AUC Undergraduate Journal of Liberal Arts & Sciences

Capstone Issue Vol. 7 2015

AUC Undergraduate Journal of Liberal Arts & Sciences
Capstone Issue 2015

Typesetting & Design
Barlock

The AUC Undergraduate Journal of Liberal Arts and Sciences is a biannual, interdisciplinary publication showcasing outstanding undergraduate academic papers. The Journal aims to demonstrate the strength of undergraduate scholarship at AUC, to reflect the intellectual diversity of its academic programme, to encourage best research and writing practices, to facilitate collaboration between students and faculty across the curriculum, and to provide students with opportunities to gain experience in academic reviewing, editing and publishing.

FOREWORD

The summer issue of the AUC Undergraduate Journal of Liberal Arts & Sciences is dedicated to the AUC Capstone project, the final thesis that in many senses is the culmination of teaching and study at our institution. In these pages, we aim to showcase to you the breadth and height of scholarship reached by the students of our graduating class.

Inside, you will find one piece of work from each of our majors: Humanities, Social Science and Science. Those works included in this journal are among those that have been awarded 'Thesis of Distinction' or 'Thesis of Highest Distinction' by this years Capstone Awards Committee. We mean to both celebrate the significant achievement of all those in the Class of 2015, and to encourage those who have yet to complete this step.

As the university moves into another academic year, we still have in mind the questions that were openly tabled last semester about the future and direction of higher education. What our students consistently demonstrate is their desire to be at the forefront of academic research and thinking. These authors have skilfully and boldly tackled contemporary issues, and the ambition of their work should be nothing less than an inspiration to us all.

Zinzi Mangera-Lakew

AUC Student
Editor, InPrint

CONTENTS

Curating Public Space

Museum practice on Museumplein

Lucy Cowling

ABSTRACT

Museumplein is a prime example of all that public space can be - a multifaceted, agonistic space within which the dominant hegemony is challenged by counter-hegemonic practices. Art plays a pivotal role in this battle of representation, as it forms part of the articulatory practices through which order is both asserted and subverted. In order to assess whether artistic interventions in space work to consolidate neoliberal structures or instead undermine this logic, public space is reframed as curated instead of as produced. This makes it possible to identify play and playfulness as an underlying common thread in most of the critical counter-discursive art visible on Museumplein. Within the contemporary curatorial methods shaping urban sights a further distinction can be made between permanent and temporary interventions, as these two practices intentionally mirror methods of display inside art institutions and thus achieve different results. Furthermore, spatial analysis in terms of curation highlights which of the recent interventions by the museums agree, and which disagree, with the municipality's desire to transform the square into the "fifth cultural institution" by 2020. The 'institutionalization' desired by the municipality results in commodification by facilitating the mass-consumption of culture. The practices by the museums often also combine artistic and commercial goals. This is most clear with the activities of the Rijksmuseum, but the permanent 20th and 21st century collection in the appropriated public space of its garden forms an exception. These objects disrupt order through play. The Van Gogh Museum’s artistic contribution to Museumplein comes in the form of its exhibition wing, but the expansion of this space can be regarded as an authoritative consolidation through commercializing an aesthetic structure. As the Stedelijk Museum Amsterdam is most active in commissioning and curating interventions that undo daily order and hegemonic structure, it is the most successful at undermining commodification on Museumplein.

INTRODUCTION

The three main visual art museums of the Netherlands, namely the Rijksmuseum, Stedelijk Museum Amsterdam and Van Gogh Museum, collectively attracted over 4.8 million visitors in 2014 (van Lent and van Os). This made the location that houses all three - Museumplein - the Netherlands' biggest visitor attraction. In 2008 the municipality announced a plan to remodel the square and its surrounding area to cope with the rising numbers of visitors. In this proposal they state that Museumplein itself will become the "fifth cultural institution" (ter Berg et al. 7). If Museumplein can indeed be regarded as a cultural institution in its own right, then it must produce cultural and artistic meaning. Furthermore, a space structured according to an institutional model implies a less open and more privatized character than the classic conception of public space. Hence, at surface level it seems that the conglomerate, hegemonic, and commodified nature of Museumplein can be solely accredited to the municipality's careful urban planning. However, the museums are significant stakeholders and instigators of change in the structure of Museumplein. A prime example of this is that the main motivation given by the municipality for their on-going plan to renovate the square was the earlier announcements by the Rijksmuseum and Stedelijk Museum to renovate and expand their institutions (Karsten and van Diepen 7).

These expansions have significantly changed the way a visitor to the square experiences the space. All three of the major visual art museums have relocated, or in the case of the Van Gogh Museum are still in the process of relocating, their main entrance onto the Museumplein side of their building. These very acts can be simultaneously read in terms of curation and commodification of space as they function to aestheticize and assert the eminence of the institution, but also allow higher numbers of visitors to pass through their doors. Furthermore, these moves signify a shift of focus from closed, individual institutions to an open, accessible, conglomerative space of national culture. It is not only these architectural changes that have left their mark on the space, as the redesigning of the institutions also resulted in other interventions in space, both through landscape design and artworks placed around the buildings.

Due to the recent and significant changes, Museumplein forms an ideal case-study within which to analyse the effects of artistic interventions in public space. Museumplein is a prime example of all that public space can be - a multifaceted space of conflicting interests within which art plays a pivotal role in shaping order or disorder. Therefore, departing from Henri Lefebvre's imagination of space as something produced, this research proposes to re-imagine the spatial activities on Museumplein as 'curated'. The conceptualization of Museumplein's spaces as a palimpsest of meanings and various modes of production underlines the commodified nature of the public sphere. By reframing Museumplein as curated, its very nature shifts to focus on artistic interventions, thereby incorporating the interconnections and disjunctions between the commodified space of use and the exhibitionary space of display. Commodification and aestheticization of space are entangled, both in practices by the museums and the municipality, yet analysis using the lens of curation shows there are differences in intention and achieved results. Such an understanding of space highlights which of the recent artistic interventions by the art institutions agree, and which disagree, with the municipality's desire to transform Museumplein itself into a "fifth cultural institution".

In order to synthesize the concepts of curation and public space this research will first analyse the critical debate regarding the philosophical perception and production of public space and its relation to artistic practice. The re-framing of space as curated will be aided by determining which curatorial strategies have been appropriated for artistic practice in the public sphere, hereby making a distinction between acts that playfully disrupt or consolidate order, as well as between permanent and temporary interventions. Finally, this research will turn to the urban conception of Museumplein and identify what permanent and temporary interventions were visible during the month of May 2015, or will make themselves manifest in the near future. In doing so, a dialogue is staged between the discourses of philosophy of space, urbanism, and curation, in order to expose and analyze the underlying tensions between the public and the private and between commercial commodification and artistic expression - both in the entire square and within the institutions. The vocabulary is intrinsically shaped by the case-study of Museumplein, as the urban discourse is shaped by a site-specific historical development, policy, and presence in the city of Amsterdam. However, the theoretical framework and methodology that proposes a reframing of space from produced to curated can be applied to a wide variety of institutional artistic practices in urban public spaces.

2. PRODUCED PUBLIC SPACE

2.1 PUBLIC SPACE

What it means for space to be public has, in recent decades, surfaced as a complex conundrum. This question has attracted much discussion, as the definition and functions assigned to public space have great implications for the relation between subjects and their society, politics, and culture (Deutsche Evictions 269). Michel Foucault even identified that there was a "spatial turn" during his lifetime, where scholarship moved away from an interest in time to usher in the "epoch of space" (2). An important first work to emerge in this logic of space was Jürgen Habermas' The Structural Transformation of the Public Sphere, which looked back at the late 18th and early 19th century emergence of free press and coffeehouses as spaces outside of state control, thereby forming an alternative public space for the formation of culture and ideology. A Habermasian conception of the public sphere sees this space as a place where consensus can be reached. However, it has now become clear that the idea of urban public space being unified is less applicable to modern and postmodern societies. Arguably, in current neoliberal societies the opposite is true as consensus may well be attempted from the top down, but is not easily achieved on street level. Here tactical expressions can resist strategic, dominant narratives (de Certeau 21). Jacques Rancière even introduces the term 'dissensus' to describe these activities that disrupt forms of domination - be it political or aesthetic (Dissensus 35). In keeping with this idea, in the 1970s and 1980s a neo-Marxist discourse emerged framing public space as a disunity. Instead, public space is an active battleground where the dominant hegemony tries to assert itself, but oppositional voices find room to challenge common assumptions and the prevailing symbolic order (Mouffe 11). The development of this discourse went hand in hand with the firm establishment of neoliberalism, and thus the dominant hegemonic structure to be challenged became one of increased commercialization. The neo-Marxist "agonistic model" of public space imbues it with political meaning (Mouffe 13). Public space is thus to be understood as a social construction, or a democratic invention, where different political, social, and cultural meanings are produced, asserted and contested (Deutsche Evictions 274; Lefebvre 15; Lefort 19). Lefebvre intensifies this conception by arguing that "(social) space is a (social) product" (26).

Understanding space as a product does not mean it is created within a neutral spatial geometry - it implies space is shaped through various social relations and their mutual interests and that it is inherently commodified. Different modes of production can function simultaneously in one space. Lefebvre discusses a tripartite split in space, between the conceived space of urban planning; the perceived space of everyday life; and the mental or imagined space, which can manifest itself through art and culture. Within perceived space a further distinction can be made between space produced for domination, or the reproduction of capital, and space produced for "appropriation", or the way subjects use their surroundings in everyday life (Lefebvre 230 - 4). Particularly important for this research is the idea that mental or imagined space has the potential to transcend and then restructure the balance between these various modes of spatial production (Lefebvre 92 & 382). Artistic interventions in space - however hegemonic and commercial that space may originally be - can offer "a counter-gaze" and thus have the potential power to alter the perception of space (Lefebvre 382). At the same time, however, rather than subverting the conceived and intended perception of space, an artistic intervention can also be utilized to assert the dominant symbolic structure. In order to recognize which result an artistic intervention achieves it is first necessary to gain an understanding of what art being 'public' actually entails, as well as how public artworks arrive in space.

2.2 PUBLIC ART IN PUBLIC SPACE

Unsurprisingly, with public space being a multifaceted arena of conflicting power structures, art in public space is an equally complex notion. Public artworks are not simply characterized by their open access or by being a public good, as they are still owned and financed by parties with private interests (Hein

3 - 5; Phillips 332). Public art can therefore not be differentiated from art in private or museum collections based on clear distinctions between public and private interest. However, art placed in public space is distinctive from other artistic expressions, and this is most evident when seen through its spectators. Public art confronts a public that may not choose to look at art. This forges relationships on various levels: between the public viewer and the work, between the work and the space it occupies and between the viewer and the space they use (Serra qtd. in Kwon 86). What then truly differentiates public art from other art is not just the active occupation of communal space, but the fact that spaces within the public sphere enforce various engagements that result in the interchange of communication and critical discourses (Hein 5). Thus, public art is not defined by democratic notions such as "accessibility or volume of viewers", but through the "questions it chooses to ask" and the thoughts it raises in the minds of its varied body of spectators (Phillips 332). When reversing this idea away from the spectators and back to the artwork itself, public art becomes public "because it is a manifestation of art activities and strategies that take the idea of public as the genesis and subject for analysis" (Phillips 332). For these reasons, public art in public space is a direct intervention in daily life that holds the potential to challenge or even overcome the political.

Art and politics are inevitably entangled and are hard to separate. Rancière advocates for the interconnectedness of the two in his The Distribution of the Sensible (13). The book describes how all activities within society are distributed and result in the interrelation of various modes of production. He argues that the 'distribution' of art in public space enforces a common or shared ground within that space and between groups of users (Distribution 12). Art in public space thereby welcomes participation whilst simultaneously defining itself as exclusive parts or "segmented common ground" (Distribution 12). Public art in public space is thus always politicized, as it is politics that mediates the distribution. This works due to Rancière's conception of politics being exercised by marginalized groups battling for equal recognition within the dominant hegemony (51). Art and aesthetics become weapons in this battle as representation and recognition in society can come through the possibility to display a particular image, object or performance. Framed in this manner, public space becomes agonistic. The political theorist Chantal Mouffe agrees that in this agonistic model of space art or artistic interventions can serve an antagonistic function. Art interventions in space becomes the "articulatory practices through which a certain order is established . . . [or] challenged" (Mouffe 9). She argues that public space is an active and disunified arena, where the dominant hegemony can attempt to assert itself through artistic practice, but equally oppositional voices can make themselves manifest through critical public art or artistic expression (11).

Artistic practice in public space are forms of disruption of everyday life and can either subvert or assert the established order. To establish how this may become visible in cities it becomes important to distinguish between institutional practice, city commissions and artist initiatives in open urban space. Each type of artistic intervention arrives in space due to very different interests. A large portion of public art is commissioned by local government bodies or businesses and only gets created and placed in space after being approved by citizens and stakeholders (Wong 44). This type of public art externally communicates a vocabulary of democracy and inclusion, by making art publicly visible and involving different parties in the decision-making. However, ultimately the works cannot uphold this vocabulary as the location and artist's brief remain with the powerful bodies that decide who or what to include and exclude (Deutsche Evictions 273; Phillips 332). The function of this public art can be diverse; it can seek to commemorate, monumentalize, gentrify or simply embellish space, but it does so in line with the dominant hegemonic order. It is thereby often invested in a spatial construction, which is informed by capitalist production. There are less mediated forms of public art, including graffiti and similar street art, that do not always receive approval from external stakeholders before occupying space (Wong 44). In these cases the artist alone can select the work and site. This type of art, as well as various other forms of critical public art has the potential to instigate dissensus and make "visible what the dominant consensus tends to obscure and obliterate (Mouffe 12). Thus, these two examples show that the spectrum public art forms ranges from government approved 'communal' decision making to individual 'counter-curation' that seeks to be critical of the establishment. Therefore, whether an artistic intervention corresponds to a dominant order or serves a counter-discursive agonistic function, is not only dependant on how the public interacts with the piece, but equally relies on the artist's intention as well as that of potential external stakeholders (Mouffe 13).

Within this wide range of types of public art, institutional interventions stand in between the two extremes. Institutional practice in public space holds the potential to aestheticize space, breaking with the assertion of space as produced for commercial gain and transforming it into a space of exhibition. This type of intervention in everyday life breaks through hegemonic structuring that seeks to commercialize space. However, art institutions themselves cannot be fully external to cultural hegemonic structures. If importance is ascribed to this, interventions in space can become an assertion of control and authority over the space outside of the institution as well as the walls inside. The intention or used method when 'curating' public space through institutional interventions thus becomes relevant when analyzing the results of placing art objects outside the institutional interior.

The discussion regarding public art and public space has made it clear that artistic interventions in space embody different combinations of conflicting yet at times simultaneous functions; they are public and private; can assert or subvert a hegemonic structuring of space; and are a politicized social product yet form a poignant visual reminder of importance beyond modes of production and capital. Therefore, to make sense of all this conflicting interest, it is useful to understand how these different types of public art relate to contemporary models of curation, as this type of framing is a tool that can help identify how artistic interventions function in space.

3. CURATED PUBLIC SPACE

3.1 CURATION

Since the establishment of art institutions in the 19th century, the term 'curator' has been understood as applicable to an individual who is responsible for managing and caring for the valuable objects in the collection of a museum (Huyssen 13; Hooper-Greenhill 3). As artistic media and styles developed throughout the 19th and 20th centuries the curator became increasingly involved with the modes of display and exhibition management, becoming a voice of authority in the art world. The verb 'to curate' came about in the course of the 20th century and further activated and mobilized the display of collections (Huyssen 21). Contemporary developments have seen curation move even further away from management of specific collections to processes of display, interpretation and representation in a much wider variety of media and locations (Chaplin and Stara 1). The broadening of the term curation, now arguably incorporating modes of commercial display and branding, can be interpreted as "yet another fad of consumerist postmodernity, requiring ever changing ways of selling everything" (Chaplin and Stara 1). Indeed curation can make itself manifest in spaces that include commercial interests, but it not always utilized in order to increase commercialization of space (Wong 43).

With curatorial practice expanding outside the institution, the broadening of contemporary curation also moves further away from a concern with objects, towards a model which incorporates the original meaning of 'cura', the Latin root of the word. 'Cura' translates as care, but it was used not in relation to objects but in a spiritual and sacred sense (Gough 95). Used in this sense, curation moves beyond a primary concern with the appearance and form of objects, instead seeing the arts as the visual manifestation of the 'sacred' in everyday life. Extending this definition to urban environments and contemporary times, the 'city curator', urban planner, or spatial architect is responsible for aptly utilizing social constructions to take care of a place, in order to ensure its 'sacredness' or inherent meaning in a world which is increasingly preoccupied with capital and material goods (Gough 101). Following this definition, contemporary curation becomes less concerned with authoritarian voices, and more invested in a sense of community and place. If this is the case, curation is not appropriated to adhere to privatized and commercial interests, but it can undermine these (Mac an Bhreithiún 270).

3.2 CURATING URBAN SPACE

The curation of the urban fabric adds meaning and narrative, transforming an undefined space into a destination or place (Storrie 3; Mac an Bhreithiún 255). As has already been assessed, public art is partly defined by the communication exchange between the object and its broad public. Curation is the active hand in placing artistic interventions and mediates the tripartite relationship between viewers and space. For this reason, the contemporary development that broadens the definition of curation away from a single voice inside a museum into a pluralistic understanding suits the shaping of urban sites through public art and artistic interventions, as it can give the public a voice (Wong 49).

Curation can shape the agonistic public space. According to Michel de Certeau, urban environments form a stage for users of space to perform a "tactical art that plays with the structures of place available" (21). As de Certeau views 'tactics' as a counter-discursive device, playful art can thus be a form of critical art that can undo structure (21; Mouffe 9). Therefore, curation of the urban space can be approached from two premises: it can treat the city as an authoritative, institutionalized space or as a site of discovery, exploration, and exhibition (Wong 41 & 49). This exhibitional model particularly can create a communal sense of place, or imbue 'sacredness' into space, by reminding the public of importance beyond capitalist modes of production. Following these models, artistic interventions can function to consolidate strategic structure and hegemonic order, or disrupt these ideas by inciting critical thinking, for instance through the notions of play, surprise and imagination (de Certeau 21; Roode and Lefaivre 7). It is however not possible to make a definitive split between playful disruptions functioning as counter-discursive subversion of structure and voices of authority asserting existing power relations. Since public space is complex and multifaceted, there are interventions that utilize the concept of play yet still consolidate the dominant hegemony.

As the definition of curation shifts away from a sole-concern with museum practice, the curation of public space undertaken by art institutions is of particular interest. Art institutions have historically been regarded as closed, authoritative, stable entities that through intervening in space are opening up a discussion with agonistic and dynamic public life. Their current interventions in public space signify that art institutions are adapting themselves to the more pluralistic form of curation, thereby acknowledging contemporary museological tendencies that place diverse spectators at the centre (Wong 108; Huyssen 14; Hooper-Greenhill 82). They are thereby in a sense breaking down their walls, in order to open up and interact with a larger audience (Huyssen 21).

The curation of space can direct how art functions within the public sphere. When utilized as a counter-narrative by non-institutional bodies the exhibitional use of space can, for instance, question or undermine the authoritative power still largely held by visual art institutions (Mouffe 13). When utilized by those institutions it can further foster a more incorporative form of curation and art display that fits within the current movement visible throughout art institutions to become less authoritative and elitist, and instead include or interact with the greater public's voice. This can be done both in order to strengthen their position and to playfully undermine daily order.

3.3 PERMANENCE AND TEMPORALITY

Museums have been regarded as the antithesis of public art (Deutsche "Art" 37; Hein 6). The claimed division between practices inside museums and public art is asserted with the notion of temporality in combination with public life. Impermanence or adaptability are seen as positive characteristics of public art as the constant flux within the public sphere gives it a temporal nature (Hein 2; Phillips 331). 'The public sphere' is an invention - or to revive Lefebvre's argument; a production - which is constantly being reinvented (Sennett qtd. in Phillips 331). According to the curator, art critic and academic Patricia Phillips "enduring objects [have] become as quixotic as time itself" due to the accelerated societies of the 20th and 21st century that have resulted in the increased compression of time. Therefore the desire society has to stay "timely" is best expressed through temporary visual markers (331). Within this logic, art functions as the material symbol of the sentiment of a certain time and place, which can be an active locus in the constantly transforming, and antagonistic public sphere. Art then becomes more historicized and monumentalized once moved into the private sphere (332). It is for this reason that museum practice has been placed in direct contrast with public art, as adding art objects to a museum collection ingrains those pieces with permanence and historic value (Deutsche "Art" 39).

This line of argumentation legitimately acknowledges the agonistic space of the public sphere and the critical function art can have. However, it is unjust to accredit this potential solely to temporary pieces and not permanent objects. Permanence and historic value do not equate to static meaning, especially not when a permanent object is placed within the flux of disunified and antagonistic space. These works do explicitly contribute to the evolving critical discourse and communication exchange in public space that is vital for public art, namely by reminding the public of what to remember or forget. Both permanent and temporary art objects contain the potential to politicize space and can function as agonistic disruptions so no clear distinction can be made between their desired and realized effects in space on these grounds.

