5 DISCUSSION

5.1 What is the Molecular Basis of DS Phenotypes?

Scientists have struggled for many years about the mechanism of action in DS. Aneuploidy arises either from meiotic errors leading to germline aneuploidy or from early mitotic errors resulting in aneuploidy in most cells of an individual (Hall *et al.*, 2006; Hassold and Hunt, 2001). However, we know very little about why the abnormal gene dosage has such profound and often lethal effects on the development of an individual. Two principal hypotheses were formulated in an attempt to explain how trisomy 21 may cause Down syndrome. Both hypothesis are based on the assumption that if a gene is present in three copies rather than two the level of expression of that gene will be elevated by 50%.

The amplified developmental instability hypothesis, proposed by Burton L. Shapiro, states that, in general, the most important cause of the array of phenotypic features that are associated with Down syndrome does not actually involve direct contributions of specific genes on HSA21, but rather that elevated activity of sets of genes, regardless of their identity, will lead to a decrease in genetic stability or homeostasis (Pritchard and Kola, 1999; Shapiro, 1999). This hypothesis was proposed in an attempt to explain the similarities between the phenotypes of different congenital aneuploid states (e.g. trisomy 13 or 18) and the observation that all of the phenotypic traits in DS are also seen in the general population. However, these signs are generally less frequent and are never present simoultanously in the general population.

A second hypothesis, the gene dosage-effect hypothesis, states that elevated expression of specific trisomic genes directly leads to specific features of Down syndrome. This model underlies the hypothesis of “critical regions” on HSA21, chromosome segments believed to include one more dosage sensitive gene that are directly responsible for a given aspect of the DS phenotype. This hypothesis was based on data pertaining to DS patients carrying only a partial trisomy. The comparison of the chromosome anomalies and physical characteristics shared among those patients with partial trisomy 21 has led to the concept of a critical region for certain features of Down syndrome (see Introduction). The best described Down syndrome critical region (DSCR) extended

about 5.4 Mb on HSA21 (McCormick *et al.*, 1989). This region was associated with several of the major DS phenotypes, including flat facies, mental retardation, short stature, protruding tongue, etc. Although controversial, the idea of the DSCR implies that much of Down syndrome could be caused by extra copies of one or a small number of genes in this region (Delabar *et al.*, 1993; Korenberg *et al.*, 1994).

It should be noted that the molecular analysis of partially duplicated patients was published more than ten years ago. These patients should be re-investigated taking advantage of more modern technologies and DS marker maps spanning HSA21. It should be said that most of partially duplicated patients had additional chromosomal abnormalities. Nevertheless, the notion that a few genes might be of critical importance in this syndrome seems more likely because each different type of other congenital autosomal chromosomal anomaly results in a different phenotype with characteristics specific for that aneuploidy/segmental aneusomy (e.g. Cri du chat syndrome, trisomy 9 or 13). However, where multiple genes contribute in a complex way to an aneuploidy phenotype, there will be no simple dosage to phenotype correlation.

Using mouse models, recent work by Olson and coworkers (Olson *et al.*, 2004) demonstrated that the DSCR is not sufficient neither indispensable to cause specific Down syndrome phenotypes. Data from studies of cranofacial dysmorphology in the Ts1Rhr model of DS show that there is no simple relationship between individual genes and discrete aspects of a phenotype (Olson *et al.*, 2004). In such cases, an iterative deletion/duplication analysis may prove inadequate to identify the causative genes, and more complex strategies will have to be developed.

However, a recent study showed that the onset Alzheimer's disease was caused by duplication of a region of HSA21, which contains only four genes, including the amyloid plaque precursor protein, *App* gene (Zhang *et al.*, 2006). Such findings in human converge to toward similar conclusions than several studies performed on the mouse models, including ours, that support the concept that some specific phenotype will be due to a limited number of genes. The mechanisms whereby genes that are present in three copies might contribute to changes in cell function directly or indirectly to cause specific DS phenotypes is likely to represent

the full range of genetic mechanisms seen in other complex traits, with some additional aspects specific to trisomy (Figure 5-1).

