. The methylated ICR is on the paternal allele. *Figs. 10f and10h* summarize the differential regulation of Gs-alpha (GNAS1) expression from maternal and paternal alleles, and how the regulatory circuity should look like if the Igf2/H19 mode of regulation by CTCF is applied to Gs-alpha expression. The model is complex and at this time could only partially explain the tissue specific imprinting of Gs-alpha. To 'save" the model we have postulated a down-stream enhancer B (*Fig. 10h*). We expect to be able to locate enhancers through deletion analysis of the GNAS1 transgene. At present the GNAS sequence down stream of exon 1 extends 7.3 kb to include exon2 and 3. If new results would indicate this to be desirable we could extend the genomic sequence further. Transgenes are of course analyzed by making transgenic mice, which will be done either in house or under contract at the University of Rochester, NY.

As for the initial experiments testing for involvement of CTCF we will begin by searching for CTCF binding sites on the BAC with the GNAS locus. This has to be done experimentally, because by informatics there is no good consensus CTCF biding site. This is due to the fact that CTCF has 11 Zn fingers and uses different combinations of these DNA binding motifs at different sites of the genome (*Filippova et al., 1996*). In preparation for this study we have obtained the CTCF cDNA from Gary Felsenfeld, transferred it to the transcription competent pAGA plasmid and obtained by *in vitro* translation CTCF labeled with high specific activity [^{35}S]Met and [^{35}S]Cys. The plan is to make a bacterial library of sheared GNAS BAC DNA in pCR-blunt, and by "colony hybridization" search for plasmids with inserts positive for CTCF binding. These clones will then be amplified and sequenced

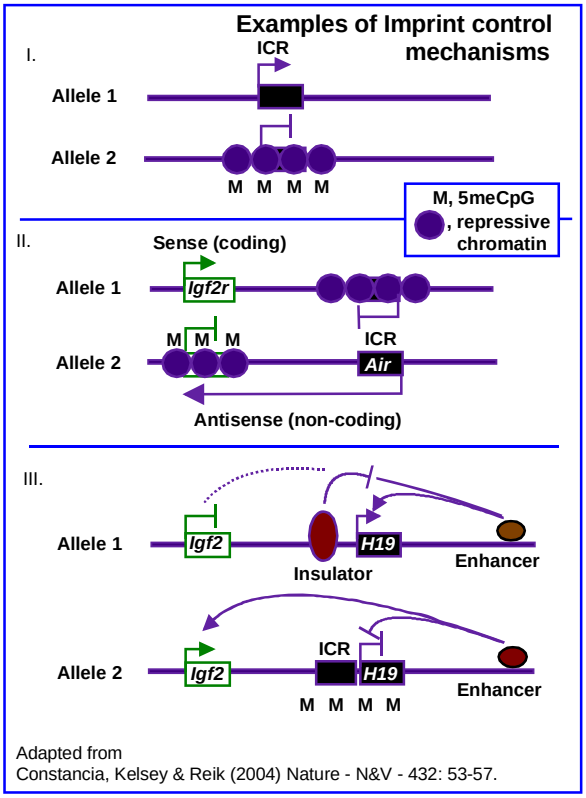


Fig. 15: Summary of mechanisms under consid-ration to explain imprinting at the GNAS locus.

In case we do not find a CTCF binding site in the GNAS locus, we will consider other models of imprinting. *Fig. 15* is a summary, adapted from *Constancia et al.'s 2000* News & Views commentary in Nature, depicting three mechanisms that may account for imprinting. They have in common that the transcription initiation site of the imprinted allele is located in a "repressive" region of the chromosome, which may but need not be fully heterochromatic with Polycomb proteins loaded onto the methylated DNA. All models, eventually call for an analysis of the chromatin structure that underlies silencing. Model I is a general model with no further assumptions, Model II is based on the imprinting of the Igf2 receptor, in which the heterchromatic state is induced through the intermediary synthesis of an untranslated antisense RNA termed Air, transcribed from an unmethylated promoter, which presumably catalyzes the transformation

of the chromosome to a "repressive" state by attracting histone modifying and DNA binding proteins.

D.. *Can DNA methylation be used as a biomarker of environmental exposures?*

A fourth and last aim of our future research in epigenetics is an expansion of the first in which we investigate the imprinting variability of Gs-*alpha*, and will test whether changes in somatic imprinting are sensitive to environmental exposures. We will explore specifically whether exposures leave a footprint in the epigenome (methylation pattern) of self-renewing lympho-hematopoietic compartment of bone marrow. If so, the DNA of the circulating lympho-hematopoietic compartment (blood cells) could serve as an indicator of adverse environmental exposures: pathogens, allergic and inflammatory insults, heavy metal intoxications, endocrine disruptors and polychlorinated compounds will be insults looked at in an initial pass.