A distinction between permanent and temporary interventions remains valuable when framing artistic interventions in space through modes of curation, as the two notions are foundational for curatorial practice. This is particularly relevant for public art practices by visual art institutions. Museums almost always display both a permanent collection and a temporary exhibition - attributing equal legitimacy to the historicity of fixed objects and the fleeting significance of the ephemeral convergence of objects within the narrative of an exhibition (Huyssen 20). In this sense, museums do not function as the antithesis of public art, but rather mirror artistic activities in public space as they actively practice both types of object display and thereby constantly reinvent their own historical and critical narrative. The curated display of art inside an institution can therefore be seen to translate to the display of public art in urban space.

The perceived disunity between temporary and permanent public art is not visible within institutional practice in space, as their actions help to legitimize permanence in spaces of flux. It is important to acknowledge that it is not just temporary works that can act counter-discursively. By treating the public sphere as a space of exhibition that can be shaped through curation - be it of temporary or more permanent nature - museums play a role in shaping the binary between the agoric 'democratization' of space and the privatization and commodification of public space that is increasingly visible in neoliberal societies (Deutsche Evictions 276). The theoretical discussion has shown that artistic interventions can politicize space, either by consolidating hegemonic structures or undermining these through playful or critical disruption. Both permanent and temporary objects can serve these functions, but their results are often achieved through different artistic ideals and intentions. The various types of artistic interventions can all be found on Museumplein, and therefore, to aid the analysis, the various practices will be categorized according to these distinctions.

4. MUSEUMPLEIN - THE DEVELOPMENT OF URBAN PUBLIC SPACE AND ITS MUSEUMS

4.1 THE CREATION OF MUSEUMPLEIN

Already in 1877, immediately after plans were announced to build a new state museum now known as the Rijksmuseum on the outskirts of Amsterdam, plans were developed by the museum's architect Pierre Cuypers to model the open area behind the site into a public event space (van der Werf 19). Cuypers shaped the space to suit the needs of 'his' Rijksmuseum, which included influencing

Figure 1. Aerial photograph Museumplein, 2012. Image courtesy of Stedelijk.nl

the municipality's plans in such a way that the Rijksmuseum was at the centre of all viewing axes. The integration of a traffic thoroughfare and the later inclusion of a garden on all sides further strengthened the museum's interaction with its direct surroundings (ibid 25-8). The space behind the Rijksmuseum promptly became the sight for all large-scale public cultural and sport events, including parades, skating competitions, and the world exhibition of 1883 (ibid 8). By the end of the 19th century, urban development - including the building of the Royal Concertgebouw and the Stedelijk Museum - framed the boundaries of what had now become a square. This period also saw the shaping of the entire Museumkwartier into the urban fabric that has remained largely unchanged up until the present day. However, Museumplein itself has undergone various stages of development, the most notable of which being Sven-Ingvar Andersson and Stefan Gall's large scale restructuring planned and implemented between 1993 and 1999.

The landscape architect Andersson was responsible for designing the green open public space that the municipality desired and Gall supported his designs with urban analyses (ibid 67). Andersson had already suggested in his original plan that all the museum's entrances should be on the square, and the space in front of the Royal Concertgebouw should be more incorporated with the rest of Museumplein, elements that were rejected at the time, but have since been implemented or are under review in the current plans of the municipality (van der Werf 16; ter Berg et al. 7). Andersson did succeed in redirecting almost all flows of traffic around or under the square instead of through the centre of the space, achieving a combined effect of vast green open space and the hustle of pedestrian city life on the parameters passing the museum buildings. The centre of the square was to be freed from the obstruction of built structures, but Andersson did allow for new buildings along the edges, including space for the Stedelijk Museum and Van Gogh Museum to build extensions onto the square. The Van Gogh Museum immediately took up this opportunity, as Kisho Kurokawa - the architect appointed to design their new wing - collaborated with Andersson on the placement, height, and shape of his design (van der Werf 81). Furthermore, Andersson unified and opened up Museumplein by replacing the individual areas sectioned by the Stedelijk Museum's garden and dense tree lines with one open space. This consisted largely of grass, sometimes intersected with stone lines that connect circulation axes, viewpoints and cultural institutions with each other (figure 1 and figure 4).

Historically, the three museums have always been significant stakeholders in the shaping of Museumplein. Initially this can be seen through the arrival of the institutions, as they created the physical boundaries of the square. Shortly thereafter, when the Stedelijk Museum and Rijksmuseum opened gardens, the museums strongly contributed to the closed 'back-garden' feeling that Museumplein was accused of having before Andersson and Gall's developments (van der Werf 31). Lastly and most recently, the creation of a car free open public space prompted all three of the museums to relocate their entrance on the Museumplein side of their building. In turn, these changes caused the municipality and the district office Amsterdam Zuid to draft a new 'vision' for Museumplein that could support the increasing numbers of visitors the museums can now accommodate (Karsten and van Diepen 7).

The 'masterplan' for Museumplein was published in 2008 and is projected to be completed by 2020 (ter Berg et al. 47). The plan aims to connect all the institutions, transforming them from "loose pearls to one necklace" (ibid 23). This will be achieved on a bureaucratic level by facilitating cooperation and discussion between the institutions, and on a practical level by transforming the current tram thoroughfare on the van Baerlestraat into the "urban foyer" of the Royal Concertgebouw, thereby fusing the square in front of their building with Museumplein (ibid 7). The public space of Museumplein itself is thereby supposed to grow into the "fifth cultural institution" (ibid 7). Theoretically, this statement has a major implication. The desired institutionalization of Museumplein strongly detracts from its public nature, as institutions have physical, economic and authoritative boundaries that are not characteristics of open urban public space. The institutionalization of Museumplein would thus further increase private interest within the square. In practice, the only cultural addition actually proposed in the plan is a pavilion which should arise on the east side of the square, and that will program cultural and musical activities and performances (ibid 31). Other than that, the municipal plans only incorporate commercial goals, such as the increased amount of shops, galleries and particularly restaurants and cafés. It is almost exclusively concerned with how to increase tourist consumption (ibid 12). Thus advocating for structural changes that will transform Museumplein into a "cultural institution", is in fact an implementation of practices that seek to privatize and commodify. The claim functions as a mask for mass-commercialization, and fundamentally asserts hegemonic structures.

The institutions themselves are also all concerned with enlarging their own capacity. The continuous desire to increase visitor numbers is clearly reflected in the decisions to renovate and expand their buildings. These choices are intimately tied to their artistic practices, and therefore it is vital to assess how these and other strategic choices motivate or influence the artistic interventions by the three museums on Museumplein. The various mission statements and methods of display implemented inside the spaces of the museums will be discussed below as they provide an insight into how they view their own role within a context that is greater than their own walls. In this manner it can be assessed if the desires expressed in their strategies and curatorial choices inside reflect or contradict their practices and influence in the space around their buildings, as well as whether these practices subvert or assert the municipality's vision.

4.2 THE MUSEUMS

The Rijksmuseum promotes itself as nothing less than "the museum of the Netherlands" that sets out to "link individuals with art and history" ("Vision and Mission"). Fitting with this mission statement, the building's architecture and interior design, with its neo-Gothic grandeur and Catholic embellishment, presents itself as a castle or temple for art. Furthermore, Cuypers succeeded in reiterating the perceived importance of the Rijksmuseum through its gateway function into the city and its central location on Museumplein. This has remained unchanged. Since the reopening in 2013 the curatorial model inside has been completely altered, and their current display strategy also reflects the museum's mission statement. Before the renovation all objects were displayed within their distinct medium. With the renovation came a complete reordering of every single object and model of display, with the only piece to remain in its old place being the Night Watch by Rembrandt van Rijn (Pijbes). The methodological structuring of the collection now displays all media together within their shared historical context, thus not only connecting aesthetic value with historical importance, but also creating a curated narrative of Dutch history for the viewing public.

However, despite the museum presenting itself as a temple devoted to art, the renovation has also highlighted the commercial aspect of the institution. Not only is the new entrance Atrium built to accommodate many more visitors, newly launched initiatives are also indicative of a more commercial standpoint. A very direct example of this is Rijksstudio, where online visitors are invited to design and purchase a wide variety of products using digital reproductions of objects in the collection. Within the physical space, a significant wing of the museum has been transformed into a restaurant called RIJKS®. The 'registered trademark' symbol represents the R of 'restaurant', but it is also a clear visual signifier that emanates the commercialized branding of art institutions. The food served in the restaurant is supposed to mirror the narrative of the collection by being the "Rijksmuseum served up on a plate" ("RIJKS®"). With the Rijksmuseum's new curatorial model and new commercial activity departing from the historicity of their collection, these two practices are strongly entangled - both inside the museum and in the manner it presents itself to the outside world. The commercial nature of the museum's recent activity also has implications for their artistic activities in space, which will be discussed in the following chapters.

Contrary to the Rijksmuseum concerning itself with Dutch history, the aim of the Stedelijk Museum Amsterdam is to be an international institute dedicated to modern and contemporary art and design. The museum prides itself on not only collecting, but also stimulating and collaborating with various artistic talents, for instance by directly commissioning works ("Mission and Vision"). As a collection of contemporary art, the Stedelijk has the necessity to continuously acquire, adapt and re-evaluate its collection to fulfil its own mission statement of staying at the forefront of artistic developments. This is reflected in the curatorial model of the museum. The modern art section up until the 1950s is medium-specific as the various galleries display design objects separately from visual art objects. Furthermore, it is largely chronological, grouping works by art style and movement, whilst the galleries exhibiting art after the 1950s has no such formal grouping. These galleries are highly subject to change, with objects frequently forming new combinations of works and thus different collisions of meaning as they change from room to room. Alongside the inside of the building continuously changing, the outside has also been subject to a re-evaluation since the renovations plans were announced in the early 2000s and the new design was revealed in 2004 (Coevert). The large white addition to the building's architecture emphasizes the progressive and mobile identity of the museum and shows that the Stedelijk is not afraid to make bold interventions in space.

The youngest institution on Museumplein is the Van Gogh Museum, which opened in 1973. The museum defines itself as the place that makes "the life and work of Vincent van Gogh and the art of his time accessible to as many people as possible in order to enrich and inspire them" ("Mission and Strategy"). Accessibility is a key point in the rest of their strategy as the museum emphasizes an engagement with a wide public through stipulating that they reach "as many visitors and non-visitors as possible, creating dialogue with the public alongside a high degree of public involvement" ("Mission"). The most significant example of the Van Gogh Museum increasing its accessibility to cope with more visitors is the on-going work to create a new entrance, which will open in the second week of September 2015 ("Nieuw Entree"). Apart from the commercial aspect of upscaling the amount of visitors that are capable of passing through the doors, the new entrance also serves other functions. First of all, the Van Gogh has moved to be in line with the other institutions, which signifies a shift undertaken by all the cultural institutions on the square from closed off individual institutions to an open, conglomerate, and collaborative group, in line with each other and the municipality's desire ("Museumpleindebat"). Furthermore, this shift reshapes the nature of Museumplein from being the back garden of the museums to being the central heart off all the institutions. It also changes the way Museumplein is used, as visitors now have to cross the square to enter one of the museums. The extension onto Museumplein can be read as the Van Gogh Museum's most radical permanent intervention in space.

5. MUSEUMPLEIN - PERMANENT ARTISTIC AND CURATORIAL PRACTICES

5.1 PLAYFUL DISRUPTIONS

Perhaps the most visible semi-permanent structure on Museumplein is the 'Iamsterdam' sign. The three-dimensional red and white letters that stretch out for 23,5 meters have stood behind Andersson's lake since 2004 and are not set to move any time soon (Mac an Bhreithiún 255). The sign is a literal playful obstruction of daily movement and order, as tourists interact with the sign by using it as a climbing frame on which to have their picture taken (figure 2). However, the slogan 'Iamsterdam' is part of a controlling and commercialized discourse of homogenized city branding that sells Amsterdam to the outside world (Mac an Bhreithiún 258). The sign thus exemplifies that even a physical playful disruption in a hegemonically structured space by no means has to resist or undermine that discourse, and in this case can strongly assert it through its ideological underpinnings. Simultaneously, Museumplein houses art objects which function within the hegemonic structure and succeed in using the notion of playful disruption in order to undermine the commercialization of space. In the case of Museumplein these are museum initiatives.

Figure 2. Visitors interacting with the sign. Image courtesy of urbancapture.com

Directly behind the sign visitors are greeted with the Rijksmuseum and its gardens. Before the renovation, the gardens of the Rijksmuseum were part of the museum grounds, but after the renovation there was an important change. Despite the fact that the gardens are still subject to opening times, are lined by a fence physically separating them from the rest of the square and are maintained by the museum, there is one very significant difference. The gardens have been opened up to the public, now being free and accessible to enter without the need to purchase a museum ticket. The completely redesigned gardens opened to the public on the 22nd of June 2013, but were only fully completed in April 2015 (Starink). The landscape architects responsible for the remodelling called their creation an "open-air museum", built to accommodate the collections of sculptures and building fragments with the flowers and vegetation functioning as "a green stage for the art" ("Rijksmuseum"). Similarly, the current general-director of the Rijksmuseum, Wim Pijbes, pointed out that by making the gardens "semi-public" they have created an green outdoor gallery ("Museumpleindebat"). Pijbes is right to point out that the gardens stand in a literal and figurative transitioning position between the public nature of Museumplein and the semi-private institution that is the Rijksmuseum. Therefore, he accredits the capacity of the gardens to integrate the entire institution more into the sphere of open public space. As Pijbes told the NRC: "the Rijksmuseum should not just stand on the edge of the square, it should become part of the square. Through a different structuring of space we want to emphasize that the gardens of the Rijks are public. In other words, the public space of Museumplein should extend all the way to the Stadhouderskade" (qtd. in Smallenburg). Thus the gardens can be acknowledged as not only "semi-public", but as appropriated public space, that extend the parameters of the entire Museumplein beyond its previous border. The exhibitions and permanent exhibits staged in the gardens are therefore perhaps the most explicit example of curated public space visible on Museumplein.

The curatorial strategy of the Rijksmuseum's outside spaces departs from their strategy inside. The chronological historical contexts inside the rooms of the museum stand in stark contrast with the fusion of time periods and historical importance visible in the gardens (see for instance figure 3.1). This is not only the case for the objects placed in the gardens, but also for the garden styles themselves. A variety of classic garden styles have been juxtapositioned to create a nonlinear interaction between different time periods ("Rijksmuseum"). The garden features 'classic' outdoor art objects, including garden vases, sculptures, and busts (figure 3). These have generally been on display since before the renovation. Additions to the permanent collection since the reopening are art or design objects that can be directly used and interacted with. These include a PTT-telephone booth from 1933, a large chessboard and post-war playground equipment designed by Aldo van Eyck. Van Eyck considered play to be foundational for cultural-societal renewal, and believed that the hundreds of playgrounds that he created for the city of Amsterdam would help a generation overcome the trauma of war (Roode and Lefaivre 29). This playground equipment, including these two pieces that originally stood in Amsterdam Nieuw-West in the 60s, disrupted the strategic functionalism of the dominant modernist discourse, instead transforming structured space into a place of free imagination and creativity (Roode and Lefaivre 7; "Speeltoestellen" - figure 3). The curator of 20th century art at the museum, Ludo van Halem, is also interested in this idea and his acquisition of the objects can be seen to revive van Eyck's concept of subverting the dominant order through play (Schoenberger). An interactive water installation called Hide and See(k), created by contemporary artist Jeppe Hein in 2013, adheres to similar principles (Starink - figure 3). It is simultaneously an "integral part of the existing design" and a playful addition, disrupting the quiet garden space (J. Hein). The rising and falling water walls invite visitors to interact and traverse the object, incorporating the elements of play and surprise in order to question order and structure (J. Hein). The newly acquired 20th and 21st century playful use objects stand alongside building parts and ornamentation which have also been given interactive functions aside from their aesthetic value. An early 17th century city gate from Groningen marks the back entrance of the museum; an 18th century garden dome from Haarlem has been converted into an espresso bar; and a newly acquired French 19th century castle greenhouse grows vegetables for the RIJKS® restaurant (Starink - figure 3). These objects serve more overtly commercialized function rather than undermining that narrative through play, in particular the dome directly selling food and drinks. The juxtaposing of all these objects exemplifies the entanglement of artistic and commercial interests visible throughout the Rijksmuseum's practices in space.

Like the Rijksmuseum, the Stedelijk had a museum garden surrounding its grounds from the early 20th century (van der Werf 9). The garden had always housed a variety of sculptures, but since 1972 there

Figure 3. Rijksmuseum gardens 3.1 Bronze statue by L. Gasne (copy after Adriaen de Vries) Mercurius and Psyche, 1890 with Hein's installation behind
3.2 Child playing on Aldo van Eyck's climbing frame, 1962
3.3 19th century castle greenhouse filled with vegetables for RIJKS®

stood one so tall it could not be missed. It consists of three monumental steel sheets of almost twelve meters high, which lean on each other to create a triangular work called Sight Point (for Leo Castelli). Sight Point was not only Richard Serra's first of many vertical pieces, but it was also his first work created specifically to be in public space (Rutte). Furthermore, the director at that time, Edy de Wilde, asked him to create the work for a specific spot in the garden, so Serra decided to align the openings of the sculpture with the paths in order for Sight Point to connect various points with each other (Serra - figure 4). De Wilde was director at the museum between 1963 and 1985 and his decision can now be read as a typical choice of that period as throughout the 1970s and 1980s site-specific art was commissioned to challenge the dominant perception that art was universal, autonomous and above all a mobile good (Kwon 13). It was believed that site-specific works offered a counter-discursive narrative to the commodification of space and art, due to their intimate relationship with the space they occupied and the dialogue this forges between the viewer, site and space (ibid 13). Site-specific works are defined by the relationship they have with their direct surroundings, and can imbue their space with an inherent sacredness and symbolic meaning.

Figure 4. The Stedelijk Museum's garden designed by Hans Warnau, with Sight Point in its original position. Image courtesy of cascasde1987.nl

Figure 5. Richard Serra posing with the reinstalled Sight Point in front of the 'bathtub', 2012

However, the immediate site-specificity of the work has perished as the Stedelijk's garden disappeared in 1997 due to Andersson's restructuring. After being in a depot for 15 years, Sight Point returned to coincide with the reopening of the Stedelijk Museum in 2012. Moving the sculpture into the open fabric of Museumplein transforms its meaning, as it can no longer fulfil the original criteria of breaking through the commodification of space by being an immobile structure. Yet the Stedelijk's collection curator of paintings and sculpture, Bart Rutten, admits that despite these reservations, its return was immediately incorporated into the extension plans for the museum. Whilst these plans were being made, Serra was asked whether he would like Sight Point installed again, and if so where, so that the foundations could be fortified in order to carry the 51 ton steel construction (Rutten). Sight Point now stands by the northern corner of the newly shaped square space that the 'bathtub' extension created. Here it functions as a beacon for the museum - not only due to its sheer volume in space, which frames the large new wing, but also due to the stark contrast created by offsetting the rusty orange with the white of the new facade (figure 5). This makes it an attention grabbing reference point for any visitor to the square. The reinstallment can thus be read as a commercial move, drawing attention and thereby visitors towards the Stedelijk. Whilst overseeing the reinstallment, Serra himself is dissenting of this reading when talking about the new location of Sight Point:

> I like the intervention in urban life; I do not like to make art rarefied. I think what museums do is pull pieces out of a living situation and put them in a situation where they are framed by an ideology that says 'aesthetics'. This piece mediates between what we call normal life and aesthetic life, and the way this piece is placed accomplishes that task. (Serra qtd. in Rutten)

Specifically by being moved into open public space, Serra accredits his own work with the capacity to demystify art from the ideological pedestal surrounding museum curation, thereby going far beyond mere eye-catching commercialization of space. He therefore calls this new location better than the original spot (Serra qtd. in Stigter). Despite the loss of site-specificity, the sculpture is still intentionally ascribed with the inherent potential to playfully intervene in the life of unsuspecting passers-by, perhaps even more so now it is in open urban space rather than a secluded garden. Sight Point transforms a normal commute into an aesthetic experience freed from mediation by the walls of an authoritative institution. Its urban location and scale result in the artwork functioning as a fruitful agonistic intervention.

5.2 AUTHORITATIVE CONSOLIDATION

Of all the artistic interventions on Museumplein the longest standing objects are those commissioned or placed there directly by the municipality.