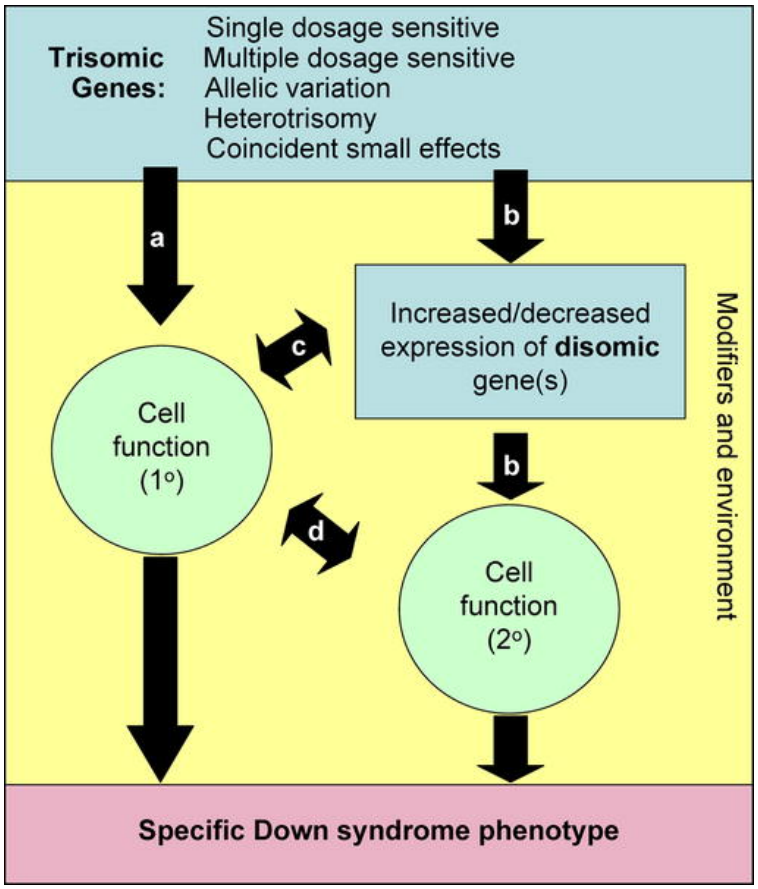

Figure 5-1: Possible Phenotypic Consequences of Gene Action in Down Syndrome. (A) A trisomic gene or genes might directly affect cellular function in a fully differentiated cell to cause a functional phenotype of DS or in an immature cell to produce a developmental phenotype. (B) Trisomic genes may alter expression of disomic genes, leading to a cellular manifestation and a DS phenotype. A trisomy-induced change in cellular function altering the relationship of that cell to surrounding cells leads to a secondary distortion of (C) disomic gene expression or (D) function in neighboring cells. Modifier genes or environment (yellow box) might interact at multiple points to initiate, ameliorate, or exacerbate phenotypes**.** This picture & legend was taken from Roper *et al.* (Roper *et al.*, 2006)

In summary the understanding of DS pathology requires to identify unambiguous associations between the overexpression of specific chr21 genes with specific phenotypic components. Three basic questions have to be addressed: 1). What is the level of expression of genes at dosage imbalance? 2) Is this level of overexpression likely to be pathogenic for all genes considering that some level of expression variation is not systematically deleterious for the organism? 3) What are the pathological effects associated with dosage sensitive genes? It is moreover

essential to remember that trisomy also affects all the non-protein-coding genes, microRNAs, regulatory regions, structural regions, etc., which lie within the overdosed region. Furthermore, for some genes, negative feedback may exist to regulate the level of transcript or protein such that the altered gene dosage may not translate into the altered levels of protein. However it is reasonable to iniatially focus the study on analyzing transcript levels of protein coding genes.

As with all other human genetic disorders the mouse is the model organism of choice for carrying out sophisticated manipulations to recapitulate (as far as possible) the human condition and to start to understand the mechanisms by which dosage abnormalities affect the human biology. The Ts65Dn mice produce phenotypes with direct parallels to those seen in DS. The commonality of phenotypes indicates that the same developmental genetic pathways are similarly disrupted by the corresponding dosage imbalance in the two species.

One may ask at this step why do we measure mRNA levels at all, as messenger RNA is only an intermediate molecule? It is essential to measure the mRNA levels, as it is the primary level of change in trisomy. Differences in gene expression are responsible for both morphological and phenotypic differences as well as indicative of cellular response to environmental stimuli and perturbations. Changes in the quantity of some gene products may also have profound qualitative consequences on the target protein (post-translational modifications and processing for instance). The transcriptome is highly dynamic and can change rapidly and dramatically in response to normal cellular events or external perturbations. So, mRNA levels are immensely informative about cell state and the activity of genes, and for most genes, changes in mRNA abundance are related to changes in protein abundance. Further, protein-based approaches are generally more difficult, less sensitive and have a lower throughput than the RNA-based ones. In terms of understanding the function of genes, knowing when, where and to what extend a gene is expressed is crucial to understand the activity and biological role of its encoded protein.