*An attempt to sequence the epigenome: global analysis of DNA methylation. - W*e are attempting to develop methods that would allow us to examine DNA methylation at a global, genomic scale. One approach we are testing is to collect 20 nt sequences that flank methylated CpG dinucleo-tides using linkers with the *Mme*I TCCRAC recognition sequence (R=A/G). *Mme*I is a restriction enzyme that cuts 18nt 3' of the recognition sequence leaving a 2nt cohesive end [TCCRAC$(N)_{18}$'NN_]. The idea is to shear genomic DNA, digest with *Hpa*II (C'CG_G), then ligate to the 5' ends of the *Hpa*II fragments a linker formed (5' to 3') of PCR primer tag 1, the *Mme*I recognition sequence followed by the cohesive ends of the methylation sensitive *Msp*I recognit-ion sequence. These cohesive ends are 2-nt 5' overhangs to which the *Hpa*II digests can be ligated. *Msp*I and *Hpa*II have the same C'CG_G recognition sequence with identical cohesive ends, but differ in that *Hpa*I is insensitive to methylation of the site's CpG, while *Msp*I will not cut either fully or hemimethylated sites. The tagged DNA fragments are then digested with *Msp*I. If the genomic fragment had been cut at an unmethylated CpG site, it will be released and lost. If the ligated *Hpa*II fragment was methylated, it will not come off the linker, leaving an *Mme*I 2 nt cohesive end preceded by18nt of meCpG flanking sequence. Ligating a second linker to the *Mme*I site with PCR tag 2 should leave fragments of 60-70 nt (including primers) that can be amplified and analyzed.

We will consider two types of analysis. One, is to amplify the fragments by PCR and hybridize the products to a Perlegen (or similar) genomic chip. If this proves unfeasible due to price restrictions, we will consider switching to a SAGE (sequential analysis of gene expression) using shorter linkers and identifying the products by concatenation, cloning and sequencing. An alternative, less appealing method, would be to analyze electrophoretic patterns of arbitrarily primed PCR products for which the DNA whose methylation changes one is interested in evaluating is the template. This is the approach used by Anway and colleagues to detect transgenerational changes in DNA methylation caused by vinclozoline (*Anway et al., 2005*).

Since the approach based on the *Hpa*II/*Msp*I restriction enzyme pair to differentiate methylat-ed from non-methylated CpG's only interrogates CpG's flanked by C and G, we would be inter-rogating only about 8% of the total number of CpG's. With the tools at hand we consider this as being pretty good. Analysis of genomic sequences for allelic methylation differences (DMRs) using the *Hpa*II/*Msp*I screen has been used successfully to identify DMR's in specific genes, including the Igf2/H19, the Igf2r and the complex GNAS and Gnas loci, and should be adequate for a genomic screen since toxic effects are likely to be at the level of regions, instead of defined CpGs.

Literature Cited

Abramowitz, J., Grenet, D., Birnbaumer, M., Torres, H.N. and Birnbaumer, L. (2004) "XLαs, the Extralong Form of the α Subunit of the Gs G Protein is Significantly Longer than Suspected, and So is Its Companion Alex." Proc. Natl. Acad. Sci. U.S.A. *100*, 8366-8371.

Anway. M.D., Cupp, A.S., Uzumcu, M. and Skinner M.K. (2005) "Epigenetic transgenerational actions of endocrine disruptors and male fertility." Science *308*, 1466-1469.

Bastepe, M., Frohlich, L.F., Hendy, G.N., Indridason, O.S., Josse, R.G., Koshiyama, H., Korkko, J., Nakamoto, J.M., Rosenbloom, A.L., Slyper, A.H., Sugimoto, T., Tsatsoulis, A., Crawford, J.D and Juppner, H. (2003) "Autosomal dominant pseudohypoparathyroidism type Ib is associated with a heterozygous microdeletion that likely disrupts a putative imprinting control element of GNAS." J. Clin. Invest. *112*, 1255-1263.

Bastepe, M., Frohlich, L.F., Linglart, A., Abu-Zahra, H.S., Tojo, K., Ward, L.M. and Juppner, H. (2005) "Deletion of the NESP55 differentially methylated region causes loss of maternal GNAS imprints and pseudohypoparathyroidism type Ib." Nat. Genet. *37*, 25-27.

Bastepe, M., Pincus, J.E., Sugimoto, T., Tojo, K., Kanatani, M., Azuma, Y., Kruse, K., Rosenbloom, A.L., Koshiyama, H. and Juppner, H. (2001) "Positional dissociation between the genetic mutation responsible for pseudohypoparathyroidism type Ib and the associated methylation defect at exon A/B: evidence for a long-range regulatory element within the imprinted GNAS1 locus." Hum. Mol. Genet. *10*, 1231-1241.

Cattanach, B. M. and Kirk, M. (1985) "Differential activity of maternally and paternally derived chromosome regions in mice. Nature *315*, 496-498.

Chaillet, J. R., Vogt, T. F., Beier, D. R. & Leder, P. (1991) "Parental specific methylation of an imprinted transgene is established during gametogenesis and progressively changes during gametogenesis." Cell *66*, 77-83.