Figure 6. Vrouwen van Ravensbrück, 1975

Figure 7. Hel en Vuur, 1978

Two Second World War memorials have stood on Museumplein ever since the mid-70s. The first is Vrouwen van Ravensbrück, which commemorates those that did not return from the Ravensbrück concentration camp and was revealed in 1975 (figure 6). The second, Hel en Vuur was unveiled in 1978 and commemorates another "forgotten group" of the war, the gypsies (Levano – figure 7). By memorializing, these objects narrate an authoritatively selected history. However, according to Mouffe, artworks that explore "marginality, oppression or victimization" are in a position to critically challenge dominant order (13). Thus, by calling attention to specific social minorities, these objects alter the politics of representation. They are literal manifestations of authoritative structure, by being placed in space by a powerful authority, but do not function to reiterate hegemonic discourse. Just like the 'Iamsterdam' sign is a playful disruption, but not a counter-discursive one, these artworks are authoritative but in a manner that is critical of dominant perceptions.

As the Stedelijk Museum has placed Sight Point in the middle of the new space their extension created, the Van Gogh Museum similarly have a permanent structural intervention which extends into the heart of the square. In this case it is the extension itself. In the late 1990s Kisho Kurokawa designed the original extension that functions as an exhibition wing. It is this building that is now the basis for the further extension forming the new entrance (figure 8). Kurokawa originally planned to create a round building, but Andersson convinced him that the realized semi-ellipse would function better within the rest of the space (van der Werf 81). The current building work will reshape the ellipse back into a full circle (figure 8). The extension completed in 1999 aesthetically transformed Museumplein, and this process is being reinvented in 2015. But the exhibition wing is more than an architectural and functional feature of the Van Gogh Museum, it serves as an artistic intervention into the public space of Museumplein. As figure 1 shows, on the street level of Museumplein Kurokawa's extension looks completely detached from the main museum building and stands within the grassy area

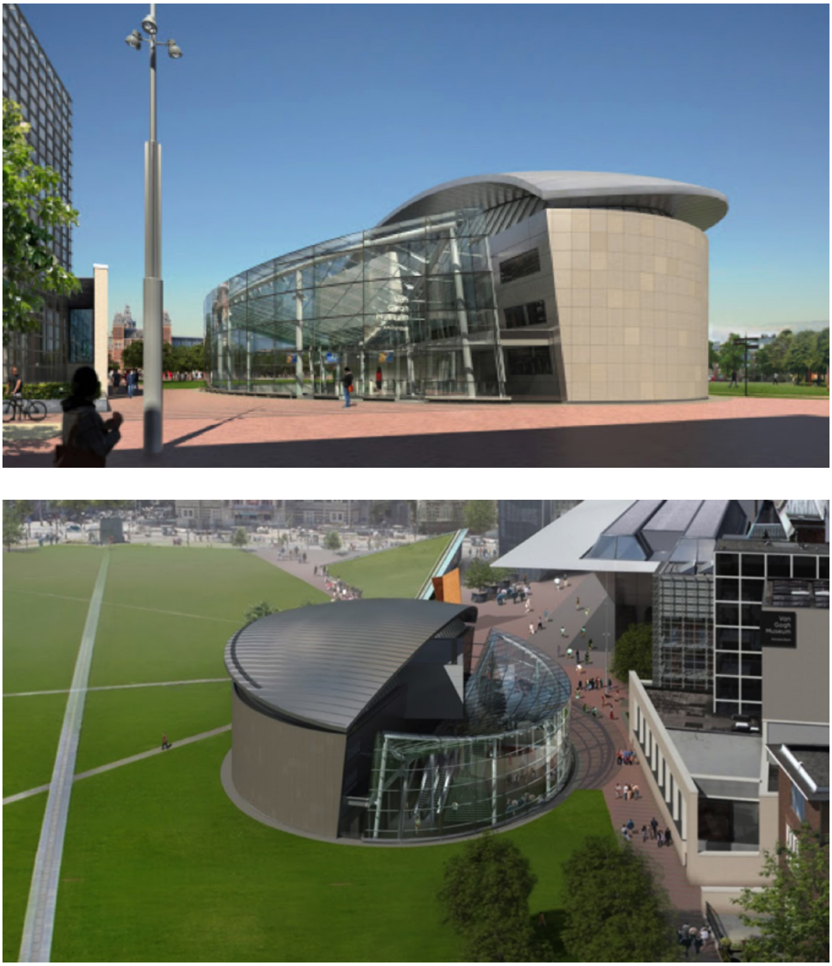


Figure 8. Renders of finished entrance Van Gogh Museum, 2015. Images courtesy of Vangoghmuseum.nl.

- which signifies the heart of the public open space. It is therefore not part of the built up parameters of Museumplein, but a substantial segment of the fabric of the square. Kurokawa's design is a very permanent artistic and aesthetic contribution, freed from any mediated content or museum branding. This will change with the addition of the new entrance. The extension will change the function of the structure from solely housing art objects to accommodating increased visitor numbers, thus becoming a commodified space. Furthermore, the new glass facade will be branded with the museum's name and logo, subverting the aesthetic of the original building and outwardly branding that space as a part of the museum territory. Despite the other museums also having relocated their entrance, the move of the Van Gogh Museum is a stronger and more overt consolidation of its presence on Museumplein, as it reappropriates an aesthetic contribution on the heart of the square to increase commercial interests.

Figure 9. KAWS, Along the Way, 2015

6. MUSEUMPLEIN - TEMPORARY ARTISTIC AND CURATORIAL PRACTICES

6.1 PLAYFUL DISRUPTIONS

Since its conception, Museumplein has been the city's prime location for big temporary events (van der Werf 8). This means there has been a constant stream of projects that disrupt daily order on the square, including festivals, fairs, exhibitions, and installations. All these events add a new layer of meaning to space, but the actual result and function of these curated interventions are highly dependent on who the initiator and stakeholders are. Critical or light-spirited subversions of hegemonic structure are rare, as temporary interventions are relatively expensive and hence highly reliant on neoliberal finance structures, thereby implicitly reiterating these values ("Museumpleindebat"). However, even initiatives embedded within this hegemonic logic can offer a counter-discursive gaze, once again using elements of play to achieve this. Such an example can be found in the center of Museumplein, where two 5,5 meter tall wooden sculptures stand on the water of Andersson's lake between May and September 2015 (figure 9) (Davis).

The pieces, called Along the Way, were created by KAWS who is a contemporary artist and designer known for his toy-like figures (Davis). The figures are large scale versions of the type of limited edition 'toys' KAWS made his name with, thus on surface level the sculptures reference playfulness in a very direct way. Furthermore, these works act against the hegemony by destabilizing Museumplein's status of being the cultural centre of the Netherlands. These works form a series together with two other KAWS pieces and a whole range of other sculptures dispersed along a 2,5 kilometre route through residential neighbourhoods in Amsterdam Zuid (Davis). By exhibiting big international artists in these diverse settings equal importance is given to the square and to quiet streets (Fuchs). ArtZuid, the name is the sculpture route, transforms daily environments into provoking sights of exhibition (Fuchs). Similarly, institutional practice on Museumplein can restructure the square into a sight of playful display.

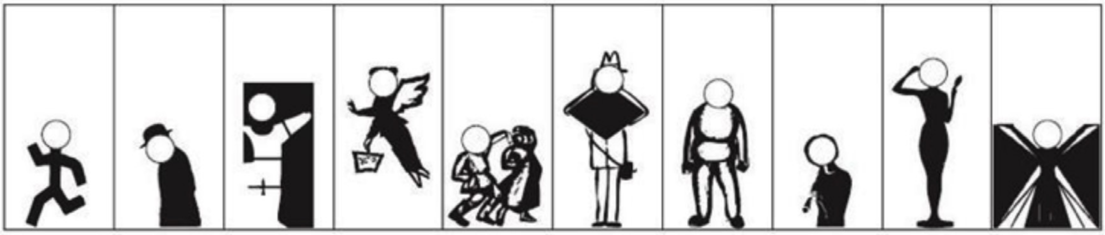
In keeping with their mission statement to remain contemporary and innovative, the Stedelijk Museum is the most active museum of the three when it comes to artistic interventions in the central space. Sometimes, the disruptive nature of these interventions is not directly visible. In 2010, for example, the Stedelijk Museum worked together with the MediaLAB and Your Own Reality to create an exhibition of 3D virtual installations on Museumplein that were only visible through the mobile application Layar ("MediaLAB"). As the Stedelijk Museum commissioned, selected and 'placed' the installations, they directly curated urban public space in a manner that suits their desire to be progressive. However, sometimes these projects have a massive impact on the experiential feel of Museumplein. This was the case during the Holland Festival, that happened between the 30th of May and the 23rd of June 2015. In this period an enormous installation diagonally traversed the square, running in front of the Stedelijk Museum towards the Rijksmuseum ("Stedelijk"). The Stedelijk Museum and the Holland Festival commissioned the British artist Liam Gillick to create the public work titled All-Intimate-Act. Characteristics of Gillick's work are its public nature, and constant redefinition of space and meaning based on how the public interact with the work ("Stedelijk"). This piece is no different, as it forms a very physical disruption, blocking visitors from entering Museumplein from the northwestern side. Furthermore, it actively invites participation. The installation consists of a long row of head-in-the-hole panels featuring adaptations of various figures from early 20th century artists, including Oskar Schlemmer and Kazimir Malevich, as well as incorporating popular culture and costume design between the 1910s and 1940s drawn from the graphic collection at the Stedelijk Museum ("Stedelijk" - figure 10).

Due to the faces being cut out, any visitor to Museumplein is invited to participate by sticking their head in the holes in order to have their picture taken. As the installation is turned away from the museum, these pictures will include the museum building in the background. According to the Stedelijk Museum itself, visitors to the square "physically engage" with its history through the mode of exhibition, expressing that "anyone who sticks their head through a hole becomes part of a work from the Stedelijk collection" ("Stedelijk").

The invitation to interact and take photos with a three dimensional installation on Museumplein draws a clear parallel to the 'Iamsterdam' sign. The constant reframing of the museum as a backdrop on user's photos is therefore suggestive of a similar tactic of commercialized museum branding. However, no such commodification of space is to be found with All-Intimate-Act itself. Instead, it has a strong tie to the performing and visual arts, for instance through the title being derived from the Greek word for 'pantomime' and the figures directly relating to art objects ("Stedelijk"). This installation utilizes similar ideas of inviting playful interaction through the occupation of space, but without the underlying commercial narrative, and thus achieves an entirely different effect. The selection of this sight and object means this curation of space creates an open participatory relationship between visitors to Museumplein, the Stedelijk Museum's collection and an art object, which together act as a counter-narrative.

The municipal wish for Museumplein to foster stronger relationships between the cultural institutions on the square can perhaps seen to be partly embodied by the collaboration between the Stedelijk Museum and Holland Festival. However, the municipality wants a strong, self-contained "necklace" created by the institutions on the square and Holland Festival is an external force connecting

Figure 10. Liam Gillick, All-Intimate-Act (fragment), 2015.
Digital reproduction courtesy of Stedelijk.nl.



various sights of culture throughout the city with each other (ter Berg et al. 23). The Van Gogh Museum is potentially even more conforming to this desire as the Art Department of the museum has stressed that the Van Gogh is keen to participate in artistic projects or events in the future, but only if they can do this in collaboration with the neighbouring museums (Blok). Currently the museum does not have any art objects displayed in public space, so there is no active performance of curatorial practices. Thus, despite the Van Gogh Museum stressing public involvement and outreach to non-visitors as an integral part of their mission statement, their involvement with the public on Museumplein has not materialized further than Kurokawa's structure. The Art Department justifies this by stating that objects in their collection are not suitable for outside display (Blok). However, as the various departments of the three museums regularly meet, future artistic collaboration on Museumplein will be discussed once the building work has been completed (Blok). Confirming this idea, the current managing director Axel Rüger stressed that he is eager to further connect the museums with each other and Museumplein through short-term collaborative artistic interventions ("Museumpleindebat"). This desire sounds to be completely in line with the municipality's concept of transitioning from individual "pearls to one necklace", thereby strengthening the underlying commercialization of public space (ter Berg et al. 23). Whether this will be the case remains to be seen, as the Stedelijk Museum has already shown the capacity to stage artistic interventions that function within this logic but still playfully undermine ideals of hegemonically structured public space.

6.2 AUTHORITATIVE CONSOLIDATION

Museumplein hosts artistic interventions of all shapes, sizes, and nature. Interventions disrupt and alter space, and may at times come across as playful subversions of daily structure, but in fact serve to further assert commercial and authoritative control. Annual festivals, such as the art fair 'Art Square' or the festive presentation of the theater season that is the 'Uitmarkt', have an outwardly commercial character as they are staged with the intention to sell art and culture. Art Square actively contributes to the institutionalization of Museumplein as it brands itself as a place to "meet the living Dutch Masters in a temporary museum", whereas really it functions as a gallery space (van der Lichte). This is in line with the municipality, who has been giving out numerous event licenses that contribute to their desired 'institutionalization' or 'museumization' of Museumplein. In May 2015 alone there were three concurrent temporary installations and exhibitions. The longest running of these is the toy-like KAWS sculptures. A similar reference to toys and play can be seen in the series of sculptures called Miffy Art Parade. Yet, despite Miffy being the fictional character from children's books, the playful connotation has been lost to a reimagining of the figure that aids the institutionalization and commercialization of space. To celebrate the 60th anniversary of Miffy, sixty sculptures of the female rabbit were created, each one sporting a design by a different artist. Twelve of these sculptures form a row between the Rijksmuseum and Van Gogh Museum, one stands in the gardens of the Rijksmuseum, another has been placed inside the Philips Wing of the same museum, the Van Gogh Museum has one, and finally, two stand in the museum shop of the Stedelijk ("Nijntje"). It is telling that the Stedelijk Museum decided to display the sculptures in their shop, as the Miffy exhibition is a highly commercialized enterprise. Not only will all the sculptures be auctioned off once the display period is over, there is even a web-shop selling small replicas and numerous paraphernalia of the

Figure 11. Miffy Art Parade, the front six sculptures of the row on Museumplein. The front object is Needje by Florentijn Hofman.

designs ("Nijntje"). The sculptures use Museumplein as a 'space of exhibition', adding meaning through aesthetic value and directly linking inside and outside museum space. Yet, this institutionalization of space mirrors the municipal use of the term, where curation and exhibition are used as masks for underlying commercial gain and commodification.

The last of the three May exhibits on Museumplein was a light installation by the Dutch artist Daan Roosegaarde called Waterlicht (figure 12). Between 22:00 and 00:00 on the 11th until the 13th of May Museumplein was virtually flooded with blue light waves, signifying the height of the water level if there was no human regulation and sea defence. Once again, despite being staged in the public space of the square, there was museum involvement. ING - a main sponsor of the Rijksmuseum - also financed Waterlicht, and did so in order to commemorate that the museum purchased a large canvas by 17th century artist Jan Asselijn in March 2015. The supposed link between these two works is their shared subject matter of human interaction with technological innovation and water (Smulders). The canvas, called The Breach of the St. Anthony’s Dike in Amsterdam, was also acquired with a financial contribution from ING (Smulders). As Waterlicht received support from the Rijksmuseum, it can be seen as a contemporary and outward visualization or assertion of the Dutch historic narrative that is so important inside the museum. Furthermore, the high involvement of a commercial partner, both in the acquisition of objects and in the spatial intervention, is a manifestation of hegemonic economic structures that strengthen the production of a commodified space.

The Rijksmuseum's intimate connection with the space surrounding it has been evident ever since Cuypers' interference in the shaping of Museumplein. The consistent incorporation of its surroundings during special events and short-term projects is a conscious attempt to strengthen its ties with the city and assert its own authority and the value of its collection. In the past this has happened particularly through the Golden Age narrative and legacy of Rembrandt that is strongly present in the Dutch historical context (de Waard 143). Manifestations of this include the ongoing project 'Walking with Rembrandt', where visitors are invited to explore spots in Amsterdam that relate to Rembrandt's life through a self-guided tour and the 'Rembrandt Year' held in 2006, which included the Rijksmuseum co-producing

Figure 12. Daan Roosgaarde, Waterlicht, 11 May 2015. Image courtesy of StudioRoosgaarde.net

a film about the Night Watch (de Waard 148). In the years since the re-opening, the gardens have become an ideal sight to propagate the museum's importance, being presented as an outdoor gallery staging annual exhibitions of celebrated international artists. Each summer the permanent collection shares the gardens with temporary sculpture exhibitions. 2013 saw twelve monumental works by Henry Moore arrive, in 2014 the largest amount of Alexander Calder's modernist sculpture was brought together in an outside exhibition and in the summer of 2015 the garden will host a series of sculptures by Joan Miró, on loan from museum and private collections from all over the world ("Rijksmuseumtuinen"). Whereas the fixed collection in the garden display a strong use value, the temporary exhibits occupy space solely to be viewed. This does however not mean that they solely serve to aestheticize the gardens and imbue the space with inherent artistic value. The artists are all big names and thus the display of their objects serves as an outdoor 'blockbuster' exhibition, effectively attracting more attention and visitors to the museum. Therefore, despite adding aesthetic value to the playful objects permanently in the garden, the installation of these exhibitions equally serves as a calculated commercial act. All these various activities seem to suggest a clear outcome: the Rijksmuseum is the most active institution of the three in mediating space for its own benefit.

7. CONCLUSION

Museumplein is not a static space, its sense of place is in continuous flux, being made and then remade by its users and stakeholders (de Certeau qtd. in Cresswell 39). These stakeholders include three of the most significant art institutions of the Netherlands. The Rijksmuseum, Stedelijk Museum Amsterdam, and Van Gogh Museum all make interventions on the square, which contribute to a larger urban phenomenon that sees city spaces mirroring exhibition spaces by being inscribed "by the same textual and visual markers" (de Waard 146). For this reason, this research suggests to analyse these interventions as acts of curation. The social, political and cultural factors shaping urban public spaces in neoliberal societies can be accessed in terms of production. This neo-Marxist framework sees space as a hegemonic structure, within which art can function to assert or subvert the dominant order. Viewed in this manner, Museumplein becomes an agonistic arena where the battle tied to the politics of representation is openly fought. To curate space is to actively shape artistic interventions, both in terms of permanent and temporary contributions. Thus the vocabulary of curation is capable of making the effects of museum practice in space visible, as well as assessing whether an art object in a public and agonistic setting consolidates the dominant symbolic order or undermines it.

On Museumplein the concept of play, or playful disruption, through art is utilized as a form of critique that can undo or lead to the questioning of structure. However, physical playful disruptions of space can still propagate an authoritarian discourse, as is the case with the 'Iamsterdam' sign. This object, however, has an underlying municipal policy of commercialized city branding. Perhaps paradoxically then, it is not the museums that are the biggest perpetrators of the institutionalization of space, but the City of Amsterdam and the district office Amsterdam Zuid. By 2020 they want Museumplein to be reshaped into an 'institution' which will be capable of accommodating the mass consumption of culture. The desire to institutionalize space will result in the increased commodification of Museumplein, through stimulating privatized commerce and moving the square towards a more closed, self-contained space. In this case increased commodification means an increased assertion of hegemonic structure. The museums partly support this model by promising more inter-institutional collaboration and through starting their own commercial enterprises. However, as these are individual museums rather than a conglomerate space, they are also moving in an opposite direction by transforming themselves from self-containing units to increasingly porous entities. Analysing the municipality's and institutions' activities through a lens of curation helps differentiate between these various types of 'museum practice' happening in public space, and shows that the two 'movements' meet in the middle on Museumplein. This opening up of visual art institutions mirrors the broadening out of the term 'curation', which sees its definition shifting from a single voice of authority towards a pluralistic understanding within which a greater public can give meaning to places. By opening up to this broad definition of curation, which allows for the curation of city space, institutions can simultaneously adapt themselves to the contemporary museological question of how to open up to a more diverse public.

The museums do not contribute to the undoing or strengthening of hegemonic structure equally. The Van Gogh Museum claims that their collection and narrative does not allow for them to make artistic contributions in public space. Instead, their architectural and aesthetic contribution to Museumplein is currently being rebuilt to increase commercial gain, leading to an authoritative and territorial claim in the square's open space. Similarly, the Rijksmuseum displays an increased entanglement of commercial and artistic practices inside the institution, which have also been transported to many of its outside activities. The newly acquired modern and contemporary objects in the garden form an exception, as these use public interaction and playfulness to counteract order and structure. The Stedelijk Museum's mission to be a progressive and innovative museum also makes it the most successful of the three at staging counter-hegemonic interventions on Museumplein. Already in the 1970s, with the commission of Sight Point, the museum showed it was consciously breaking through the commercialization of space. Sight Point continues to intervene in daily life now it has returned to a new urban context. By inviting Liam Gillick to create All-Intimate-Act, the Stedelijk demonstrates that of the three museums it is still the leading force in supporting and creating counter-narratives that ensure established structure and the commodification of space is disrupted.