5.2 Dosage Sensitive Genes

The postulate that genes present in triplicate in the genome would be over-expressed by 50% compared with disomic genes has never been formally demonstrated in a systematic fashion in a panel of tissues from a mammal with trisomy. As a first step towards understanding the molecular consequences of trisomy, we have analyzed the expression of 136 mmu21 transcripts in nine tissues of Ts65Dn trisomic mice and euploid littermates. We found that 82% of the tested orthologs of human Chr21 were expressed in at least one tissue (Kahlem *et al.*, 2004). Many were expressed in the brain (64%) whereas relatively few genes were active in skeletal muscle (15%). Hierarchical clustering of quantitative array data in this work clearly demonstrated that the mmu21 genes are differentially expressed in the nine tested tissues, showing that expression profiling reflects tissue specificity. Previous expression maps of Chr21 orthologs in mice demonstrated that 94-98% of all mmu21 genes are expressed in adult mice as observed by RT-PCR (Reymond *et al.*, 2002) or by estimation from EST mining (Gitton *et al.*, 2002). The somewhat lower result from array analysis is explained by the fact that previous expression studies surveyed a broader range of tissues and were based on techniques that are more sensitive than arrays in detecting low levels of gene expression. RT-PCR data from the previous studies were not quantitative.

Our observations support a simple model for global changes in Chr21 transcript levels in DS. Most genes at dosage imbalance in Ts65Dn mice were up regulated by about 50%. That is, their steady state transcript levels were directly proportional to the number of copies of the corresponding structural genes across multiple tissues, consistent with previous observations reported for several individual HSA21 genes. This result indicates that there is little or no feedback to regulate steady state mRNA levels in response to gene copy number by controlling either transcription or any of the other steps determining mRNA quantity in the cell (e.g. splicing, RNA processing or degradation). Rather, template availability appears to determine transcript levels for a given gene, even though the actual level of transcript varies among tissues. Recent observations confirm that transcript levels are elevated about 1.5-fold for the majority of trisomic genes in a few tissues or cells from humans with trisomy 21 (Bahn *et al.*, 2002; FitzPatrick *et al.*, 2002; Giannone *et al.*, 2004; Gross *et al.*, 2002; Mao *et al.*, 2003) and across a broad

range of tissues that can be measured in trisomic mouse models (Amano *et al.*, 2004; Chrast *et al.*, 2000; Dauphinot *et al.*, 2005; Kahlem *et al.*, 2004; Lyle *et al.*, 2004; Saran *et al.*, 2003). These studies uses mostly pooled RNA samples and the results are recapitulated in Table 5-1.

The assumption that 50% more template will result in 50% more steady state RNA has been central to DS research for many years. And its demonstration in a variety of cells and tissues is a reasonable indicator that, for the most part, this level of overexpression will occur in all cells where that gene is expressed throughout development. For those genes whose elevated expression alters a function in fully differentiated cells, the presence of elevated expression in adults may be considered directly in determining the mechanism by which overexpression of that gene contributes to a phenotype of DS. However, overexpression of a given gene will not necessarily affect development and function in every cell type and at every developmental time point. It is likely that overexpression of some genes is detrimental only at a specific time during development, and then only in a specific cell type. Further, a trisomy-induced change in one cell population could affect neighboring cells, resulting in aberrant development as a secondary consequence of trisomy. Nonetheless, it is also a somewhat surprising observation on several levels. Transcription for many genes is tightly regulated with regard to tissue type and developmental stage. For a number of gene products (e.g., transcription factors, cell surface receptors and their ligands, signal transduction molecules, rate-limiting enzymes in metabolic pathways), the amount of protein produced is critical to cellular and higher function. It is also curious that RNA levels are not particularly subject to feedback or other type of regulation from an evolutionary context, given the prevalence of aneuploidy and the significant perturbations that result. Trisomy 21 is the most frequent live-born aneuploidy and even here, only a small fraction of DS conceptuses survive to term. Prenatal mortality is nearly universal for the more common human trisomies of Chr13 and 18 and for any other full chromosome trisomies.