Filippova, G.N., Fagerlie, S., Klenova, E.M., Myers, C., Dehner, Y., Goodwin, G., Neiman, P.E., Collins, S.J. and Lobanenkov, V.V. (1996) "An exceptionally conserved transcriptional repressor, CTCF, employs different combinations of zinc fingers to bind diverged promoter sequences of avian and mammalian c-*myc* oncogenes." Mol. Cell. Biol. *16*, 2802-2813.

Geyer, C.B., Kiefer, C.M., Yang, T.P. and, McCarrey, J.R. (2004) "Ontogeny of a demethylation domain and its relationship to activation of tissue-specific transcription." Biol. Reprod. *71*, 837-844.

Hayward, B.E., Moran, V., Strain, L and Bonthron, D.T. (1998) "Bidirectional imprinting of a single gene: GNAS1 encodes maternally, paternally, and biallelically derived proteins." Proc. Natl. Acad. Sci. USA *95*, 15475-15480.

Kozasa, T., Itoh, H., Tsukamoto, T. and Kaziro, Y. (1988) "Isolation and Characterization of the Human G_s α Gene." Proc. Natl. Acad. Sci. USA 85,2081-2085.

Larijani, M., Frieder, D., Sonbuchner, T.M., Bransteitter, R., Goodman, M.F, Bouhassira, E.E., Scharff, M.D, and Martin, A. (2005) "Methylation protects cytidines from AID-mediated deamination." Mol. Immunol. *42*, 599-604.

Linglart, A., Gensure, R.C., Olney, R.C., Juppner, H. and Bastepe, M. (2005) "A novel STX16 deletion in autosomal dominant pseudohypoparathyroidism type Ib redefines the boundaries of a cis-acting imprinting control element of GNAS." Am. J. Hum. Genet. *76*, 804-814.

Mayer, W., Niveleau, A., Walter, J., Fundele, R., Haaf, T. (2000) "Demethylation of the zygotic paternal genome." Nature *403*, 501-502.

Olek, A., Oswald, J. and Walter, J. (1996) "A modified and improved method for bisulphite based cytosine methylation analysis." Nucleic Acids Res. *24*, 5064-5066.

Peters, J., Wroe, S.F., Wells, C.A., Miller, H.J., Bodle, D., Beechey, C.V., Williamson, C.M. and Kelsey G. (1999) "A cluster of oppositely imprinted transcripts at the Gnas locus in the distal imprinting region of mouse chromosome 2." Proc. Natl. Acad. Sci. USA *96*, 3830-3835.

Reik, W., Collick, A., Norris, M.L., Barton, S.C. and Surani, M.A. (1987) "Genomic imprinting determines methylation of parental alleles in transgenic mice." Nature *328*, 248-251.

Warming, S., Costantino, N., Court, D.L., Jenkins, N.A. and Copeland, N.G. (2005) "Simple and highly efficient BAC recombineering using galK selection." Nuc. Acids Res *33*:e36.

Weaver, I.C., Cervoni, N., Champagne, F.A., D'Alessio, .A.C, Sharma, S., Seckl, J.R., Dymov S., Szyf, M. and Meaney, M.J. (2004) "Epigenetic programming by maternal behavior." Nat. Neurosci. *7*, 847-854.

Weiss, U., Ischia, R., Eder, S., Lovisetti-Scamihorn, P., Bauer, R. and Fischer-Colbrie, R. (2000) "Neuroendocrine secretory protein 55 (NESP55): alternative splicing onto transcripts of the GNAS gene and posttranslational processing of a maternally expressed protein." Neuroendocrinology *71*, 177-86.

Williamson, C.M., Skinner, J.A., Kelsey, G. and Peters, J. (2002) "Alternative non-coding splice variants of Nespas, an imprinted gene antisense to Nesp in the Gnas imprinting cluster." Mamm. Genome *13*, 74-79.

Wroe, S.F., Kelsey, G., Skinner, J.A., Bodle, D., Ball, S.T., Beechey, C.V., Peters, J. and Williamson, C.M. (2000) "An imprinted transcript, antisense to Nesp, adds complexity to the cluster of imprinted genes at the mouse Gnas locus." Proc. Natl. Acad. Sci. USA *97*, 3342-3346.

Xie, K., Sowden, M.P., Dance, G.S., Torelli, A.T., Smith, H.C. and Wedekind, J.E. (2004) "The structure of a yeast RNA-editing deaminase provides insight into the fold and function of activation-induced deaminase and APOBEC-1." Proc. Natl. Acad. Sci. USA *101*,

8114-8119.

Yu, S., Gavrilova, O., Chen, H., Lee, R., Liu, J., Pacak, K., Parlow, A.F., Quon, M.J., Reitman, M.L. and Weinstein, L.S. (2000) "Paternal versus maternal transmission of a stimulatory G-protein alpha subunit knockout produces opposite effects on energy metabolism." J. Clin. Invest. *105*, 615-623.