The recent and future structural changes and artistic interventions on Museumplein have far-reaching implications. This period of change could soon be viewed as one of the biggest turning points in the history of Museumplein. Exactly how these developments will be remembered is still an ongoing question. The Rijksmuseum gardens have only just reached their full potential as construction was only completed in April 2015, the Van Gogh Museum is still in the process of building its new entrance, the precise effects of the municipality's 'vision' for Museumplein are unclear, and the exact type of future collaboration between the institutions is still up for discussion. It seems that increased commercialization and mass consumption of culture is imminent; the forecast is that by 2025 the annual visitor rate to Museumplein will quadruple to almost 20 million ("Museumpleindebat"). However, the strengthened assertion of a hegemonically structured space is not likely to result in less artistic practices playfully disrupting order. Public space is fascinating precisely because it is characterized by its unstoppable conversation, staging a dialogue between authority and critical counter voices, and the future for Museumplein should be no different.

LIST OF ILLUSTRATIONS

Figure 1: Aerial View of Museumplein. 2012. Stedelijk Museum. Web. 12 April 2015.

Figure 2: Vinckens, Roel. Iamsterdam Letters. 28 June 2014. Urban Capture. Web. 27 May 2015.

Figure 3: Cowling, Lucy. Various Images of the Rijksmuseum Gardens. 15 May 2015. Photographs.

Figure 4: Aerial View of Stedelijk Museum's Former Garden. n.d. Tuin Historisch Genootschap Cascade. 4 Feb 2007. Web. 17 May 2015.

Figure 5: Sculptor Serra with his sculpture 'Sight Point' on Museumplein Amsterdam. 14 Sep. 2012. NRC Archief. Web. 27 May 2015.

Figure 6: Cowling, Lucy. Vrouwen van Ravensbrück. 15 May 2015. Photograph.

Figure 7: Cowling. Lucy. Hel en Vuur. 15 May 2015. Photograph.

Figure 8: Kurokawa, Kisho and Associates. Renders of Finished Entrance Van Gogh Museum. 2015. Web. 14 Feb. 2015.

Figure 9: Cowling, Lucy. KAWS 'Along the Way'. 15 May 2015. Photograph.

Figure 10: Gillick, Liam. All-Intimate-Act (fragment). 2015. Stedelijk Museum and Holland Festival. Stedelijk Museum. Web. 13 April 2015.

Figure 11: Cowling, Lucy. Miffy Art Parade Museumplein. 15 May 2015. Photograph.

Figure 12: Thomas Doornbos. Waterlicht. 11 May 2015. Studio Roosgaarde. Web. 13 May 2015.

WORKS CITED

"About RIJKS®." Rijks Restaurant. 2014. Web. 19 May 2015.

Blok, Martine. "Van Gogh on Museumplein." 20 April 2015. Personal Correspondence.

de Certeau, Michel. The Practice of Everyday Life. 1980. Trans. Steven Rendall. University of California Press, Berkeley, 1984. Print.

Chaplin, Sarah and Alexandra Stara. Curating Architecture and the City. Abingdon: Routledge, 2009. Print.

Coevert, Annemarie. "Het Nieuwe Stedelijk." NRC Archief. 22 Sep 2012. Web. 16 May 2015.

Cresswell, Tim. Place: A Short Introduction. Oxford: Blackwell, 2004.

Davis, Josh. "KAWS 'Along The Way' @ ArtZuid 2015". HypeBeast. 13 May 2015. Web. 16 May 2015.

Deutsche, Rosalyn. "Art and public space: Questions of democracy." Social Text (1992): 34-53. Print.

Deutsche, Rosalyn. Evictions: Art and Spatial Politics. London: MIT Press, 1998. Print.

Foucault, Michel. "Of Other Spaces: Utopias and Heterotopias." Trans. Jay Miskoweic. Architecture / Mouvement/ Continuité (1984): 2 - 9. PDF.

Fuchs, Rudi. "ArtZuid 2015." ArtZuid. 2015. Web. 20 May 2015.

Gough, Tim. "Cura." Curating Architecture and the City. Ed. Sarah Chaplin and Alexandra Stara. Abingdon: Routledge, 2009. 93 - 102. Print.

Habermas, Jürgen. The Structural Transformation of the Public Sphere: An Inquiry into a Category of Bourgeois Society. 1962. Trans. Thomas Buger. Cambridge, MA: MIT Press, 1991. Print.

Hein, Hilde. "What Is Public Art?: Time, Place, and Meaning." Journal of Aesthetics and Art Criticism (1996): 1-7.

Hein, Jeppe. "Hide and See(k)." Jeppe Hein. 2013. Web. 16 May 2015.

Hooper-Greenhill, Eilean. Museums and the Shaping of Knowledge. London: Routledge, 1992. Print.

Huyssen, Andreas. Twilight Memories: Marking Time in A Culture Of Amnesia. London: Routledge, 1995. Print.

Karsten, Lia & Albertine van Diepen. "Van Wie Is Het Amsterdamse Museumplein? Bewoners En Toeristen in de Stedelijke Openbare Ruimte." Ruimte & Maatschappij 1.2 (2009): 5 – 26. PDF file.

Kwon, Miwon. One Place After Another: Site-Specific Art and Local Identity. Cambridge, MA: MIT Press, 2002. Print.

Lefebvre, Henri. The Production of Space. 1974. Trans. Donald Nicholson-Smith. Oxford: Blackwell, 1991. Print.

Lefort, Claude. "The Question of Democracy." Democracy and Political Theory. Minneapolis: University of Minnesota Press, 1988. Print.

Levano, Helen. "Hel en Vuur." Buiten Beeld In Beeld. 2012. Web. 16 May 2015.

Mac an Breithiún, Bharain. "Graphic Design, Globalization, and Placemaking in the Neighbourhoods of Amsterdam." Imagining Global Amsterdam: History, Culture, and Geography in a World City. Ed. Marco de Waard. Amsterdam: Amsterdam UP, 2012. 255 - 270. Print.

"MediaLAB Amsterdam Stedelijk Museum." Stichting Genen. 2010. Web. 7 May 2015.

"Mission and Strategy." Van Gogh Museum. 2014. Web. 6 May 2015.

"Mission and Vision." Stedelijk. n.d. Web. 30 April 2015.

Mouffe, Chantal. "Art as an Agnostic Intervention in Public Space." Art and Democracy 14.1 (2008): 6 - 15. PDF file.

"Nieuw Entree." Van Gogh Museum. 9 Feb. 2015. Web. 14 Feb. 2015.

"Nijntje Art Parade." Nijntje Art Parade. 2015. Web. 16 May 2015.mPhillips, Patricia. "Temporality and Public Art." Art Journal 48:4 (1989): 331 - 335. PDF File.

Pijbes, Wim. "The Renovation." Rijksmuseum. 2013. Web. 24 April 2015.

Pijbes, Wim, Beatrix Ruf and Axel Rüger. "Het Grote Museumpleindebat."
Volkskrant op Zondag. De Rode Hoed, Amsterdam, 12 April 2015. Public Round Table Debate. 12 April 2015.

Rancière, Jacques. The Politics Of Aesthetics: The Distribution of the Sensible. 2000. Trans. Gabriel Rockhill. London: Continuum, 2006. Print.

Rancière, Jacques. Dissensus: On Politics and Aesthetics. Trans. and Ed. Steven
Corcoran. London: Continuum, 2010. Print.

"Rijksmuseum." Copijn Tuin- en Landschaps-architecten. 2013. Web. 19 Feb. 2015.

"Rijksmuseumtuinen Voltooid." Rijksmuseum. 17 April 2015. Web. 1 May 2015.

de Roode, Ingeborg and Liane Lefaivre. Aldo van Eyck: The Playgrounds and The City. Rotterdam: NAi Publishing, 2002. Print.

Rutten, Bart. "Serra Back on the Museumplein." ArtTube. 31 Oct 2012. Web Video. 14 April 2015.

Serra, Richard. "Klaus Ottoman Interview." Journal of Contemporary Art (1989). PDF File.

Schoenberger, Janna. "Constant, New Babylon and the Playful Homo Ludens."
Urban Utopias. Amsterdam University College, Amsterdam. 2 April 2015. Lecture.

Smallenburg, Sandra. "Museumplein Word Omgebouwd Naar Voorbeeld Berlijn." NRC Archief. 31 July 2009. Web. 6 March 2015.

Smulders, Marc. "Kunstenaar Daan Roosegaarde Zet Museumplein Onder Water." ING Bank. 11 May 2015. Web. 16 May 2015.

"Speeltoestellen Aldo van Eyck." Rijksmuseum. 2012. Web. 15 May 2015.

Starink, Petra. "New Look Gardens Rijksmuseum." Architectuur. 20 April 2015. Web. 1 May 2015.

"Stedelijk Museum en Holland Festival Presenteren All-Intimate-Act." Stedelijk. April 2015. Web. 13 April 2015.

Stigter, Bianca. "Serra Weer Terug bij Stedelijk." NRC Archief. 14 Sep. 2012. Web. 27 May 2015.

Storrie, Calum. The Delirious Museum: A Journey from the Louvre to Las Vegas. London: I.B. Tauris. 2006. Print.

ter Berg, Dick, Hans Karssenberg, Anouk Montagne, Marianne Griffioen, Charlotte Hilbrand, Saskia Blaas, Rutger Veldhuizen van Zanten, Remco Bakker, Claire Kuijs en Alex Schoep. Visie Museumkwartier: Vormgeven aan het Culturele Hart van Amsterdam. Amsterdam: Gemeente Amsterdam Project Management Bureau, 2008. PDF file.

van Lent, Daan and Pieter van Os. "Musea Doen Het Goed: Aantal Bezoekers in 2013 Fors Gestegen." NRC Handelsblad. 27 Dec 2013. Print.

van der Lichte, Peter. "Temporary Museum." Art Square Amsterdam. 2015. Web. 16 May 2015.

van der Werf, Jouke. Plein, Park of Veld? Cultuurhistorische Verkenning Museumplein en Omgeving. Amsterdam: Gemeente Amsterdam Bureau Monumenten & Archeologie, 2008. PDF file.
"Vision and Mission". Rijksmuseum. n.d. Web. 24 April 2015.

de Waard, Marco. "Rembrandt on Screen: Art Cinema, Cultural Heritage, and the Museumization of Urban Space." Imagining Global Amsterdam: History, Culture, and Geography in a World City. Ed. Marco de Waard. Amsterdam: Amsterdam UP, 2012. 143-67. Print.

Wong, Tammy. The City as a Curated Space: A Study of the Public Urban Visual Arts in Central Sydney and Central Melbourne, Australia. Melbourne: RMIT University Press, 2011. PDF File.

Unless otherwise stated, all works in Dutch have been interpreted or translated by the author.

School Attendance of the Indian Adolescent

Uncovering the socio-economic determinants of demand for education

Jan-Paul R. Wiringa

ABSTRACT

While secondary schooling is essential to break the intergenerational transmission of poverty, structurally unequal access to secondary education in India prevents many adolescents from climbing the socio-economic ladder. Through a logistic regression (LR) of National Family Health Survey 2005–06 data (NFHS-3), this research attempts to uncover the socio-economic indicators and determinants of school attendance of children aged 15 – 17 in India (N = 13,973). With high economic returns on Indian secondary education, adolescent demand for education should be high. However, the LR results imply that Adolescent School Attendance (ASA) is hindered by a combination of household factors, including conservatism about gender roles, residence, religion, parental education, household size and wealth constraints.

The results indicate a strong effect of household wealth and parental education on ASA; a particular positive effect of educated mothers on the attendance of her daughter(s) is noted. Strong effects of the Islam religion on ASA allude to Muslim discrimination and small returns to Muslim education. Concurrently, conservative gender-biased values within Hinduism and Islam are shown to lead to female adolescent absenteeism. This study suggests that female ASA attendance suffers under increasing household size and rural residence. Additionally, controlling for wealth and parental education, caste or tribe status did not significantly affect adolescent participation. These unexpected results should be explored in the light of recent interventions that promote educational opportunities of children from backward communities.

Based on this study's findings, this research recommends an increase of Indian ASA through public information campaigns to transform attitudes about education and delayed marriage, financial assistance, and investments in educational reforms.

1. INTRODUCTION

India's recent economic achievements have bred great optimism about its general socio-economic development. However, the story of India's educational development is one of mixed success; yet the country is home to 17% of the world's population, but 37% of the world's illiterates, and a high proportion of the world's out-of-school children (Burke, 2011; Indo-Asian News Service, 2014). Admittedly, it has recently made promising progress in increasing participation in elementary education (Kingdon, 2007).

Elementary education is undoubtedly crucial, yet it is often inadequate in empowering young workers to elevate themselves lastingly out of poverty. Secondary schooling is essential to break the "intergenerational transmission of poverty" (World Bank, 2009, p. 10). However, structurally unequal access to secondary education in India prevents many adolescents from climbing up the socio-economic ladder. As free primary education is guaranteed until the age of 14, the transition to fee-based secondary school becomes particularly problematic. In the 2009 secondary enrollment rates, the World Bank found a 10% gap between boys and girls, a 40% gap between students from the highest and lowest expenditure quintile groups and a 20% gap between students living in urban and rural areas (2009). Additionally, the enrollment rates of scheduled tribes (STs), scheduled castes (SCs) and Muslims was far below their share in the entire Indian population (World Bank, 2009).

In urban India, the gender gap in school attendance seems trivial, but in rural areas, the disparity between the attendance of males and females widens and increases strongly with age (IIPS and Macro International, 2007a); to be noted in tables 1 and 2. Surprisingly, the interaction effects of gender with urban residence, religion, caste/tribe status, and household size have been considered only at a theoretical level, supported by analyses of descriptive statistics. Therefore, in an effort to add to the current body of research on Indian Adolescent School Attendance (ASA), this research aims to shed light on the aforementioned interaction effects.

Through a secondary data analysis of the National Family Health Survey 2005–06 (NFHS-3), this research attempts to uncover the socio-economic indicators and determinants of school attendance of children aged 15 – 17 in India. Recent governmental programs have focused on improving the supply-side issues associated with adolescent participation in schooling, among which are improvements in school quality, infrastructure, and accessibility (Jayachandran, 2002). However, an important part of the cause of inequality in attendance remains unexplained by institutional justifications (Kingdon, 2005). It is rather explained by various socio-economic factors that are postulated to constrain the demand for education, including biased gender roles, rural residence, religion, caste status, parental education, household size and credit-constraints. While this research limits its scope to explaining demand-side variables associated with adolescent school attendance, there exists an opportunity for future research to explain the variation in school attendance caused by supply-side issues.

2. LITERATURE REVIEW

2.1 EDUCATIONAL DEVELOPMENTS IN POST-INDEPENDENCE INDIA

When India gained independence in 1947, it inherited a postcolonial legacy of widespread illiteracy with no solid educational system in place (Kingdon, 2007). When the first post-independence census was conducted in 1951, only 9% of women and 27% of men were literate (Registrar General, 2011b). Nonetheless, the crude literacy rates, whereby the literate population is taken as a proportion of the total, show a consistent increase both for males and females from 1951 onwards, with a slowly decreasing gender gap; while in 1951, only 25% of Indian males and 7% of females could read and write, 60 years later, 71% of males and 57% of females were literate with a general crude literacy increase from 17% to 64% (Registrar General, 2011b). While literacy has increased substantially over time, with recent cohorts displaying higher literacy than older cohorts; the National Family Health Survey-3 shows that of those aged 15-19, still only 74% of women and 89% of men are literate (IIPS and Macro International, 2007a). It has been estimated that, given its current rate of development, India will not be able to attain total literacy before 2060 (Pathak & Gupta, 2013).

To combat the problems associated with illiteracy globally, the international legal sphere has progressively recognized equal access to education as a basic duty of a mature civil society (Sen, Shariff & Vanneman, 2010). The Millennium Development Goals two and three planned to warrant universal primary education for children everywhere, and to eliminate gender disparity in primary and secondary education (United Nations, 2000; United Nations, n.d.-a; United Nations n.d.-b).

Furthermore, the spread of higher education is not only an ideological objective; it is generally acknowledged to be vital for the growth of the Indian economy and reduction of poverty (Desai et al., 2010; World Bank, 2009). Since 2000, most of the economic and employment growth in India has taken place in skilled services which minimally require a secondary education degree; these industries are now finding themselves being constrained in their growth by a shortage of skilled employees (World Bank, 2009).

Statistical analyses display increasing rates of wage return to secondary education in India, significantly greater than to primary and middle levels of education, illustrating the demand for skills gained at the secondary level outgrowing the supply in the Indian labor market (Kingdon & Unni, 2001; Kingdon, 2007; World Bank, 2009). In addition to its monetary yields, education is also argued to positively influence other determinants of well-being, including health outcomes, social networks, and civic participation (World Bank, 2009). Despite high returns to secondary education, it is remarkable that large population surveys found that while 72% of children of primary-school age attend primary school, only 51 percent of secondary-school age children attend secondary school (IIPS and Macro International, 2007a).

2.2 SECONDARY SCHOOL ATTENDANCE ON A GLOBAL SCALE

Despite recent advancements, India has not caught up with other comparator countries in terms of secondary school enrollment. In 2006, 17% of Indians over the age of 25 had followed some form of secondary education, relative to 36% in China, 17 percent in Sub-Saharan Africa, and 28% globally (World Bank, 2009). Figure 1 shows an Indian GER in secondary education of 47% in 2000. Although below the level forecasted for a country of India's per-capita GDP level, secondary enrolment rates have increased (World Bank, 2009).

Though it is unclear which definitions of secondary school age the World Bank uses in different countries, figure 2 paints a clear picture of how India's secondary school attendance rates have developed since 2006. A substantial GER increase, from 56.1 to 68.5, shows that India has made promising development in absolute terms of secondary

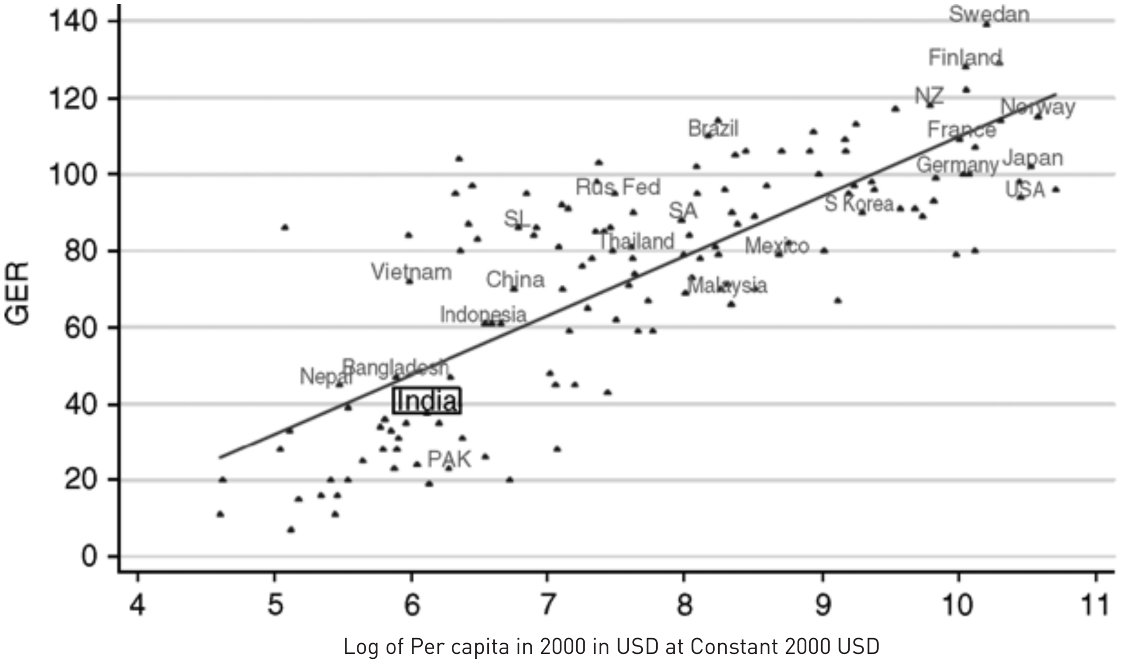


Figure 1. Cross-country comparison of gross enrolment ratios (GER) in secondary education and per-capita GDP, 2000
Note. Calculations are based on MHRD Selected Education Statistics for India and World Bank's Education Statistics Database for other countries. Reprinted from "The progress of school education in India", by G.G. Kingdon, 2007, Oxford Review of Economic Policy, 23(2), p. 170. doi: 10.1093/oxrep/grm015.