Species & Strain	Tissues	Ages	Sample size	Sample type	Method	Results	References
Human amniocytes	Amniocytes primary cell cultures	16-18 weeks gestation	3 T21 + 4 controls	RNA pools	InCyte Genomics cDNA array	Global upregulation of Chr 21 genes. Most deregulated genes are not on Chr21- Global transcriptome disregulation.	FitzPatrick et al., 2002
Human placenta	Placental tissue	16-24 weeks gestation	7 T21 + 14 controls	Pooled RNA	Custom cDNA arrays	13 genes overexpressed.	Gross et al., 2002
Human neurospheres	Neuronal precursor cells derived from cortex	8-18 weeks gestation	1 T21 + 2 controls	Non pooled RNA	Differential display PCR	Down-regulation of REST-regulated genes (not HSA21) in DS.	Bahn et al., 2002
Human fetal brain and derived astrocyte cell lines	Cerebral cortex + derived astrocyte cell lines	17 20 weeks gestation	4 T21 + 4 controls	Non pooled RNA	Affymetrix chip	Up regulation of chr21 genes. Greater variability in gene expression levels on chr21 among DS than among Euploid.	Mao et al., 2003
Human T lymphocytes cell lines	Primary cell lines of T lymphocytes from blood	12 months-40 years.	2 T21 + 2 controls	Non pooled RNA	MicroMax NEN array	Up regulation of Chr21 genes. 15 nonHSA21 dysregultaed	Giannone et al., 2004
Ts1Cje mice	Whole brain	P0	6 Ts1 + 6 controls	Non pooled RNA	Affymetrix chip	Dosage-dependant up-regulation of genes in the trisomic region. No significant perturbation of the expression levels for genes on euploid regions.	Amano et al., 2004
Ts1Cje mice	Cerebellum	P0, P15, P30	3 Ts1 + 3 controls	Pooled and non pooled RNA	Affymetrix chip	Up-regulation of genes on the trisomic region. Some exception to this rule. Differential expression of euploid genes (~4% in a time dependant manner.	Dauphinot et al., 2005
Ts65Dn mice	Whole brain, spleen	P30	3 Ts + 7 controls	RNA pools	SAGE	No statistical significant results for the triplicated genes. Deregulationion of 330 genes including ribosomal proteins	Chrast et al., 2000
Ts65Dn mice	Cerebellum	3-4 months	6 Ts1 + 6 controls	Pooled and non pooled RNA	Affymetrix chip	Up-regulation of genes on the trisomic region. Global perturbation of the rest of the genome (~7000 genes dysregulted.)	Saran et al., 2003
Ts65Dn mice	Brain, liver, kidney, heart, muscle and lung	P30 and 11 months	4 Ts + 4 controls	RNA pools	RT-PCR	Up-regulation of genes on the trisomic region. Some genes escape this trend	Lyle et al., 2004
Ts65Dn mice	Cortex, midbrain, cerebellum, liver, testis, kidney, heart, muscle,	3-4 months	4 Ts + 4 controls	RNA pools	Custon cDNA arrays	Up-regulation of genes on the trisomic region. Some genes escaped this trend.	Kahlem, Sultan et al., 2004

Table 5-1: Transcriptome studies in trisomy.

It will be interesting to see if organisms that are less sensitive to trisomy exert more control over transcript levels. In *Drosophila melanogaster* for instance, individuals trisomic for an entire chromosome arm can survive to late stages of pupal development (Devlin *et al.*, 1982). The transcriptional activity of many loci appears to be negatively regulated or compensated in both larvae and adult flies with autosomal trisomy (Birchler *et al.*, 1990). A reduction in the level of gene products per gene template by one-third from the diploid quantity has been observed for several enzymes including *Sod*. Concentration-dependant repressors of these loci often reside in the duplicated chromosome arms, as described for the *Adh* gene regulator involving a closely linked *cis*-acting element (Devlin *et al.*, 1988). However, it is thought that this reflects a consequence of control mechanisms that normally operate in diploids rather than an attempt to compensate for the genetic imbalance in trisomy (Devlin *et al.*, 1988).

Dosage compensation has been rarely described for mammalian autosomal aneuploidies. It was recently shown that a dosage compensation mechanism occurs for the *Igf2r* imprinted gene in segmental mouse trisomy 17 (Vacik and Forejt, 2003). Thus, an indication is presented for a dosage mechanism of *Igf2r* in a trisomic context. Clinical reports of both paternal and maternal uniparental disomy for Chr21 indicate that there are no imprinted loci among the Chr21 genes. The results of our quantitative analysis suggest that regulatory mechanisms that can act positively or negatively may occur in some tissues. In a simple model, a repressor at dosage imbalance would have a corresponding affect on the level of expression of the gene it regulates. The results might be markedly different when the gene and regulator are in *cis* (both at dosage imbalance), compared to genes in *trans* that are present in the normal two copies when the regulator is present in three copies. The nature of such modifiers is at present unknown, although microRNAs similar to those regulating gene expression in *C. elegans* are obvious candidates (Lagos-Quintana *et al.*, 2002). Tissue specific expression of the regulators themselves could explain the occurrence of groups of genes that are similarly differentially regulated in trisomy. For example in RNA pools, *Dscr5*, *Mrps6*, seems compensated in skeletal muscle or midbrain in contrast to liver or kidney where they are largely over-expressed. Further analysis of gene expression profiles in trisomic samples will be helpful in revealing such mechanisms. However this remains to be

verified in individual mice. From the tested genes in individual Ts65Dn brains we could not identify an obvious mechanism of compensation, at least in the brain.

Given the fact that trisomic genes are globally overexpressed by a factor 1.5 it is however not possible to predict which genes will show dosage sensitivity in mammalian development. As yet we have little idea of the normal range of variability in gene expression that is tolerated (or required) by the organism. We know far less about the phenotypic consequences of a subtle up-regulation (e.g. 1.5 fold) of a given gene than for its depletion. For instance morphogens such as *bicoid* function in a dose-dependant fashion for which an increase of 50% of the gene product has significant impact on the phenotype, albeit less drastic than the one caused by a reduced amount of the gene product in *Drosophila* larvae (Driever and Nusslein-Volhard, 1988). Identifying dosage-dependant factors will require the knowledge of the molecular and cellular functions of these genes and of the biochemical mechanisms that are involved. We can postulate that triplicated genes coding for DNA binding proteins, including transcription factors, chromatin proteins, and RNA binding proteins may represent good candidates. These factors may exert a *trans*-acting effect by altering the stoichiometry of regulatory complexes setting the level of gene expression on target genes located elsewhere in the genome. For instance we have already identified disomic genes whose expression was altered in Ts65Dn tissues. We found that *Tff3* and *C21orf56*, located on MMU17 and 10, respectively, are strongly dysregulated in Ts65Dn mice, suggesting that these genes may be regulated by modifiers encoded by the triplicated region of MMU16. Identifying their promoter targets and analysing the expression profiles of the whole transcriptome would be essential to reveal expected cascade effects and finally to understand whether these genes cause a dose-dependant phenotype.

Our ongoing expression profiling study integrating the whole mouse transcriptome will shed light onto this issue. The consequences of trisomy 21 on the expression of other genes in the genome remain to be determined with robustness. In some but not all studies, the perturbation of gene expression levels has been confirmed to extend beyond trisomic genes to those that are disomic, affecting expression levels of a substantial proportion of transcripts in trisomic tissues in mice (Dauphinot *et al.*, 2005; Saran *et al.*, 2003). In one study of trisomic mouse cerebellum, up to one-third of disomic gene transcript levels were subtly

altered (Saran *et al.*, 2003). Very few of these genes showed a statistically significant difference with euploid when considered individually, but collectively, disomic gene expression distinguished the trisomic and euploid cerebellar transcriptomes. Nonetheless the number of replicates used in this study was limited and didn't allow the establishment of robust conclusions. Conflicting results regarding the question of perturbation of disomic gene expression have been reported in human studies (Mao *et al.*, 2003). The controversy is perpetuated by the use of different analytical approaches and questionable controls and samples used for array analysis in different studies.

The systematic analysis of mmu21 gene expression profiles in a panel of tissues contributes to an understanding of how cells and organisms respond to structural gene dosage imbalance. In the few example suggestive of a higher order of regulation in trisomy, it will be interesting to investigate whether the down-regulation or compensation effects are observed at different stages of mouse development. The consequences of differentially expressed genes and regulators at dosage imbalance is an important factor in understanding phenotypic outcomes of trisomy, which may result from compensation and down-regulation as well as overexpression of HSA21 loci.

5.3 Consequences of Gene Expression Variation

Studies using pools minimize inter-individual variations and have been useful in providing an averaged measure of the overexpression level in trisomic tissues and to identify possible outliers. Nonetheless, Ts/Eu ratios must be interpreted in conjunction with the distribution of the gene expression values in trisomic and euploid individuals. In order to evaluate inter individual variation of gene expression, we further measured expression levels of 33 trisomic and 17 disomic genes in several adult Ts65Dn and euploid mice. We also assessed Ts/Eu gene expression ratios in pooled RNAs, which allowed us to compare our results directly with those of the previous experiments (Kahlem *et al.*, 2004).

We observe that evaluation of individuals reveals variation of gene expression in the range of 20-50% for a large majority of the mmu21 genes, whereas only a few genes show either tight regulation (<10% variation in expression among

individuals), or dramatically different expression levels across individuals. Consequently, pair-wise Ts/Eu ratios span a broad range of values that could deviate greatly from the 50% overexpression range, indicating that Ts/Eu gene expression ratios need to be interpreted with caution. For instance, three genes that were previously shown to escape the 1.5x rule, *Bace2* in the cortex and *Kcne2* and *Sh3bgr* in the midbrain*,* show here a wide inter-individual variation could explain these skewed ratios. However, these genes had an average Ts/Eu ratio close to 1.5. Conversely, we observe here a higher order of dysregulation in response to trisomy for three genes in the testis previously shown to escape the 1.5 x rule*, Usp16, Dscr2* and *Bace2,* whose expression levels in pools and in individuals follow a comparable trend*.* Assessment of gene expression levels in individuals also provided further evidence for dysregulation of three disomic genes in Ts65Dn, *C21orf56* in the testis, and *Cbs* and *Nrip1* in the cerebellum, and this is independent of inter-individual variation.