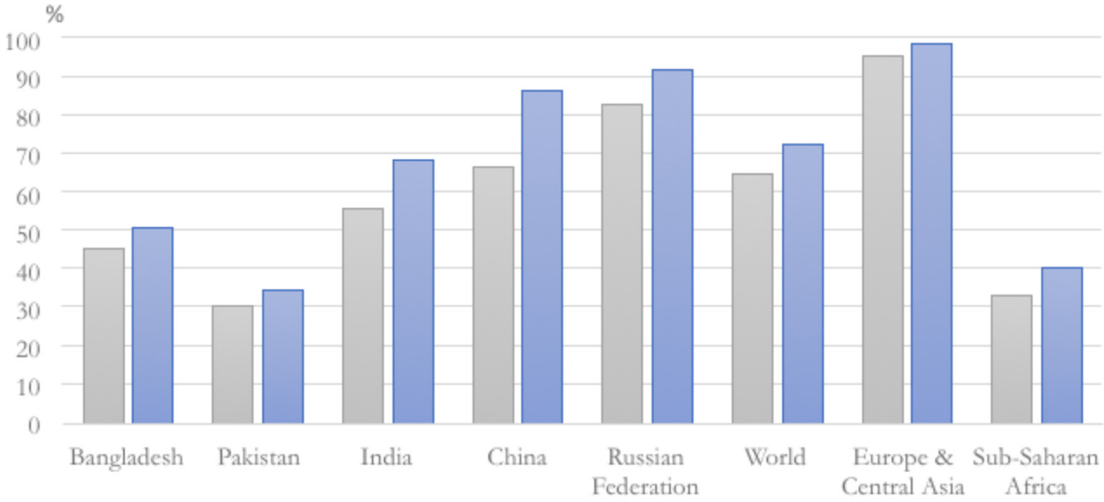


Figure 2. Cross-Country Gross Enrollment Rates (GER) Secondary School, 2006 and 2011
Note. GER is the total enrollment in secondary education, regardless of age, expressed as a percentage of the population of official secondary education age. Adapted from World Bank World Development Indicators. Data. http://data.worldbank.org/indicator/SE.SEC.ENRR/countries

school attendance. However, when these scores are benchmarked against scores of other countries, the improvement proves to be trivial. In 2006, India's GER ranked 110th out of 145 for which comparable data were available; in 2011, it ranked 107th out of 146 (World Bank, n.d.). While the other main emerging economy in Asia, China, is rapidly catching up with other large world economies, India is still below the GER world average.

2.3 GOVERNMENTAL EFFORTS TO PROMOTE ATTENDANCE

To improve attendance at the primary and secondary levels of education, the Government of India has undertaken various initiatives in recent years.

The Sarva Shiksha Abhiyan (SSA) is the national program launched in 2001 to achieve universal primary education by 2007 and universal elementary education by 2010. The focus of SSA is on "universal access and retention, bridging of gender and social category gaps in elementary education and improving the quality of learning" (Department of School Education & Literacy, n.d.-c, para. 1). SSA interventions include school and class room construction, creating toilets and drinking water facilities, academic resource and teacher support and free textbooks and uniforms (Department of School Education & Literacy, n.d.-c; National University of Educational Planning & Administration, 2008).

While the SSA program is not legally binding, the Right to Education (RTE) Act guarantees by law that every Indian child aged 6 to 14 is entitled to free, quality elementary education (Times News Network, 2010). The RTE was added to the Indian Constitution in 2002 with the passage of the 86th constitutional amendment, and has had legal force and effect as of April 1, 2010, under The Right of Children to Free and Compulsory Education Act. It was designed to warrant "that every child has a right to full time elementary education of satisfactory and equitable quality in a formal school which satisfies certain essential norms and standards" (Department of School Education & Literacy, n.d.-b, para. 1).

In conjunction with the RTE, the Rashtriya Madhyamik Shiksha Abhiyan (RMSA) was established with the intention to improve the access to secondary education for 15-18 year olds. It is envisaged to increase the secondary school enrolment rate to 75% by 2017, from just over 50% in 2005-06. Its main objectives include infrastructural improvements of classrooms, laboratories, and water and toilet facilities in Indian schools (Department of School Education & Literacy, n.d.-a).

Advancements have been made resulting from the RTE; a 2011 Unicef report found that since its implementation, 11 million children have additionally enrolled in elementary education, with 99% of India's rural population living within a one kilometer-range of a primary school (SOS Children's Villages, 2013). However, these progressed enrollments rates conceal the fact that 80 million children drop out before finishing their elementary education, and 8 million 6-14 year olds continued to be out of school (SOS Children's Villages, 2013). In a 2012 study commissioned by the SSA, the central government of India recognized that educational exclusion, based on caste status, class, gender, and special needs, affects access and completion of elementary education, and has constructed policies under the RTE Act to keep its most disadvantaged underage inhabitants inside the classroom (EdCIL, 2012; Human Rights Watch, 2014).

A comprehensive evaluation of the effects the RMSA has on enrollment and attendance of secondary school is yet to be offered, however, achievements are estimated to be behind initial targets (SOS Children's Villages, 2013). On the one hand, adolescent school attendance can be expected to rise as a result of the increased participation in elementary education. On the other hand, costs have been described as the main obstruction to school attendance amongst Indians; hence, the transition from free primary education to fee-based secondary schooling may dissuade many from attending levels beyond the elementary years (Jayachandran, 2002).

Interestingly, the RTE emphasizes on the importance of educational exclusion based on socio-economic factors in elementary education; however, RMSA improvements of secondary school attendance are mostly sought on the supply-side. Yet, when children reach adolescence, socio-economic factors like early marriage, household work, restrictive gender roles, and household wealth in context of the school fees, may play a role in the large discrepancy observed between primary and secondary school attendance. In the RMSA's focus on infrastructural improvements of schools, these demand-side factors seem relatively disregarded. In consideration of the value of secondary education to break through the intergenerational poverty trap, this research aims to find the socio-economic demand-side determinants of adolescent school attendance, which will be individually discussed subsequently.

2.4 DEMAND-SIDE DETERMINANTS OF SCHOOL IN CURRENT LITERATURE

The Indian Human Development Survey discovered positive associations between ASA and educated adults living in the household, higher income, urban residence, and effects of religion and caste on attendance (Desai et al., 2010). Vast interregional differences were thereby noted. The strongest attendance and lowest drop-out rates were observed for the wealthy, urban, Hindi, Forward caste males with educated parents (Desai et al., 2010). In the following sections, the literature on attendance discrepancies will be explored per individual factor. Accordingly, for each factor, hypotheses will be constructed based on the conclusions of previous research.

2.4.1 GENDER

While there is little gender-disparity in age appropriate education at the primary school level; a much higher proportion of boys (53%) than girls (43%) aged 15-17 are in school, as can be noted in table 2. An explanation offered by Borooah & Iyer is that many families in India display a "preference for sons" (2005, p. 133), which effects the size of the family investments in girls. The necessity to have female children educated only to a level that ensures their marriage – a few educational levels below that of their prospective husbands – creates a gender-bias in their participation in schooling (Borooah & Iyer, 2005). Therefore, females are hypothesized to have a lower ASA than males.

2.4.2 URBAN OR RURAL RESIDENCE

From the current literature it is still unclear whether living in urban areas has a significant positive or negative effect on adolescent school attendance. Jayachandran failed to prove significant associations between primary school attendance and urban residence, after controlling for other variables, including wealth and parental literacy (2002). Yet, the results from NFHS-3 in tables 1 and 2 show that, for both genders, attendance is higher in urban areas than in rural areas.

Table 1 provides definitions for net attendance ratio (NAR), gross attendance ratio (GAR), and the Gender Parity Index (GPI), and illustrates important discrepancies between school attendance in urban and rural areas. The NAR is slightly higher for students in urban residences, at around 71.9%, while the GAR is slightly higher for rural pupils, averaging around 83.1%. For both ratios measuring school attendance, the Gender Parity Index (GPI), calculated as the ratio of females attending school to males, approximates 1, indicating virtually equal attendance for both genders. For middle, secondary, and higher secondary school attendance, urban school attendance maintains a high GPI of 0.97 to 0.98. However, in rural areas, the gender gap drastically increases at higher levels of education, for which we see a fall of the GPI to 0.73 to 0.74. This research hypothesizes that rural adolescents will have lower school attendance than their urban counterpart. Additionally, rural residence is expected to interact with gender, such that the negative effect of rural residence on ASA is stronger for females than for males.

2.4.3 RELIGION

In 1950, Article 45 of the Indian Constitution stipulated that within 10 years from its commencement, free and compulsory education until the age of 14 would be ensured. In 2001, the GOI formally institutionalized free and compulsory elementary education for children of ages 6 to 14 as a fundamental right in the 93rd amendment. Yet, there exists considerable variation in the take-up of school education between Hindus, Muslims and SCs and STs (Borooah & Iyer, 2005).

Table 1. School attendance ratios
Net attendance ratios (NAR), gross attendance ratios (GAR), and Gender Parity Index (GPI) for the de facto household population by level of schooling and sex, according to residence, India, 2005-06

	Net attendance ratio[1]			**Gender**	**Gross attendance ratio**[2]			**Gender**
Residence	Male	Female	Total	Parity Index[3]	Male	Female	Total	Parity Index[3]
				PRIMARY SCHOOL				
Urban	73.9	73.7	73.8	1.00	82.9	80.7	81.9	0.97
Rural	72.9	69.4	71.2	0.95	85.6	81.2	83.5	0.95
Total	73.2	70.5	71.9	0.96	84.9	81.1	83.1	0.95
				MIDDLE, SECONDARY, AND HIGHER SECONDARY SCHOOL				
Urban	61.4	59.5	60.5	0.97	71.5	70.1	70.8	0.98
Rural	54.7	40.1	47.2	0.73	64.8	47.7	56.0	0.74
Total	56.8	45.6	51.2	0.80	66.9	54.1	60.5	0.81

Note. Adapted from National Family Health Survey (NFHS-3), 2005–06: India: Volume I, p. 31, by IIPS and Macro International, 2007a
In this table, children's age refers to their age at the start of the 2005-06 school year (assumed here to be April 2005).

[1] The NAR for primary school (standards 1-5) is the percentage of the primary-school age population (6-10 years) that is attending primary school. The NAR for middle, secondary, and higher secondary school (standards 6-12) is the percentage of the population in the appropriate age group for those school levels (11-17 years) that is attending those school levels. By definition the NAR cannot exceed 100.0 percent.

[2] The GAR for primary school (standards 1-5) is the total number of primary school students, expressed as a percentage of the official primary-school-age population (6-10 years). The GAR for middle, secondary, and higher secondary school (standards 6-12) is the total number of students in those school levels, expressed as a percentage of the official population that is the appropriate age to be attending those school levels. If there are significant numbers of overage and underage students at a given level of schooling, the GAR can exceed 100.0.

[3] The Gender Parity Index is the ratio of the NAR (GAR) for females to the NAR (GAR) for males at the same levels of school.

Table 2. School attendance by age group
Percentage of de facto household population age 6-17 years attending school in the 2005-06 school year by sex, residence, and age, India, 2005-06

	Male			Female			Total		
Age	Urban	Rural	Total	Urban	Rural	Total	Urban	Rural	Total
6-10 years	87.6	83.6	84.6	88.3	78.5	81.0	87.9	81.1	82.9
11-14 years	82.8	78.6	79.9	80.8	66.4	70.4	81.9	72.6	75.3
15-17 years*	55.7	50.3	52.8	54.5	34.4	42.9	55.2	42.0	47.8

Note. Adapted from National Family Health Survey (NFHS-3), 2005–06: India: Volume I, p. 33, by IIPS and Macro International, 2007a
* indicates corrected rates as found by this research, missing data removed
In this table, children's age refers to their age at the start of the 2005-06 school year (assumed here to be April 2005.

According to the 2011 Census, Muslims were the least literate among all religious communities in India (Registrar General, n.d.). 23.1 percent of all Muslim males and 20.1 percent of Muslim females had not obtained any educational credentials, apart from literacy. Moreover, a substantial proportion of Muslims – male (18 percent) and female (15.4 percent) – had attained only primary education. Meanwhile, at the higher levels of education, upper primary and above, Muslim attendance was significantly lower than attendance among all SCs and STs. Muslims, both males and females, have the highest figures for having never attended school and the lowest figures for those enrolled in primary or above level (Fazal, 2013).

In an estimation of wage returns, Unni (2007) has estimated that returns to education are insignificantly different from zero for Muslim men and are significantly smaller for Muslims and Christians than for the majority Hindu group. Yet, empirical data emphasize that specifically women from Muslim and SC and ST households were the least likely to have studied up to standard six, i.e. to have finished primary school, which students are expected to do around age 11 (Unni, 2008). Using a more theoretical approach, Unni explained that poor school infrastructure, increased involvement in domestic work and child labor, early marriage and child birth, and perceptions of parents, stimulate, through institutions of caste and religion, poor attendance of girls in school (2008). Based on the current body of research, this study will explore the negative association between being Muslim and low adolescent school attendance, and the reinforcing effect that being female has on that negative association.

2.4.4 CASTE OR TRIBE STATUS

Perhaps even stronger than the religious division in the Indian society is the stratification on the basis of caste hierarchy. Historically, Hindu society is separated into thousands of castes, mirroring the discrepancies between Hindus on socioeconomic, cultural, and educational levels. In the Indian Constitution, three of the major socioeconomic categories are distinguished by caste groupings: Forward Castes (FCs), Scheduled Castes and Scheduled Tribes (SCs & STs) – also known as Dalits, and Other Backward Castes (OBCs) (Chauhan, 2008).

The term Scheduled Caste was devised by the British government to designate all classes and castes formerly contained within the 'depressed classes'. Because of the negatively-regarded nature of their work, their criminal background, or their disease, they were labeled 'untouchables' and thus were denied equality of opportunities in all socioeconomic fields, including education (Chauhan, 2008). Scheduled Tribes mostly live in remote areas, where they remain geographically isolated from society at large. Due to their impoverished transportation and communication facilities, the STs have continuously been deprived of opportunities in every aspect of development (Chauhan, 2008).

Before 1950, SCs and STs were suffering under their formal segregation from the rest of the castes in public amenities, including schools, and were trapped into slavery through bonded labor. Since education has been widely acknowledged to be a potent tool for upward social mobility, by reducing socioeconomic inequalities, numerous efforts have been made to enable disadvantaged classes of the Indian society to flourish after "untouchability" was legally abolished in the 1950 National Constitution of India (Desai et al., 2010, p. 75; World Bank, 2009, Chauhan, 2008). Examples of these incentives for positive discrimination are the provision of constitutional protections, reservation of seats in educational and governmental institutions, and the facilitation of scholarships and student coaching classes (Zachariah, 1972).

In spite of the GOI's constitutional ban on the practice of untouchability and strenuous post-independence efforts to bridge the socioeconomic gap between the privileged and underprivileged groups, caste-based discrimination persists in India. SCs and STs have remained at an economic, cultural, and social disadvantage as a result of their specific geographical and occupational conditions (Chauhan, 2008). Chauhan's analyses indicate that after 60 years of affirmative action, participation of the lower castes in higher education has still not caught up with their share in the overall population. Jayachandran additionally offered an explanation for low school attendance associated with STs, SCs and Muslims. She argued that education may not be as strong of a cultural value relative to the higher classes. It may also be possible that discrimination within the educational system plays a role (Jayachandran, 2002).

In addition to the aforementioned demand-side issues that impact school attendance, Nambissan considers the influence the formal educational system has on Dalit students. Despite encouragement from underprivileged household members, supply-side issues, among which inadequate infrastructural facilities, the absence of pedagogic support, and an indifferent attitude of teachers and school staff negatively affect the school attendance of Dalit children (Nambissan, 1996). While these structural problems within the educational system of India, including student discrimination which make up an important part of the variation in ASA among castes and tribes are important to recognize, this research paper focuses on the demand-side issues associated with the lower demand for education by Dalit adolescents.

In wage returns estimations, the wage returns to education are significantly lower for ST groups than for the majority Hindu group after controlling exclusively for observed characteristics (Unni, 2007). Remarkably, Jayachandran (2002) did not find significant differences in attendance among castes and tribes. Still, to fully explore the associations between SC/ST-membership, including the weak position of the female Dalit student, as generally agreed upon by the current body of research, it is hypothesized that SC/ST-membership has a negative effect on ASA. Additionally, females are expected to be especially negatively affected by SC/ST-membership in terms of ASA.

2.4.5 PARENTAL EDUCATION

Jayachandran ran several regressions in an attempt to uncover socio-economic determinants of school attendance in the 5-14 age group, whereby she focused on the gender bias in primary education. Based on census data for 1981 and 1991, she found parental literacy to be positively related to primary school attendance (2002). Additionally, following from the NFHS-3, the level of mother's education was especially helpful in explaining school attendance variation. While in households where the mother had no education, children aged 5-14 were most likely involved in paid/unpaid/household work instead of school; those with mothers who had over twelve years of education showed barely any children engaged in any type of work (IIPS and Macro International, 2007a). More generally, an increase in the years of education of the mother meant a decrease in the proportion of children engaged in work instead of school. Doraisamy maintained that an educated mother specifically improves the chances of the female child attending school (as cited in Unni, 2008). This study will investigate two separate hypotheses on the positive direct effects of both father's education and mother's education on ASA. Thereby, mother's education is projected to interact with gender, such that being female strengthens the positive association between mother's education and ASA.

2.4.6 WEALTH

It has been pointed out that despite the ostensible link between mother's education and her potential influence on getting her offspring into school, the increased ASA may also partially reflect that the probability for women in wealthy households to have educational credentials is highest compared to less affluent families, and simultaneously, children from prosperous households are not required to work (Unni, 2008). Following this line of reasoning, wealth is explored as a distinct variable.

The biggest inequalities in ASA are found between children from different wealth quintiles. For children from the wealthiest 20% of all households, the NAR of children aged 11-17 school is 83% compared to a NAR of only 29% among children from the poorest households (Unicef, 2011). Jaychandran (2002) named poverty, in relation to schooling costs, as the main barrier to education, and found a significant negative effect of poverty on both male and female school

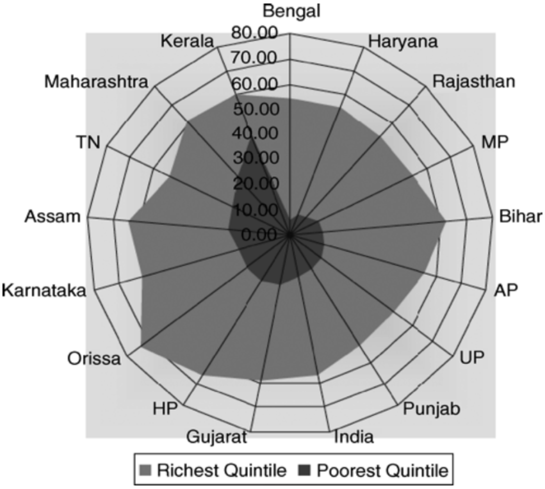


Figure 3. Differential access (to secondary schooling) between the top and bottom income quintiles
Note. Reprinted from "The progress of school education in India", by G.G. Kingdon, 2007, Oxford Review of Economic Policy, 23(2), p. 174, . doi: 10.1093/oxrep/grm015.

attendance rates. After confirming the positive influence of household prosperity, Doirasamy reasoned that the greater the household resources, the higher the demand for the children's education with a decreasing requirement for children to work to supplement their parents' income (as cited in Unni, 2008).

Based on National Sample Survey data for 1999-2000, Kingdon (2007) found that there exists considerable interstate variation in the magnitude of inequality in access to secondary schooling, as shown in figure 3. The inequality between the top and bottom quintiles of the distribution of household income is largest in Haryana, Andhra Pradesh, and the 'BIMARU' states. The BIMARU states are the literally 'sick', largely backward, states of Bihar, Madhya Pradesh, Rajasthan, and Uttar Pradesh, which are behind in many other fields of social development besides educational indicators. The inequality gap is smallest in the politically left-leaning states of Kerala and West Bengal (Kingdon, 2007). In line with the current literature, wealth has a hypothesized positive effect on ASA.

2.4.7 HOUSEHOLD SIZE

As household size increases, the gender gap in school attendance widens further to the advantage of males (Jaychandran, 2002). Furthermore, Doraisamy found that especially female participation in education is adversely affected by the presence of young siblings (as cited in Unni, 2008). In order to test the effects of household size on ASA, in particular for females, it is hypothesized that household size has a negative effect on ASA, and through a gender interaction, being female strengthens the negative association between household size and gender.

3. METHODOLOGY

The purpose of this section is to outline the methodological approach of this research. To recapitulate, the objectives of this research are to: a) develop an improved understanding of Indian adolescent school attendance; b) verify and improve the set of determinants established in the current literature to be included in an expanded conceptualization of Indian ASA; and c) explore the demand-side factors variables and interactions that advance and obstruct Indian ASA.

Given the primary objectives of the research, the availability of large school population survey data, and the limited amount of existing statistical analyses on the determinants of ASA in India, a secondary data analysis is utilized. First, the secondary data obtained from the National Family Health Survey (NFHS) and its collection are discussed, after which an explanation of this study's operationalization will follow.

3.1. DATASET: NFHS-3

Based on two factors, the availability of raw data and sample size, the most recently available data from the National Family Health Survey was chosen for secondary analysis.

The 2005-06 National Family Health Survey (NFHS-3) is the third in a series of nationwide surveys conducted with a representative sample of households throughout India; previous NFHS surveys date back to 1992-93 (NFHS-1) and 1998-99 (NFHS-2). The Ministry of Health and Family Welfare (MOHFW), and the Government of India (GOI), initiated the NFHS surveys with the purpose of providing high quality data on population and health indicators. The NFHS surveys assist policymakers and programme administrators in designing and executing population, health, and nutrition programmes. The MOHFW appointed the International Institute for Population Sciences (IIPS), Mumbai, as the nodal agency for each of the three rounds of NFHS (IIPS and Macro International, 2007a).

The NFHS surveys use standardized questionnaires, sample designs, and field procedures to collect data uniformly across all states (IIPS and Macro International, 2007a). The NFHS-3 employed stratified sampling, whereby the rural and urban samples within every state were drawn independently and, where possible, the sample within each state was allotted proportionally to the size of each state's rural and urban populations (IIPS and Macro International, 2007a).
In NFHS-3, 18 research organisations conducted interviews with a total of 109,041 households from all 29 states, with a household response rate of 97.7% (IIPS and Macro International, 2007a). A total of 515,507 individuals who stayed in the household the night before the interview were deemed de facto household members, and were enumerated in the 109,041 sample households (IIPS and Macro International, 2007a). Additional interviews conducted with 124,385 women age 15-49 and 74,369 men age 15-54 are outside of this research's scope, which focuses its secondary data analysis on the household data.

Fieldwork for NFHS-3 was conducted in two phases from November 2005 to August 2006. Technical assistance for NFHS-3 was provided by Macro International, Maryland, USA. From the eighteen research organizations that conducted fieldwork for NFHS-3, thirteen are private sector research organizations and five are Population Research Centres (PRCs) established by the GOI. Each research organization carried the responsibility for collecting the data in one or more states (IIPS and Macro International, 2007a)

NFHS-3 was funded by the United States Agency for International Development (USAID), the United Kingdom Department for International Development (DFID), the Bill and Melinda Gates Foundation, UNICEF, UNFPA, and the GOI (IIPS and Macro International, 2007a)

In 2014-2015, India will implement the fourth round of NFHS-4. Like its predecessors, NFHS-4 will be conducted by the MOHFW, GOI, coordinated by the IIPS and implemented by local research institutes. At the time of conducting this research, the data collection for NFHS-4 has not been finalized (IIPS and Macro International, 2014).

3.2 OPERATIONALIZATION

To test the aforementioned hypotheses, a logistic regression (LR) will be performed with the aim to determine which socio-economic factors influence Indian adolescent school attendance. The operationalization of the included variables are explained subsequently.

3.2.1 DEPENDENT VARIABLE: ADOLESCENT SCHOOL ATTENDANCE

Previous research shows two main concepts related to school participation rates of Indian social groups: enrollment and attendance. The World Bank utilizes the definition of gross enrolment ratio (GER), “the total enrollment in secondary education, regardless of age, expressed as a percentage of the population of official secondary education age”, which can exceed 100% due to the “inclusion of over-aged and under-aged students because of early or late school entrance and grade repetition" (World Bank, n.d.). However, enrolment rates are solely measured at the beginning of the school year, and conceal drop-out and non-attendance later on (Kingdon, 2007). The current school attendance is considered a more reliable indicator of schooling participation rather than enrolment.

Basing attendance analyses on attendance at the age appropriate to the level of education, as has been done in previous research, may make the definition of ASA too narrow. As age widely varies within Indian school grades, adolescent school attendance should not be restricted to the definition of adolescent attendance of secondary school. This research defines Indian Adolescent School Attendance (ASA) as Indian school attendance of 15-17 year olds at any school level, as recorded in the NFHS-3 for the school year 2005-06. The dependent variable of the LRs in this research will hence be the dichotomous variable measuring whether 15-17 year old Indian adolescent attended school or not.

3.2.2 INDEPENDENT VARIABLES

Based on the current literature, the following independent variables will be included in the logistic regression models:

1. Gender: a dichotomous variable to indicate the adolescent’s gender.

2. Residence: a dichotomous variable to indicate whether the adolescent lives in an urban or rural area.

3. Religion: while variation in religion within households may exist, the NFHS-3 recorded only the religion for every household head, not for individual household members. In the analyses, the religion given for the household head will be assumed to be the same for all individuals within his or her household.
   Based on the current literature and the hypotheses this research designed, the religion variable has been divided into three levels: Muslim, Hindu, and other. The last category includes Christian, Sikh, Buddhist/Neo-Buddhist, Jain, Parsi/Zoroastrian, No religion, Donyi polo, other.

4. Caste or tribe status: again, in the dataset, only the caste status for every household head is recorded, and will be assumed to be the same for all household members. The caste or tribe status variable includes three levels: Scheduled Caste (SC), Scheduled Tribe (ST) and other.

5. Mother's education: a continuous variable measuring the number of single years of education the child's mother has completed. Since the original NFHS-3 dataset did not include a variable for mother's or father's education, it was created for the purpose of this research. A variable for education in years for either parent could only be constructed if, for a given case, both parents were alive, registered in the same household sample, and each recorded their education details in the NFHS-3 survey. As will become apparent in the results section, this led to a greater number of missing cases in the 15-17 year olds' sample (table 6), yet both parental education variables proved to be important explanatory factors of ASA variance.

6. Father's education: a continuous variable to indicate the number of single years of education the child's father has completed, constructed identically to the mother's education variable.

7. Wealth index: the NFHS-3 constructed a wealth index, by combining information on 33 household assets and housing characteristics such as ownership of consumer items, type of dwelling, source of water, and availability of electricity into a single wealth index. As the household population is divided into five equal quintiles at the national level from 1 (poorest) to 5 (wealthiest), the index is treated as a continuous variable.

8. Household size: a continuous variable measuring the number of usual household residents.

9. State: dummy variables were created for the 29 levels of the control variable 'State', to indicate which state the subjects live in.

4. RESULTS

4.1 DESCRIPTIVE STATISTICS

This study's sample came from the larger NFHS-3 sample (N = 515,507). Among the respondents, only 33,253 (6.5%) respondents were aged 15 to 17 at the start of the 2005-06 Indian school year. After removing the cases with missing values for certain variables, this study's sample counts 13,973 individuals. A comprehensive breakdown of the sample is given in table 3. A detailed table including the frequency and percentage of missing data per variable, for all 33,253 NFHS-3 respondents aged 15 to 17 can be found in table 6, Appendix A.

As has been outlined in the methodology of this report, variables for mother's and father's education could only be constructed for individuals whose parents were both alive, in the same household sample, and had registered their schooling details. Table 6 shows that, in line with the chosen approach for the creation of parental education variables, for approximately half of all 15-17 year olds in the NFHS-3 sample parental education variables could not be computed. Any deviations that may exist between demographic characteristics of the study sample and the demographic characteristics of all 15-17 year olds in the NFHS-3 sample may originate from this sampling method and will be highlighted accordingly.

The sample included 51.9% males and 48.1% females. Table 3 shows male (52.8%) and female (42.9%) attendance rates for 15-17 year olds averaging around 47.8%, similar to those rates found in table 6 (appendix A), while the latter still take into account several missing cases. However the attendance rates in the study sample are considerably higher for both males (65.3%) and females (58.8) at around 62.2 on average. The limitations section will further expand on a certain selection bias that may exist here, resulting from the missing case deselection.

Around 73% of the sample was drawn from the Hindu group. The remainder was drawn from the Muslim group (14%) and a group including all other religions and atheists (12.9%), as detailed in Table 3. These numbers are virtually equal in the larger samples of table 6 and the overall NFHS-3, though the 2011 census showed a higher proportion of Hindus (80.5%), yet a comparable proportion of Muslims (13.4%) (Census, n.d.).

The proportion of SCs (18.7%) and STs (12.6%) is also highly similar across all 15-17 year olds in the NFHS-3 sample, while the 2011 census recorded slightly lower figures for both SCs (16.6%) and STs (8.6%) as fractions of India's population (Census, 2011a). According to United Nations Development Program (UNDP) data, the average (over 25 years old) adult's schooling was four years in 2005; the average parental education in this research sample was just under 5 years (United Nations Development Program, 2013). The difference in figures may depend on the definition

Table 3: Demographic characteristics of 15-17 year olds at the start of the 2005-06 school year within the NFHS-3 sample, missing data removed

	Male (n = 7255)		**Female** (n = 6718)		**Total** (n = 13,973)	
Variables	Frequency	Percentage	Frequency	Percentage	Frequency	Percentage
Total school attendance	4736	65.3	3949	58.8	8685	62.2
Primary	120	1.7	88	1.3	208	1.5
Secondary	4466	61.6	3609	53.7	8075	57.8
Higher	150	2.1	252	3.8	402	2.9
Residence						
Urban	3339	46.0	3314	46.4	6453	46.2
Rural	3916	54.0	3604	53.6	7520	53.8
Religion						
Hindu	5384	74.2	4828	71.9	10212	73.1
Muslim	958	13.2	994	14.8	1952	14.0
Other	913	12.6	896	13.3	1809	12.9
Caste or tribe type						
Scheduled Caste	1351	18.6	1256	18.7	2607	18.7
Scheduled Tribe	926	12.8	836	12.4	1762	12.6
Other	4978	68.6	4626	68.9	9604	68.7
Literate father	5292	72.9	4909	73.1	10201	73.0
Literate mother	3434	47.3	3225	48.0	6659	47.7
Father's education in years						
0	1970	27.2	1818	27.1	3788	27.1
1 – 5	1464	20.2	1328	19.8	2792	20.0
6 – 10	2518	34.7	2337	34.8	4855	34.8
11 +	1303	18.0	1235	18.4	2538	18.2
Mother's education in years						
0	3850	53.1	3505	52.2	7355	52.6
1 – 5	1138	15.7	1125	16.8	2263	16.2
6 – 10	1648	22.7	1519	22.6	3167	22.7
11 +	619	8.5	569	8.5	1188	8.5
Wealth Index						
Poorest	848	11.7	713	10.6	1561	11.2
Poorer	1079	14.9	948	14.1	2027	14.5
Middle	1457	20.1	1387	20.6	2844	20.4
Richer	1737	23.9	1631	24.3	3368	24.1
Richest	2134	29.4	2039	30.4	4173	29.9
Number of household members						
0 – 5	4952	68.3	2850	42.4	6512	46.6
5 – 10	2026	27.9	3481	51.8	6675	47.8
11+	277	3.8	387	5.8	786	5.6

Source: 2005–6 Household data from the National Family Health Survey (DHS Program, 2006).

of years of education. The UNDP statistics represent a number that has been converted from education attainment levels using official durations of each level, whereas the NFHS-3 recorded single years of completed education, and may be considered more accurate for our purposes. In the entire NFHS-3 sample, adults over 25 had 5.7 years of education on average. A slightly lower average for parental education in the study sample is explained by the observation that parents with children aged 15 to 17 are an older group than the collection of 25 year old Indians. As older cohorts are less educated than more recent cohorts, an average education of five years for parents of individuals within the sample seems to be compatible with an over-25 average of 5.7 years for those included in the NFHS-3 (IIPS and Macro International, 2007a).

The rural-urban distribution in our sample – 46.2% against 53.8% respectively – is highly similar to proportions found in the overall NFHS-3 sample, of 44.2% and 55.8% respectively. The discrepancy between the NFHS-3 figures and the distribution found in the 2011 Census, of 68.8% rural residents against 31.2% urban citizens, currently remains unexplained (Registrar General, 2011b).

The average household size of 5.6 for the all-encompassing NFHS-3 sample, is very similar to what was found in the 2001 (5.3) and 2011 Census (4.9) (Nayak & Behera, 2014). In the study sample, a slightly higher average of 6.2 was noted. It logically follows that people with more children will have larger households.

Approximately 11.2 percent of the sample was classified in the poorest rank of the Wealth Index, with increasing proportions for each consecutive wealth rank, leading up to a 29.9 percentage of the richest Indians as a proportion of the complete population, which are found very similar to proportions over the sample in table 6 and the larger NFHS-3 sample. Since the Wealth Index was constructed by NFHS-3, it is problematic to compare these proportions to other population survey data.

On average, the typical respondent in the sample was a rich Hindu male attending secondary school from a father with between 6-10 years of education, an uneducated mother, and living in a 6-person rural household.

4.2 CORRELATIONS

Inspection of the correlations in table 4 between the DV and all IVs shows that school attendance is most strongly correlated with an individual's father's education in years ($r(13971) = .41, p < .001$) and his or her mother's education ($r(13971) = .41, p < .001$). Thereby, someone's position on the constructed wealth index plays a comparably important positive role ($r(13971) = .37, p < .001$).

The negative correlation between rural residence and school attendance ($r(13971) = -.14, p < .001$) confirms previous findings that rural 15-17 year olds are less likely to be in school than their urban counterparts. The female gender, SC/ST-membership, and household size are all weakly negatively related to school attendance.

It may be interesting to note that both rural residence and being Muslim are significantly negatively associated with parental education. Also, as was suggested by Unni (2008), there is a strong relationship between household wealth and the education of parents.

4.3 LOGISTIC REGRESSIONS

4.3.1 DIRECT EFFECTS – MODEL 1

A logistic regression was completed to determine the direct relationships between the 15-17 year olds' demographic variables, as outlined in the methodology section, and school attendance in the 2005-06 school year. The model includes dummy variables for various religious groups whereby the Hindu and Muslim groups were contrasted against the collection of all individuals following another or no religion. Dummy variables SCs and STs were contrasted against the category that contains those who do not fall under any SCs/STs. The results of the majority of direct effects are presented in model 1 of table 5. Since no interaction terms were produced for 'Father's education in years' and 'Wealth index', the interpretation of the direct effect of these variables will refer to model 2 of table 5. An indication of the size of the effects is given by the odds ratios (OR), which show the change in likelihood of school attendance associated with a unit change per IV. An OR value of greater than 1 indicates a positive association of the variable with school attendance, while an OR smaller 1 shows a negative association.

While the female gender significantly confirms

Table 4. Correlations between variables

				Religion			Caste status					
				dummy variables)			(dummy variables)					
	School attendence 05-06	Gender	Residence	Hindu	Muslim	Other religion	Scheduled Caste (SC)	Scheduled Tribe (ST)	Other/ none	Years education father	Years education mother	Wealth index
Gender (female = 1)	-.067***											
Residence (rural = 1)	-.140***	-,003										
Hindu	.063***	-,026**	,053***									
Muslim	-.155***	,023**	-,111***	-,664***								
Other religion	.076***	,011	,044***	-,635***	-,155***							
SC -.068***	,001	,007	,159***	-,166***	-,037***							
ST	-.025**	-,005	,157***	-,241***	-,126***	,449***	-,182***					
Other/no C/T	.075***	,003	-,119***	,040***	,230***	-,290***	-,710***	-,563***				
Years education father	.408***	,003	-,313***	,087***	-,112***	,001	-,119***	-,134***	,196***			
Years education mother	.384***	,005	-,369***	,008	-,100***	,092***	-,155***	-,090***	,195***	,662***		
Wealth index	.373***	,019*	-,563***	-,072**	,025**	,069***	-,113***	-,181***	,225***	,553***	,558***	
House-hold size	-.078***	,052***	,067***	-,156***	,186***	,015	-,021*	,051***	-,019*	-,108***	-,172***	-,058***

*p < .05; **p < .01; ***p < .001

Source: 2005–6 Household data from the National Family Health Survey (DHS Program, 2006).

the hypothesized negative effect on attendance (B = -.398, p < .0001), we fail to confirm a negative effect of rural residence. In fact, model 1 shows the existence of a positive direct effect of rural habitation on school attendance (B = .409, p < .0001).

Muslims were over two times more likely than members of any other religion to be absent from the classroom (B = -.875, p < .0001). However, the LR failed to confirm the hypothesized negative effect of SC/ST-membership and household size on school attendance.

The LR confirmed significant positive effects of father's education (B = .107, p < .0001) and mother's education in years (B = 114, p < .0001). The larger impact was for 'mother's education in years; children from the most educated mothers, with 20 years of education, were 22 times more likely to attend school than an unschooled mother's offspring. A similar effect was apparent for father's education (model 2). In addition to that, a significant positive coefficient on 'Wealth index' confirms the hypothesized positive effect of wealth on ASA.

Overall, the combination of direct effects explain a fair amount (36.4%) if variance in the school attendance of 15 to 17 year old Indians. Table 5, model 2, adds the hypothesized gender interaction effects, which slightly increases the explanatory power of the model accounting for around 37.2% of the variance.

4.3.2 INTERACTIONS – MODEL 2

A significantly positive coefficient of the 'gender' and 'residence' interaction (B = -.606, p < .0001) confirms the hypothesis that being female strengthens the negative association between rural residence and ASA. This interaction decreases the likelihood of school attendance by rural female adolescent almost by a factor 2. A similar negative effect can be noted for Muslim females (B = -.502 p < .01). Thereby, while no interaction between gender and the Hindu religion was hypothesized, model 2 shows a negative effect on school attendance for the combination of being female and Hindu (B = -.691, p < .0001). Besides not finding a direct effect of SC/ST-membership on ASA (model 1), model 2 fails to confirm the hypothesized interaction effect of SC/ST-membership with gender.

There is, however, a significant interaction between Mother's education and gender (B = .030, p < .05). Females born from an educated mother thus have an increased likelihood of school attendance. The significant interaction term of "number of household members" and "gender" (B = -.033, p < .05) also confirms that being female strengthens the negative association between household size and gender, which was not significant as a direct effect.

5. DISCUSSION

5.1 GENDER

The direct effect gender has on ASA showed a negative coefficient, as hypothesized. These findings confirm that a combination of gender-biased traditional values within households favor male education. Apart from the cultural ideas behind a child's schooling, there is also a key economic argument; parents may only value that part of the return to their child's education which benefits them personally. On top of the dowry parents pay at a daughter's wedding, the returns to a daughter's education are enjoyed mostly by her in-laws' family (Kingdon, 2002). With the traditionally early age of marriage for girls in Indian culture, daughters are often only educated to a level that ensures their marriage, which is below that of their male counterparts (Borooah & Iyer, 2005). When the gender-interaction effects were added to the analysis in model 2 (table 5), the direct effect of gender is no longer straightforwardly interpretable.

5.2 URBAN OR RURAL RESIDENCE

Despite low attendance rates in rural areas (tables 1 & 2) and the negative correlation between rural residence and ASA (table 4), the LR returned a positive coefficient for rural residence.

Model B1, appendix B, shows that before controlling for household wealth and parental schooling, rural residence had a negative effect on ASA. Yet adding these variables to include all direct effects, as presented in model 1, table 5, they seem to explain the negative effect that rural residence appeared to have when they were excluded from the model.

These findings allude to the conclusion that rural residence does not inherently negatively affect ASA; to the contrary, it does so positively according to model 1. Yet since rural habitation is strongly negatively correlated with household wealth, mother's and father's education, which all individually have strong adverse effects on ASA, it is unlikely to find

Table 5. Logistic Regression analysis of school attendance – Parameter estimates (B), Standard Errors (SE) and Odds Ratios from the logistic regression of school attendance for children aged 15-17, India, 2005-06

	Model 1			Model 2		
Variable	B	SE (B)	Odds Ratio	B	SE (B)	Odds Ratio
Constant	–.875***	.224	–	–1.396***	.244	–
Gender						
Female	–.398***	.041	.671	.640**	.198	1.897
(base = male)						
Residence						
Rural	.409***	.055	1.506	.709***	.071	2.033
(base = urban)						
Religion						
Hindu	.023	.092	1.024	.371**	.115	1.449
Muslim	–.875***	.108	.417	–.628***	.137	.534
(base = other/no religion(s))						
Caste type						
Scheduled Caste	–.061	.055	.941	–.127	.075	.880
Scheduled Tribe	.058	.079	1.059	.012	.102	1.012
(base = other/no caste/tribe)						
Number of household members	–.003	.008	.997	.013	.011	1.013
Father's education in years	.106***	.006	1.112	.107***	.006	1.113
Mother's education in years	.114***	.007	1.120	.099***	.010	1.104
Wealth Index	.423***	.023	1.527	.431***	.023	1.539
Interactions						
Gender*Residence				–.606***	.090	.545
Gender*Religion						
Gender*Hindu				–.691***	.144	.501
Gender*Muslim				–.502**	.180	.606
Gender*Caste						
Gender*Scheduled Caste				.120	.108	1.127
Gender* Scheduled Tribe				.087	.143	1.091
Gender*number of household members				–.033*	.017	.967
Gender*mother's education in years				x.030*	.013	1.031
Nagelkerke pseudo R2		36.4%			37.2%	
X^2		4352.8, df = 38, p<0.0001			4457.8, df = 43, p<0.0001	

*p<0.05; **p<0.01, ***p<0.0001.

Note: The dependent variable in this analysis is school attendance in the 2005-06 school year, coded so that 0 = did not attend, and 1 = did attend school.
To account for interstate differences between India's 29 states shown in table 8 (Appendix C), 28 dummy variables have been used in all models, for which:
Model 1: Wald-stat = 517.688, df = 28, p < .0001 with 20 state dummy variables significant at α = .05.
Model 2: Wald-stat = 517.472, df = 28, p < .0001 with 21 state dummy variables significant at α = .05.