We showed that inherent limitations to the technique (i.e., Real Time PCR) were unlikely to be a significant factor contributing to the measured variation, because the techniques are sufficiently sensitive to detect differences in expression values that are substantially smaller than those observed. This is also true for DNA arrays. Though that past discussion of arrays have often centered on technical issues and specific performance characteristic (Lander, 1999), now that DNA arrays have been used successfully for many different organisms the discussion has shifted to questions concerning the experimental design and data analysis. As in most expression profiling studies, our data represent a snapshot of the expression level in one individual at the time of death. We cannot exclude that the expression of some genes may be sensitive to the local environment (e.g. nutrition, temperature, stress, light, etc.). Inherent individual variations in the “personal statistics” of the mouse (weight, size, metabolite levels, etc.), all of which affect the number and proportions of cell types in tissues and organs, may lead to changes in the RNA population, as well. We cannot rule out that some of the expression variation that we observe reflects variation over time and/or cycling of gene expression levels. However, it is unlikely that cells are synchronized in a complex tissue, and such effects are expected to be averaged out for most genes. The genetic contribution to differences in expression phenotypes is not expected to be

large between strains of mice that are relatively close (i.e. C3H and B6), but the allelic variation between individual Ts65Dn mice still represents a factor to take into account. Any or all of these factors may contribute to the observed variation. Allelic variants of Hsa21 genes are present in different ratios in an individual with trisomy than in the diploid state. In the case where a mutant allele results in lower levels of gene product, this mutation will display recessive inheritance when the presence of one wild-type allele is sufficient to carry on normal function. A trisomic condition resulting in two copies of the loss-of-function mutation plus one wild-type copy would probably not alter the phenotypic outcome in this case. However, a recessively inherited phenotype can also occur when a mutant allele produces a gain or change of function, one copy of which does not produce a detrimental effect in the presence of a single wild-type allele, but two copies of which may be sufficient to “overcome” the buffering of a normal allele in a trisomic individual. However, in human, variation in the baseline expression level of many genes has been shown to be mostly genetically determined (Morley *et al.*, 2004).

We posit that this variation of gene expression, which was masked in pools, may provide insights into those genes involved in constant or variable features of DS, especially when considered in light of a threshold effect for gene dosage. Of course, the operative mechanism will involve the actual quantity of a gene product in a cell. This may become pathogenic once it passes a specific threshold (or drops below a minimum necessary for its function). While evolution has allowed a rather loose control of the expression of some genes, others are under constraint to be tightly regulated. It is not clear, however, which level of overexpression relative to the normal state can be tolerated without ill effects for a specific gene product, nor how sharp the onset of possible deleterious effects of overexpression could be.

5.4 Toward Identifying Candidate Genes for DS Phenotypes

Starting from the postulate that most of the trisomic genes are overexpressed by a factor 1.5, speculations on candidate genes were initially based on the molecular function of the genes. Favorite candidates include for instance tightly regulated gene products exerting *trans* effects, such as transcription factor complexes establishing concentration gradients during development, molecules

involved in epigenetic mechanisms modulating the accessibility of DNA to the transcriptional machinery, receptor-ligand-signal transduction systems, or proteins modulating the activity of other proteins. However, many genes have a pivotal role in various cellular processes and it is difficult to identify dosage-sensitive genes *a priori*. Dissection of the molecular basis for aneuploid phenotypes will require a massive body of information that is still largely incomplete, including detailed gene expression patterns within developing organisms (Gitton *et al.*, 2002; Reymond *et al.*, 2002), knowledge of genome-wide genetic networks as well as allelic contributions to variability in the level, place and time of expression, and on the variation of basal gene expression levels in the population. Understanding the pathogenesis that produces features of DS will require integration of this type of gene expression data with a quantitative description of variable phenotypic outcomes in DS. Mapping the regulators of HSA21 genes in man and in mouse is essential to understanding the genetic basis of the variation of gene expression and its contribution to pathogenesis of DS.

Beyond the guess of candidate genes based on functional clues, we postulated in our study that genes most revelant to the T21 phenotypes must be significantly overexpressed in disease samples despite the natural variation of expression in the normal population.