Source: 2005–6 Household data from the National Family Health Survey (DHS Program, 2006).

high adolescent attendance rates in rural areas. Previous analyses of primary school attendance have also failed to prove significant positive associations between attendance and residence, after controlling for wealth and parental literacy (Jayachandran, 2002).

Despite the positive coefficient of the direct effect of residence, the results indicate that female adolescents in rural areas are disadvantaged by their habitation. Table 1 indicated that the Gender Parity Index (GPI) for middle, secondary, and higher secondary school attendance in urban areas was almost at the ideal value of 1, whereas the rural GPI was just over 0.7. A potential explanation of the negative interaction effect of gender and residence is that cultural norms in rural areas could be biased towards men and their schooling. Working, as opposed to learning, may be the norm for women in countryside villages. In rural India, especially in the 1970s early green revolution period, there were no direct economic returns to a woman's education, since women were not involved in occupations where education was remunerated (Kingdon, 2002).

The coefficients on residence may also capture a supply-side effect on school attendance. Distances in rural areas are generally larger and seem to impact school attendance; especially for girls going through puberty distance discourages attendance because of safety concerns (Bing, Goldschmidt, Boscardin & Azam, 2011). As the present research focuses on the demand-side determinants of school attendance, these supply-side issues will not be explicitly considered and are left to be further explored by future empirical efforts.

5.3 SCHEDULED CASTES AND TRIBES

Also contrary to expectations, the direct and gender interactions of SC/ST-membership did not prove significant. Similar to the variable for rural residence, it is negatively correlated with ASA, and has a significant negative effect on ASA when not controlling for parental education and household prosperity (model B1, Appendix B). The addition of these variables in models B2 and 1 demonstrates that the negative effect of SC/ST-membership is in effect explained by wealth and parental education. Equally, Jayachandran (2002) found no significant differences in school attendance between SCs, STs and those who fall outside of those, within the 5-14 age group.

Assumptions held in the literature about education as a weaker cultural value among SCs and STs, and strong effects of discrimination based on caste or tribe status, fail to hold (Jayachandran, 2002).
In recent decades, there has been a surge in policies devoted to increasing school attendance of backward caste and tribe members by generally focusing on increasing the quality and supply of education.

These initiatives include schemes that provide financial assistance for organizations in Educationally Backward Districts, particularly those which are predominantly inhabited by SCs, STs and other educationally backward minorities. Additionally, reservation policies for places in higher education have stimulated the admission of SC and ST members to these institutions (National University of Educational Planning & Administration, 2008).

While supply-side determinants of school attendance are outside the scope of this research, the effects the policies of affirmative action have had on demand for adolescent SC/ST education are valuable to consider – they do, however, remain unclear at this point. The prospective of improved success in secondary and tertiary education may have improved their perceptions of value of secondary education, and increased SC/ST ASA. A comprehensive evaluation of Indian affirmative action policies and its effects on education may reveal improvements in both supply and demand of ASA among SCs and STs.

5.4 HOUSEHOLD WEALTH

Household wealth, measured on a five-point scale of five equal wealth quintiles, was shown to form a main barrier to education attendance. Findings from previous research were confirmed; the greater the household resources, the higher the demand for adolescent schooling with a diminishing necessity for children to work to supplement their parents' income (Jayachandran, 2002; Unni, 2008). The newly encountered costs of education at the age of 15 – after years of free schooling – form a major bottleneck for attendance of adolescents from underprivileged households.

The World Bank (2009) named the high direct and indirect costs of schooling an important factor delimiting demand. On average, including both public and private institutions, the direct costs of secondary education are double those of primary education; the costs of higher secondary and tertiary education are four and eight times as much respectively (World Bank, 2009). Especially in recent years, the importance of wealth in relation to school costs as an ASA determinant has increased. While average costs of a single child's education have nearly tripled, from Rs 35,000 in 2005 to over Rs 94,000 in 2011, these increasing costs have not been countered by an equal rise in wages. In 2011, 65% of Indian parents indicated to spend over half of their income on items and activities central to the school curriculum, including fees, transport, uniforms, books, stationery, educational trips, extra tuitions and extra-curricular activities (Dhawan, 2013).

The opportunity costs of education may also be an important factor in deterring ASA, given the average annual wages for grade 8 completers (Rs. 16,000 in 2009) in an emerging economy. As a result of households having to sacrifice earnings and tolerate the direct costs of education, the demand for adolescent school attendance suffers (World Bank, 2009).

5.5 PARENTAL EDUCATION

Scholars have suggested that a link between mother's education, and her potential influence on getting her offspring into school, may partially reflect the fact that children from prosperous households are not required to work. Simultaneously, the probability for women in wealthy households to have educational credentials is highest compared to less affluent families (Unni, 2008). Though table 4 confirmed a strong positive correlation between wealth and both parental schooling indicators, models 1 and 2 (table 5) show strong direct effects of mother’s and father’s education on ASA, additional to wealth effects.

The solid effects of parental education on primary school attendance have been emphasized by the current literature (IIPS and Macro International, 2007a; Jayachandran, 2002; Unni, 2008). The results of this study extend the positive effects of schooled parents to ASA; educated parents are more capable of understanding the significance of being educated and in turn encourage school attendance of their (adolescent) children.

The significant interaction between gender and mother’s education underlines the importance of maternal influence on the female adolescent attendance. An educated mother has experienced the added value of education, and wants to ensure the school attendance of her daughter(s).

5.6 HOUSEHOLD SIZE

When controlling for other variables, previous research has not been able to find a significant association between the number of household residents and Indian school attendance. By the same token, the results in table 5 display no significant direct effect of household size on ASA. However, the significant interaction term between gender and household size shows that as household size increases, the gender gap in school attendance widens further to the advantage of males, as hypothesized. Larger families imply a larger amount of domestic work, for which the elder daughters seem to carry the burden. These chores may include cooking, cleaning, and looking after younger siblings, cousins, and elderly family members, tasks traditionally performed by women within the Indian household (Jaychandran, 2002). Meanwhile, boys are sent to school.

5.7 RELIGION

While the correlations in table 4 showed that Muslim households tend to be of larger sizes and may partially explain poor Muslim attendance, model 2 (table 5) showed that the isolated direct effect of being Muslim on ASA is also significantly negative.

Lower Muslim participation can partly be explained by the finding that returns to education are significantly smaller for Muslims than for other religious groups (Unni, 2007). Discrimination on the supply side of education also seems to play an important role in limiting demand for education from the marginalized Muslim group. Human Rights Watch’s research (2014) exposed various incidents in which Hindu teachers in government schools made deprecating comments about Muslim pupils. A systematic bias against Muslims in India is said to persist in the educational system, with Muslim children being neglected by the school system. Instead of becoming inclusive and unbiased institutions as intended under the RTE Act, schools end up neglecting and alienating Muslim children (Human Rights Watch, 2014).

Additionally, while no direct effect of the Hindu religion on ASA was found in model 1, the significant interaction term in model 2 illustrates that Hindu, as well as Muslim females are negatively affected by the combination of their gender and religion. As explained in the literature, certain conservative values within Hinduism and Islam may reinforce the notion of a female role in a domestic context, rather than in schools or work places. Increased household work and child labor, early marriage and child birth, and perceptions of parents, can, through institutions of religion, encourage female absenteeism of both Muslim and Hindu adolescents in schools (Parveen, n.d.; Unni, 2008).

Qualitative research in Rampur, Uttar Pradesh, found that Muslim parents in particular are reluctant to send their female children to school, because they are concerned that their daughters "will be misguided in the name of Education & indulge in objectionable situations which make them guilty" (Parveen, n.d., p. 24). Moreover, Parveen discovered that some Muslim parents fear that educated daughters will demand their rights in property and have to give their son's share to them. After the daughter's marriage, the property would then mostly benefit her husband's family, instead of the family she was born into.

6. LIMITATIONS

There are a number of limitations of this research that highlight additional areas for further exploration. These limitations include the potential bias of the research sample, the relatively dated dataset, and the concentration on demand-side variables associated with school attendance.

6.1 POTENTIAL BIAS OF SAMPLE

The specificities of variable creation and subsequent sample selection of cases of which parental education variables were available, may have created bias in the results of this research. Variables for mother's and father's education could only be constructed for individuals whose both parents were alive and included in the sample, lived in the same household, and had their years of education documented. For approximately half of all cases aged 15 to 17 in the NFHS-3 sample, parental education variables could not be computed and thus were not included in the analyses (tables 3 and 6). It may be that the removal of these cases caused a systematic error that prejudiced the research findings in a certain direction.

For example, in the sample with missing cases removed, school attendance was much higher than school attendance for all 15-17 year olds in the NFHS-3 sample. Theoretically, it may be possible that the school attendance of children who have lost one or more parents may be impeded by the loss of parental supervision. Through the mechanics of the variable creation of parental education, these cases were excluded from the sample and may have biased the sample in favor of those attending school.

Since parental education variables have proved to be substantial determinants of the school attendance of offspring, future population surveys are strongly recommended to record mother's and father's education as separate variables to prevent potential sampling bias.

6.2 CALL FOR MORE RECENT DATA

This research used data from the Indian academic year 2005-06, ever since when large changes have taken place in India's educational system.

As the RTE and RMSA Acts came into effect in 2009-10, and the gross secondary enrollment rose from 56.1 in 2006 to 68.5 in 2011, with enrollment and attendance rates presumed to have further increased over recent years. However, World Bank statistics point out that India's secondary education enrollment, relative to other emerging economies, has barely improved - if not worsened (World Bank, n.d.).

Moreover, due to the nature of large-scale population surveys, entailing an inherently extensive period of time between data collection and publication, it is not uncommon for analyses of (Indian) school attendance to be based on data from a previous decade.

In 2002, Jayachandran based her research on the determinants of inter-district variations in school attendance rates by children in the 5-14 age-group on Census data for 1981 and 1991. Kingdon's 2007 analysis of school attendance rates was based on previous editions of the National Family Health Survey, the NFHS-1 (1993) and NFHS-2 (1999).

Results from the second round of the Indian Human Development Survey (IHDS-II), scheduled to be released in June 2015, and the 2016-2017 release of NFHS-4 data may shed light on more recent developments in the attendance of the Indian adolescent.

6.3 EXPLORATION OF SUPPLY-SIDE ISSUES

Through an analysis of the demand-side variables associated with ASA, this research has explained over one third of school attendance variability. Another important part of the variability in school attendance is explained by supply-side factors, including school accessibility, infrastructure, teacher absenteeism, and teacher-pupil ratios (Jayachandran, 2002; Kingdon, 2007). While socio-economic household variables can be, and have been recorded as isolated variables on an individual level, supply-side variables are not as straightforwardly quantifiable in the context of their effects on individual school participation. Due to the limited availability of isolated supply-side variables, Jayachandran included a variable for school accessibility, which may have functioned as a proxy to capture various supply-side effects collectively. While a variable for teacher-pupil ratios was not significant, a highly significant, positive coefficient for school accessibility reinforced the vital effect of school availability and infrastructure on school attendance (Jayachandran 2002).

This research adds on to the current body of research on demand-side determinants of school attendance, and stresses the need for a more rigorous and exhaustive exploration of the variation in school attendance caused by supply-side factors. For this purpose, a stronger link between population surveys and school quality reports needs to be established. Only when a comprehensive picture is painted of an adolescent's socio-economic background, and the quality and accessibility of education in his or her direct environment, can the central deterrents of attendance be adequately identified.

7. CONCLUSION

Several important confirmations and lessons emerge from the results presented in this paper. With high economic returns on secondary education, theoretically, adolescent demand for education should be high. Nevertheless, the results suggest that adolescent school participation is hindered by a combination of factors, including: conservatism about gender roles, residence, religion, parental education, household size and credit-constraints, and a constrained supply of high-quality education.

Wealth continues to be an important explanatory factor of school attendance. Less affluent families are unable to send their adolescent children to school because of their inability to afford the direct costs of schooling, or to bear the indirect opportunity costs of their children's wages. The growing proportion of private education is placing an extra burden on poorer families; as experienced and highly educated teachers are finding higher wages in private institutions, the quality of relatively affordable public education is jeopardized. The World Bank's recommendation to provide financial and in-kind assistance for disadvantaged students may prove fruitful. By offsetting direct and indirect costs of schooling, families can surpass disinclination to send their offspring, girls in particular, to school (World Bank, 2009).

Lower Muslim school attendance can be partly explained by smaller returns to education for Muslims, relative to other religious groups (Unni, 2007). Yet, discrimination on the supply side of education also seems to play an important role in limiting demand for education from the minority Muslim group (Human Rights Watch, 2014). Hindu and Muslim females are negatively impacted by the combination of their gender and religion. Certain values within Hinduism and the Islam may emphasize a more conservative female role, leading to increased household work and child labor, early marriage and child birth, and female school absenteeism (Parveen, n.d.; Unni, 2008).

Furthermore, the results suggest that educated parents are more capable of understanding the positive ramifications of education and in turn encourage school attendance of their adolescent children. Educated mothers are understood to particularly relate to the added value of education for women, and, hence, want to ensure the school attendance of their daughters. Household size negatively affects female attendance; with larger numbers of family members, daughters are required to stay at home to carry out household chores, whilst their brothers are being educated.

The effects of rural residence and caste/tribe status were among the surprising outcomes of this research. Controlling for household wealth and schooling of both parents, rural habitation turned out to have a positive effect on ASA, as opposed to the hypothesized negative effect, whereas female adolescents in rural areas did prove to be negatively affected by the combination of their gender and residence. Cultural norms that favor men, or the supply-side issue of unsafe home-to-school journeys may play a role in rural areas. Additionally, after controlling for wealth and parental education, caste or tribe status did not significantly impact adolescent participation in schools. These unexpected results should be explored in the light of recent interventions that promote educational opportunities for children from backward communities, such as the SSA, RTE, and the RMSA.

The overall results indicated a gender bias in adolescent school attendance, to the advantage of male participation. The RMSA focuses on resolving supply-side issues associated with secondary school quality and infrastructure. However, as this research has shown, an important part of ASA variability is not explained by institutional justifications, but by various socio-economic factors that constrain the demand for education. Indian adolescent demand for schooling could be encouraged using public information campaigns to transform attitudes about the benefits of education and delayed marriages. Other methods of promotion could be investments in curriculum alterations, progressive pedagogy, technology, and examination reforms (World Bank, 2009).

Future explorations of the comprehensive range of explanatory factors of ASA should consider supply-side issues alongside household demand constraints. These analyses would benefit from a solid connection between school quality reports and upcoming population surveys. Results from the forthcoming IHDS-II and NFHS-4 will provide crucial findings for mapping out not only the future of India’s youth, but of India’s potential as a budding hegemony.

REFERENCES

Bing, W.K., Goldschmidt, P., Boscardin, C. K. & Azam, M. (2011). Girls in India: Poverty, location, and social disparities, in: Lewis, Maureen & Marlaine Lockheed (Eds.), Exclusion, Gender and Education: Case Studies from the Developing World (pp. 119–143). Washington, Center for Global Development. Retrieved from http://www.cgdev.org/doc/books/lewis-lockheed-eduCaseStudies/lewis-lockheed-chapter5.pdf

Borooah, V.K., Iyer, Sr. (2005). Vidya, Veda, and Varna: The influence of religion and caste on education in rural India. The Journal of Development Studies, 41(8), 1369–1404.

Burke, J. (2011). Census reveals that 17% of the world is Indian. The Guardian. Retrieved from www.theguardian.com/world/2011/mar/31/census-17-percent-world-indian

Chauhan, C.P.S. (2008). Education and caste in India. Asia Pacific Journal of Education, 28(3), 217–234.

Department of School Education & Literacy. Ministry of Human Resource Development, Government of India. (n.d.-a). Rashtriya Madhyamik Shiksha Abhiyan. Retrieved from http://mhrd.gov.in/rmsa

Department of School Education & Literacy. Ministry of Human Resource Development, Government of India. (n.d.-b). Right to Education. Retrieved from http://mhrd.gov.in/rte

Department of School Education & Literacy. Ministry of Human Resource Development, Government of India. (n.d.-c). Sarva Shiksha Abhiyan. Retrieved from http://mhrd.gov.in/sarva-shiksha-abhiyan

Desai, S. B., Dubey, A., Joshi, B. L., Sen, M., Shariff, A., & Vanneman, R. (2010). Human Development in India: Challenges for a Society in Transition. Retrieved from http://ihds.umd.edu/IHDS_files/HumanDevelopmentinIndia.pdf

Dhawan, H. (2013). Rising cost of education worries parents, survey shows. The Times of India. Retrieved from http://timesofindia.indiatimes.com/india/Rising-cost-of-education-worries-parents-survey-shows/articleshow/17946981.cms

DHS Program. (2006). India: Standard DHS, 2005-06. Retrieved from www.dhsprogram.com/data/dataset/India_Standard-DHS_2006.cfm?flag=0

EdCIL. (2012). Inclusion and Exclusion of Children in the School and in the Classroom in Primary and Upper primary schools. Retrieved from http://ssa.nic.in/research-studies-document_old/survey-report-on-out-of-school-children/list-of-studies/National%20Synthesis%20Inclusion-Exclusion%20study.pdf

Fazal, T. (2013). Millenium Development Goals and Muslims of India (Working Paper Series XIII). Oxfam India. Retrieved from http://www.oxfamindia.org/sites/default/files/13%20Millennium%20Development%20Goals%20and%20Muslims%20of%20India.pdf

Human Rights Watch. (2014). “They Say We’re Dirty”: Denying an education to India’s marginalized. Human Rights Watch. Retrieved from http://www.hrw.org/sites/default/files/reports/india0414_ForUpload_1.pdf

IIPS and Macro International. (2007a). National Family Health Survey (NFHS-3), 2005–06: India: Volume I. Retrieved from http://www.dhsprogram.com/pubs/pdf/FRIND3/FRIND3-Vol1AndVol2.pdf

IIPS and Macro International. (2007b). National Family Health Survey (NFHS-3), 2005–06: India: Volume II. Retrieved from http://www.dhsprogram.com/pubs/pdf/FRIND3/FRIND3-Vol1AndVol2.pdf

IIPS and Macro International (2014). National Family Health Survey (NFHS-4): Report of pretest. Indo-Asian News Service. (2014). India has 37% of world's illiterate adults. The Times Of India. Retrieved from timesofindia.indiatimes.com/india/India-has-37-of-worlds-illiterate-adults/articleshow/29562827.cms

Jayachandran, U. (2002). Socio-economic determinants of school attendance in India (No. 103). Centre for Development Economics, Delhi School of Economics. Retrieved from http://www.cdedse.org/pdf/work103.pdf

Kingdon, G. G. (2002). The gender gap in educational attainment in India: How much can be explained? The

Journal of Development Studies, 39(2), 25–53.
Kingdon, G. G. (2005). Where has all the Bias Gone? Detecting Gender Bias in the Intra-household Allocation of Educational Expenditure in Rural India. Economic Development and Cultural Change, 53(2), 409–52.

Kingdon, G. G. (2007). The progress of school education in India. Oxford Review of Economic Policy, 23(2), 168-195.

Kingdon, G. G., and Unni, J. (2001). Education and Women's Labour Market Outcomes in India. Education Economics, 9(2), 173–95.

Nambissan, G. B. (1996). Equity in Education? Schooling of Dalit Children in India. Economic and Political Weekly, 31(16/17), 1011-1024.

National University of Educational Planning & Administration. (2008). Status of Education in India: National Report. Department of Higher Education, Government of India. Retrieved from http://www.ibe.unesco.org/National_Reports/ICE_2008/india_NR08.pdf

Nayak, D.K. & Behera, R.N. (2014). Changing Household Size in India: An Inter-State Comparison. Transactions of the Institute of Indian Geographers, 36(1), 1–18.

Parveen, I. (n.d.). Exclusion of Muslim Girls from Schools: A participatory analysis in the district of Rampur. Children Rights and You (CRY). Retrieved from http://www.cry.org/resources/pdf/ncrrf/ncrrf_reportby_iram_praveen.pdf

Pathak, S. & Gupta, A. (2013). Status of Women in India with Particular Reference to Gap in Male Female Literacy Rate in India. International Journal of Environmental Engineering and Management, 4(6), 549-552.