Our results show the importance of considering gene expression in individuals, and this approach will be particularly relevant for human samples showing wider genetic background heterogeneity than do the Ts65Dn mice. Normal variation of gene expression plays a role in susceptibility to complex diseases and likewise plays a potentially relevant role in the phenotypic differences seen between individuals with DS. Although DS presents with highly variable clinical features, some phenotypes are common to all, irrespective of the genetic background. We expect that these common features derive from dysregulated gene expression that shows the same pattern in all individuals. Here, we identified three classes of genes with different expression levels relative to euploids: the first class is populated by genes whose expression levels are systematically higher in trisomic than in euploid individuals, whereas genes with low or high degrees of intermingled expression levels form the two other classes. We postulate that genes in the first class represent good candidates for the constant phenotypes of DS. Five genes from the

first category are common to cerebellum, cortex and midbrain (*Gart*, *App*, *Mrps6*, *Ifngr2* and *Ets2*), identifying these as strong candidate genes for Ts65Dn neuroanatomical defects. Indeed the increased expression of *App* in Ts65Dn and Ts1Cje has been shown recently to be the cause of NGF transport disruption and choligernic neuron degeneration (Salehi *et al.*, 2006). The degeneration of basal forebrain chlolinergic neurons contributes to cognitive dysfunction in Down syndrome and Alzheimer's disease.

In contrast, the few genes whose trisomic expression levels overlap completely with euploid appear less likely to be key players for invariant features of trisomy. Among those, we found that *Kcne2* demonstrated a dramatic variation in its expression level regardless of ploidy. Expression levels for a number of genes fell between these two extremes. This may indicate the limit of precision for this method, but could also represent a pool of candidates for partially penetrant phenotypes. If the disomic level of a given gene is close to a critical threshold, then elevated gene expression might be deleterious only to those trisomic individuals with the highest expression, contributing to variability in the occurrence of DS features.

This approach provides a logical strategy for prioritizing candidates genes likely to contribute to brain phenotypes observed in Ts65Dn. The present analysis should be consolidated further by an exhaustive expression analysis in a large number of individuals at several stages of the development. It may be that the deleterious effect of overproduction of gene products occurs mostly at a specific place and time during development, when the level of the gene product is particularly high. It also appears that variability in the levels of the expression of a specific gene is a true characteristic of some genes that must be considered in a description of how elevated expression of a particular gene contributes to pathogenesis in DS. We are well aware that the level of mRNA does not systematically reflect the corresponding protein amount but is however a very good indicator of the response of the cell to trisomy.

As shown here, the stratification of populations by expression profiling provides an essential dimension in the molecular analysis of aneuploidy syndromes (Antonarakis *et al.*, 2004). Identifying the pathways perturbed by trisomy will require thorough studies on expression phenotypes at the level of a global transcriptome

and integration of other large-scale experiments designed to decipher gene regulation networks[22]. The advances in high-throughput methods have provided us with a first snapshot of the overall structure of molecular interaction networks in biological systems. In future, instead of viewing the genetic parts list of an organism as a loose collection of biochemical activities, in the best scenario, discrete networks of function will bridge the gap between genotype and phenotype. A complete network picture will probably require the integration of data obtained from a broad range of approaches and methods with modeling efforts at many different levels of details via systems biology (Klipp *et al.*, 2005; Morley *et al.*, 2004), taking advantage of the large data sets produced by various systematic methodological approaches in functional genomics.

Our contribution to the functional genomics of Chr21 genes englobes the analysis of their functional orthologs in model organisms. One aspect is the systematic analysis of gene expression patterns in the mouse embryo by means of *in situ* hybrisization, which are carried out in the laboratory as part of European consortium (www.eurexpress.org).

5.5 Functional Analysis of Candidate Genes

Using lower eukaryotic organisms as test tubes for investigating gene function *in vivo* offer excellent tools for functional genomics reflected by the wealth of data available for *D. melanogaster*, *S. ceraevisie* and *C.elegans*. The nematode *C. elegans* is an attractive experimental model to study gene function in a living organism. Despite its apparent rudimentary organization, the main cell types affected in DS, including muscle cells and neurons, can be individually identified (Brenner, 1974).

Here, we determined the spatial and temporal expression of five HSA21 orthologous genes in *C. elegans*, and could reveal strong patterns of similarity in expression and subcellular localization of some of these proteins across phylogeny. For example, like its worm ortholog, *Sh3bgr* expression in the mouse is cytoplasmic

[22] Gene regulatory networks are the on-off switches and rheostats of a cell operating at the gene level. They dynamically orchestrate the level of expression for each gene in the genome by controlling whether and how vigorously that gene will be transcribed into RNA.

and mostly localized to heart and skeletal muscle and is also detected in some visceral smooth muscles (Egeo *et al.*, 2000; Gitton *et al.*, 2002). The results also demonstrated that *Sh3bgr* is expressed in earliest stages of development in mouse and worms. These findings support a possible role for this gene in heart morphogenesis and consequently, in the pathogenesis of congenital heart disease in DS. The *CBS* ortholog in the mouse was described to be expressed predominately in liver, skeletal and cardiac muscle and in the nervous system during early stages of development (Robert *et al.*, 2003). We showed herein that the expression of its ortholog in *C. elegans* is predominately cytoplasmic in the intestine, the pharynx and in head and bodywall muscles. For the other three orthologs that were investigated herein, no expression data are available in other organism. The specific pattern of expression of the orthologs of *HEMK2* and *TMEM50B* and the high level of homology among human, mouse, *Drosophila* and *C. elegans* indicates an evolutionarily conserved biological role. In *Escherichia coli*, for example, the *hemK* gene encodes the N^5-glutamine methyltransferase that modifies peptide release factors (Heurgue-Hamard *et al.*, 2002).

Currently functional analyses of novel genes in many species focus upon the consequences the disruption of gene function (Zambrowicz and Friedrich, 1998). Double stranded-mediated interference was found to be a potent and specific inhibitor of the endogenous gene activity in *C. elegans*, *D. melanogaster* and in mammalian cells providing a valuable method to elucidate aspects of gene function (Elbashir *et al.*, 2001; Fire *et al.*, 1998; Kennerdell and Carthew, 1998). Although certain features of the mechanism responsible for the RNAi response appear evolutionarily conserved (e.g. Dicer), the phenomenon of "spreading" exhibited by *C. elegans*, whereby the worm produces a systemic response to the localized introduction of dsRNA, is more species-specific. Delivery of dsRNA into *C. elegans* by microinjection (Fire *et al.*, 1998), soaking (Tabara *et al.*, 1998), or feeding (Timmons and Fire, 1998) can lead to the systemic depletion of targeted mRNAs, with the exception of a few resistant cell types (e.g. most neurons). Our results using the feeding method remained modest, as only the ortholog of *KiAA0179* gene had a visible phenotype when its expression was silenced and additional experiments will be necessary in order to decipher the biological function of

candidates. To further determine which of the HSA21 genes are involved in DS, the effects of disrupting and overexpressing the orthologs in model organisms, such as *C.elegans*, can be analysed. The use of C.elegans will provide an excellent experimental model for the initial characterization of gene function and may become an important tool in assessing the contribution of genes in complex phenotypes such as DS.

5.6 Perspectives

Aneuploidy and segmental aneusomy syndromes are common and complex. Improving our understanding of the molecular genetic basis of aneuploidy syndrome rely on highly informative data integrated from various aspects of functional genomics ranging from chromosome architecture and gene regulation to the effects of altering the stoichiometry of the proteome (Nobrega *et al.*, 2004). Investigations of the genetic basis for these disorders will inform our basic biological understanding of gene function and cellular pathways as well as knowledge of individual aspects of each syndrome. This is relevant not only to individuals with aneuploidy syndromes but also to the euploid population. Individuals with DS have an increased rate of leukemias and an earlier onset Alzheimer's disease compared with the euploid population, and an understanding of these disorders in DS will be informative for similar disorders in euploid humans. As with many other human disease under study, animal models are critical for our understanding of the disease preocess, and without such models, the difficult task of unraveling the causes of syndromes associated with aneuploidies and segmental aneusomies would be impossible. Defining the etiology of genetic mechanisms in DS requires knowledge of the trisomic genes, their expression patterns in time and space, and their downstream effects, direct and indirect, on the expression of other genes. This information must be linked to a precise description of phenotypic consequences, not only in fully differentiated cells, but also at all stages where euploid and trisomic developmental processes diverge. Animal models, including critically important segmental trisomies and monosomies in mice, provide a substrate for testing hypotheses about how overexpression of genes individually or in concert can affect development. The precision with which a phenotype and its etiology can be explained in mice points to potential difficulties and limitations with regards to the

extrapolation to humans, where phenotypes are defined clinically for practical applications, and not necessarily with the precision required for genetic studies.

Recent advances suggest that the molecular basis of trisomic phenotypes are perhaps even more complicated than assumed for many decades. What then is the most effective way to understand and, more importantly, to ameliorate the effects of trisomy 21 on development and function? As discussed here, no single approach will uncover the myriad sources of divergence from normal development and function initiated by trisomy. With the advent of new gene therapies for single gene dominant disoders (such as RNAi-based approaches) and recessive disorders, there is real hope that identifying key dosage-sentsitive genes may open up possibilities genetic therapies that could tackle some of the non-developmental aspects of some clinical signs of the syndromes, and thus maybe improve in furture the well being of DS patients. Small molecule therapeutic approaches may also possible, once key pathways have been identified.