Registrar General, India. (n.d.). Distribution of Population by Religions: Drop-in-Article on Census 2011 – No.4. Office of the Registrar General Census Commissioner, India. Indian Census Bureau. Retrieved from http://censusindia.gov.in/Ad_Campaign/drop_in_articles/04-Distribution_by_Religion.pdf

Registrar General, India. (2011a). A - 5 Union Primary Census Abstract – 2011. Office of the Registrar General Census Commissioner, India. Indian Census Bureau. Retrieved from http://www.censusindia.gov.in/2011census/hlo/pca/pca_pdf/PCA-CRC-0000.pdf

Registrar General, India. (2011b). Census of India 2011: provisional population totals-India data sheet. Office of the Registrar General Census Commissioner, India. Indian Census Bureau. Retrieved from http://censusindia.gov.in/2011-prov-results/data_files/india/Final_PPT_2011_chapter6.pdf

SOS Children's Villages. (2013). Why are 39% of the world's illiterate adults found in India? — SOS Children. Retrieved 19 May 2015, from http://www.soschildrensvillages.org.uk/news/international-literacy-day/why-are-forty-percent-of-the-worlds-illiterate-adults-found-in-india

United Nations Development Programme. Human Development Reports. (2013). Mean years of schooling (of adults) (years). Retrieved from hdr.undp.org/en/content/mean-years-schooling-adults-years

Times News Network. (2010). Will RTE fulfil the SSA dream?. The Times Of India. Retrieved from http://timesofindia.indiatimes.com/home/education/news/Will-RTE-fulfil-the-SSA-dream/articleshow/5761551.cms

Unni, J. (2007). Earnings and Education among Ethnic Groups in Rural India (Working Paper Series No. 79). Retrieved from National Council of Applied Economic Research website: www.ncaer.org/free-download.php?pID=89

Unni, J. (2008). Are Gender Differentials in Educational Capabilities Mediated through Institutions of Caste and Religion in India? (Working Paper No. 186). Retrieved from Gujarat Institute of Development Research website: www.gidr.ac.in/pdf/WP-186.pdf

United Nations, General Assembly. (2000). United Nations Millennium Declaration (A/RES/55/2). Retrieved from http://www.un.org/millennium/declaration/ares552e.htm

UNICEF (2011). The situation of children in India: A profile. Retrieved from http://www.unicef.org/sitan/files/SitAn_India_May_2011.pdf

United Nations. (n.d.-a). Goal 2: Achieve Universal Primary Education. Retrieved February 22, 2015, from http://www.un.org/millenniumgoals/education.shtml
United Nations. (n.d.-b). Goal 3: Promote Gender Equality and Empower Women. Retrieved February 22, 2015, from http://www.un.org/millenniumgoals/gender.shtml

World Bank. (2009). Secondary Education in India: Universalizing Opportunity. Retrieved from http://datatopics.worldbank.org/hnp/files/edstats/INDstu09a.pdf

World Bank. (n.d.). School enrollment, secondary (% gross). Retrieved February 24, 2015, from http://data.worldbank.org/indicator/SE.SEC.ENRR/countries

Zachariah, M. (1972). Positive Discrimination in Education for India's Scheduled Castes: A review of the problems, 1950-1970. Comparative Education Review, 16(1), 16-29.

APPENDICES

APPENDIX A: RESEARCH SAMPLE INCLUDING MISSING DATA

Table 6: Demographic characteristics of 15-17 year olds at the start of the 2005-06 school year within the NFHS-3 sample

	Male (n = 16520)		Female (n = 16733)		Total (n = 33253)	
Variables	Frequency	Percentage	Frequency	Percentage	Frequency	Percentage
School attendance						
No attendance	7657	46.3	9373	56.0	17030	51.2
Total attendance	8576	51.9	7029	42.0	15605	46.9
Primary	178	1.1	137	.8	315	.9
Secondary	7724	46.8	6113	36.5	13837	41.6
Higher	669	4.0	776	4.6	1445	4.3
Missing	292	1.8	334	2.0	626	1.9
Residence						
Urban	7642	46.3	7038	42.1	14680	44.1
Rural	8878	53.7	9695	57.9	18573	55.9
Missing	0	.0	0	.0	0	.0
Religion						
Hindu	11556	70.0	11566	69.1	23122	69.5
Muslim	2643	16.0	2824	16.9	5467	16.4
Other	2316	14.0	2340	14.0	4656	14.0
Missing	5	.0	3	.0	8	.0
Caste or tribe type						
Scheduled Caste	2918	17.7	2991	17.9	5909	17.8
Scheduled Tribe	2301	13.9	2448	14.6	4749	14.3
Other	10622	64.3	10575	63.2	21197	63.7
Missing	679	4.1	719	4.3	1398	4.2
Literate father						
Yes	5819	35.2	5372	32.1	11191	33.7
No	2228	13.5	2038	12.2	4266	12.8
Missing	8473	51.3	9323	55.7	17796	53.5
Literate mother						
Yes	4108	24.9	3897	23.3	8005	24.1
No	4650	28.1	4402	26.3	9052	27.2
Missing	7762	47.0	8434	50.4	16196	48.7
Father's education in years						
0	2234	13.5	2043	12.2	4277	12.9
1 – 5	1619	9.8	1469	8.8	3088	9.3
6 – 10	2773	16.8	2557	15.3	5330	16.0
11 +	1413	8.6	1328	7.9	2741	8.2
Missing	8481	51.3	9336	55.8	17817	53.6
Mother's education in years						
0	4677	28.3	4409	26.3	9086	27.3
1 – 5	1354	8.2	1355	8.1	2709	8.2
6 – 10	1997	12.1	1843	11.0	3840	11.6
11 +	9476	4.4	684	4.1	1412	4.3
Missing	7771	47.0	8442	50.5	16213	48.8

	Male (n = 16520)		Female (n = 16733)		Total (n = 33253)		
Variables	Frequency	Percentage	Frequency	Percentage	Frequency	Percentage	
Wealth Index							
Poorest	1858	11.2	2098	12.5	3956	11.9	
Poorer	2490	15.1	2636	15.8	5126	15.4	
Middle	3445	20.9	3591	21.5	7036	21.2	
Richer	4174	25.3	4012	24.0	8186	24.6	
Richest	4553	27.6	4396	26.3	8949	26.9	
Missing	0	.0	0	.0	0	.0	
Number of household members							
0 – 5	8698	52.7	7791	46.6	16489	49.6	
5 – 10	6886	41.6	7870	47.0	14756	44.4	
11+		36	5.7	1072	6.4	2004	6.0
Missing	0	.0	0	.0	0	.0	

Source: 2005–6 Household data from the National Family Health Survey (DHS Program, 2006).

APPENDIX B: LR MODELS OF DIRECT EFFECTS – ADDING PARENTAL EDUCATION & WEALTH VARIABLES

Table 7: Logistic Regression analysis of school attendance: adding parental education & household wealth variables: – Parameter estimates (B), Standard Errors (SE) and Odds Ratios from the logistic regression of school attendance for children aged 15-17, India, 2005-06

	Model B1			Model B2			Model 1 as presented in table 5		
						Odds			Odds
Odds									
Variable	B	SE (B)	Ratio	B	SE (B)	Ratio	B	SE (B)	Ratio
Constant	2.914***	.186	–	.775***	.202	–	–.875***	.224	–
Gender									
Female	–.291***	.037	.748	–.369***	.041	.691	–.398***	.041	.671
(base = male)									
Residence									
Rural	–.762***	.041	.467	–.085		.918	.409***	.055	1.506
(base = urban)									
Religion									
Hindu	–.318***	.082	.728	–.072	.091	.930	.023	.092	1.024
Muslim	–1.556***	.096	.211	–.930***	.106	.395	–.875***	.108	.417
(base = other/ no religion(s))									
Caste type									
Scheduled Caste	–.632***	.049	.532	–.162**	.053	.850	–.061	.055	.941
Scheduled Tribe	–.593***	.072	.553	–.120	.077	.887	.058	.079	1.059
(base = other/ no caste/tribe)									
Number of household members									
	–0.16*	.007	.984	.010	.008	1.010	–.003	.008	.997
Father's education in years	.								
				134***	.005	1.143	.106***	.006	1.112
Mother's education in years									
				.138***	.007	1.148	.114***	.007	1.120
Wealth Index							.423***	.023	1.527
Nagelkerke pseudo R2	16.3%			33.9%			36.4%		
X2	1778.6, df = 35,			4005.3, df = 37,			4352.8, df = 38,		
	p<0.0001			p<0.0001			p<0.0001		

*p<0.05; **p<0.01, ***p<0.0001.

Note: The dependent variable in this analysis is school attendance in the 2005-06 school year, coded so that 0 = did not attend, and 1 = did attend school.
To account for interstate differences between India's 29 states shown in table 8 (Appendix C), 28 dummy variables have been used in all models, for which:
Model 1: Wald-stat = 658.394, df = 28, p < .0001
Model 2: Wald-stat = 554.454, df = 28, p < .0001
Model 3: Wald-stat = 517.688, df = 28, p < .0001

Source: 2005–6 Household data from the National Family Health Survey (DHS Program, 2006).

APPENDIX C: ESTIMATED COEFFICIENTS OF STATE DUMMY VARIABLES

Table 8. Control variable 'State': Estimated coefficients (B), SE(B), and Odds Ratios of state dummy variables in models 1 and 2 (table 5)

	In model 1			In model 2		
States	B	SE (B)	Odds Ratio	B	SE (B)	Odds Ratio
Himachal Pradesh	.014	237	1.014	.027	.239	1.027
Punjab	-1.329***	.201	.265	-1.347***	.202	.260
Uttaranchal	-.551**	.204	.577	-.554**	.206	.575
Haryana	-.979***	.202	.376	-1.001***	.204	.367
Delhi	-.991***	.203	.371	-.995***	.204	.370
Rajasthan	-1.170***	.189	.310	-1.214***	.191	.297
Uttar Pradesh	-1.041***	.172	.353	-1.052***	.173	.349
Bihar	-.682**	.207	.505	-.706**	.209	.493
Sikkim -.049	.251	.952	-.043	.253	.957	
Arunachal Pradesh	.798**	.279	2.220	.783**	.281	2.187
Nagaland	-.400	.224	.670	-.417	.226	.659
Manipur	-.273	.218	.761	-.283	.219	.753
Mizoram	-1.826***	.262	.161	-1.783***	.264	.168
Tripura	.091	.237	1.095	.089	.239	1.093
Meghalaya	-.343	.238	.710	-.338	.239	.713
Assam -.428*	.211	.652	-.449*	.212	.638	
West Bengal	-.768***	.186	.464	-.796***	.188	.451
Jharkhand	-.381	.210	.683	-.428*	.212	.652
Orissa -1.847***	.196	.158	-1.877***	.198	.153	
Chhattisgarh	-.814***	.192	.443	-.825***	.194	.438
Madhya Pradesh	-.856***	.185	.425	-.883***	.186	.414
Gujarat	-2.053***	.196	.128	-2.072***	.198	.126
Maharashtra	-1.031***	.178	.357	-1.054***	.180	.349
Andhra Pradesh	-.852***	.181	.427	-.889***	.183	.411
Karnataka	-.663***	.188	.515	-.675***	.190	.509
Goa	-.577*	.242	.562	-.602*	.244	.548
Kerala .084	.260	1.087	.089	.262	1.093	
Tamil Nadu	-.796***	.196	.451	-.814***	.198	.443

(base = Jammu and Kashmir)

*p<0.05; **p<0.01, ***p<0.0001

Note. The state coefficients are in contrast with the state Jammu and Kashmir.
Source: 2005–6 Household data from the National Family Health Survey (DHS Program, 2006).

Unravelling Genetic Pathways Controlling Membrane Trafficking and the Formation of a New Endomembrane Compartment in Petunia

Analysis of Genes Regulated by the Transcription Factors AN1, PH3 and PH4

Eva van der Heijden

ABSTRACT

The Petunia is a flower that exists in many colours. For many years, it has been used as a model to study processes involved in flower colouration. At the genetics lab at SILS (Swammerdam Institute of Life Science), where I did my bachelor research, the researchers were looking into three genes (AN1, PH3, PH4) known to be involved in flower colouration when they discovered a new process inside the flower cell that is regulated by these three genes. They found a compartment that they called 'vacuolino'. Every cell contains a structure called the central vacuole, which is a large membrane compartment filled with water where molecules are stored. The vacuolino is smaller than the central vacuole, and it is involved in transporting proteins from their site of production to the central vacuole. The vacuolino is not present in flowers that have a non-functional version of AN1, PH3, or PH4, whereas it is present in normal flowers. Not much is known about this process yet, so the goal of my research was to look into all genes that are regulated by AN1, PH3, and PH4, which should lead to more information on the formation of vacuolinos. This was done by comparing the gene expression of normal flowers and the gene expression of flowers with mutations in AN1, PH3, and PH4, respectively. When one of the genes has a mutation, all the genes that are regulated by the protein this gene codes for will be differently expressed compared to a flower without mutations. Since a normal flower has a vacuolino, and a mutant flower does not have a vacuolino, the formation of the vacuolino has to be caused by the genes that are differently expressed in both flowers. By comparing the expression levels in normal flowers and mutant flowers, 41 genes were identified that are likely to be involved in the formation of the vacuolino. They were examined using databases and via some experiments in the lab. Genes involved in many processes in the cell, amongst others in protein transportation, were identified. This study should give more insight into the functioning of pathways regulated by AN1, PH3, and PH4 and complement other research that is being done. Based on the findings in this thesis, new possible directions for research may open. Improving our knowledge of plant development can be important for agricultural and commercial purposes, such as plant breeding and the flower industry.

1. INTRODUCTION

Petunia hybrida is a flowering plant species that can produce flowers of several colours, including purple, red, pink, white or blue. Additionally, flowers can have combinations of colours simultaneously or at different moments in their development. Flowers can, for instance, lose their colour a couple of days after bud opening during a process called 'fading' (Quattrocchio et al., 2006). Flower colour plays a role in the attraction of pollinators; therefore it is important in the reproduction strategy of the plant. The colour is determined by various factors such as the presence of certain flavonoids (e.g. anthocyanin pigments) and the pH in the vacuole of the plant cell where the pigments accumulate. Mutations that influence the anthocyanin biosynthesis pathway can be observed easily, so these phenomena in Petunia-species have been a subject of research for several decades. Since pigments are accumulated in the vacuolar lumen and the colour of the anthocyanin pigments is pH-dependent, flower colour can also be used as an indicator of the acidification of the vacuole in petal cells. The analysis of flower pigmentation can thus give insight into several areas such as intracellular transport of proteins, regulation of gene expression and – through comparison of different species – evolution of biochemical pathways and regulatory networks (Koes et al., 2005).

Several transcription factors involved in anthocyanin pigmentation and vacuole acidification have already been identified and characterized. ANTHOCYANIN1 (AN1) is a transcription factor involved in the regulation of pigmentation in Petunia tissues because it acts on pathways concerning synthesis of anthocyanin pigments, acidification of vacuoles in petal cells, and seed coat development (Spelt et al., 2002). AN1 interacts with several other transcription factors. One of these is a transcription factor called PH4. Together, AN1 and PH4 regulate several genes; one gene they regulate codes for the transcription factor PH3 (Quattrocchio et al., 2006). It has been shown that PH3, PH4 and AN1 form a complex that regulates several genes involved in vacuolar acidification, as well as a recently discovered protein-sorting pathway involving small organelles called vacuolinos (unpublished data: Faraco M, 2011).

The main objective of this research is to analyse a set of mRNA sequences whose expression is regulated by AN1, PH3 and PH4. This should provide more insight into the pathways regulated by these transcription factors and complement other research that is being done. Analysis will be carried out with the aid of tools like BLAST in the Petunia and NCBI databases, phylogenetic trees, and literature review concerning the proteins encoded by the candidate genes.

When the aforementioned transcription factors are mutated, the pathways they influence do not work properly anymore. Much knowledge about these pathways already exists, but it is not complete yet. The pathway of the vacuolino has been discovered only recently, so there is still a lot to learn about it (unpublished data: Faraco M, 2011). Since the RNAseq analysis will disclose all genes controlled by AN1, PH3, and PH4, this will be relevant for the study of all pathways regulated by these transcription factors. This study is part of a larger research initiative; some members of the research group are doing detailed research into specific molecules that are influenced by AN1, PH3, and PH4, whereas others look into interactions between them. All the research is in some way related to flower pigmentation in Petunia, acidification of the vacuole, and the pathway involving vacuolinos in flower cells. Based on the findings in this thesis, new possible directions for research may open. Improved knowledge of vacuole acidification, flower pigmentation, and vacuolino protein trafficking can aid in providing insights concerning plant development. The pathways present in Petunia can be compared to pathways in other species, and this will provide information into the evolution of those pathways. Improving our knowledge of plant development can be important for agricultural and commercial purposes such as plant breeding and the production of new colours for the flower industry.

2. LITERATURE REVIEW

Petunia is a flowering plant that can produce flowers in several colours. The colour of the flowers is determined by the presence of certain anthocyanins (flower pigments), co-pigments, and the pH within the vacuole of the plant cell. For the biosynthesis of the anthocyanin pigments, several genes (e.g. AN-genes) are necessary, and when some of those genes are mutated, flower colour is affected. Roses, for example, lack flavonoid 3'5'-hydroxylase and are thus not able to synthesize purple delphinidin derivatives, whereas Petunia is able to produce these pigments (Mol et al., 1998)(see Figure 1). Co-pigments such as flavonols and flavones can cause a shift in colour when accumulated together with anthocyanins. Metal-ions that can form complexes with anthocyanin-molecules can change the flower colour as well. PH-genes (PH1 – PH7) are involved in the acidification of the vacuole. When the pH of the vacuole is increased, the flower turns somewhat blue (Verweij et al., 2008) due to the pH indicator properties of the anthocyanin molecules. Fading of colour can occur when a mutation of PH4 is present in the genetic background together with a dominant allele for a gene called 'FADING' (de Vlaming et al., 1982).

One of the AN-genes, namely ANTHOCYANIN1 (AN1), codes for a transcription factor with a helix-loop-helix domain that is not only involved in the regulation of pathways involved in synthesis of anthocyanin pigments, but also in the acidification of vacuoles in petal cells, seed coat development, and probably much more (Spelt et al., 2002). When AN1 is mutated, all structural anthocyanin genes are not expressed. One of these is called dihydroflavonol reductase (DFR), which was demonstrated to be transcriptionally activated by AN1 directly (Spelt et al., 2002). This causes loss of colour because all structural anthocyanin genes are necessary for the production of anthocyanins (see Figure 1). AN1 controls anthocyanin pigmentation by interaction with several proteins, including AN2 (a MYB-domain containing protein), AN11 (WD40-protein), and JAF13 (HLH-containing protein). AN1 controls acidification by another set of interactions, namely with JAF13 and PH4 (another MYB-domain containing protein) (Quattrocchio et al., 2006; Koes et al., 2005) (see Figure 2).

Figure 1 shows an overview of the biosynthesis pathway of flavonoids. The pathway begins with the multi-step conversion of L-phenyalanine to dihydrokaempferol. Dihyrokaempferol can be converted into several products by different steps that involve other enzymes. For example, the enzyme F3'5'H (flavonoid 3' 5' hydroxylase) leads to the production of delphinidins. AN-genes regulate that pathway (e.g. DFR, which is regulated by AN1, catalyses the reaction from dihydromyricetin to leucodelphinidin). Other enzymes of the pathway lead to the production of flavonols or proanthocyanidins.

PH4 is a transcription factor involved in the vacuolar acidification pathway, but it does not influence anthocyanin production (Quattrocchio et al., 2006). As mentioned above, PH4 is a member of the MYB-family and interacts with AN1. The complex containing PH4 and AN1 regulates the transcription of PH3. PH3 itself is also a transcription factor (of the WRKY-family) involved in vacuolar acidification; it takes part in the complex containing AN1 and PH4. Genes that are regulated by the AN1-PH4-PH3-complex are, amongst others, PH1 and PH5, which encode for a H^+ pump of the family of P3A-ATPases located on the vacuolar membrane and a protein of the P_{3B}-ATPase family resembling Mg^{2+} transporters, respectively (Verweij et al., 2008).Together, PH1 and PH5 are necessary and sufficient to acidify the vacuole; it is thought that they form a heterodimer involved in transporting H+ across the vacuolar membrane to acidify the vacuolar lumen.

PH3 and PH4 are also involved in a newly discovered protein-sorting pathway consisting of vacuolinos (small vacuoles). A vacuolino is a small (vacuolar) intermediate endomembrane compartment present next to the central vacuole in epidermal petal cells that sorts proteins such as PH5 to the tonoplast of the central vacuole. It fuses to the central vacuole with the aid of SNARE-proteins PhSYP22 and PhSYP51, and with the aid of PH1. In mutants for the transcription factors PH3 and PH4, the vacuolino is absent and the canonical protein-sorting pathway is active like in leaf cells (unpublished data: Faraco M, 2011). Since a ph5-mutant still has the vacuolino, its absence is not due to a failure of vacuolar acidification since this can be caused by ph5 as well as ph3 and ph4-mutants. Additionally, a vacuolino is not required for vacuolar acidification, as acidification can still occur when PH3 and/or PH4 are mutated, and PH5 and PH1 are expressed from transgenes. PH1 and PH5 are necessary and sufficient for vacuolar acidification. Since PH1 and PH5 expression does not restore formation of vacuolinos, their formation depends on target genes of AN1, PH3, and PH4 that are distinct from PH1 and PH5 (unpublished data: Faraco M, 2011). There are two possible mechanisms leading to the vacuolino-pathway preferentially used in petal epidermal cells. First, the vacuolino-route might sort proteins to the vacuole in wild-type petals because it is more efficient, or second, AN1, PH3, and PH4 might repress the canonical pathway.

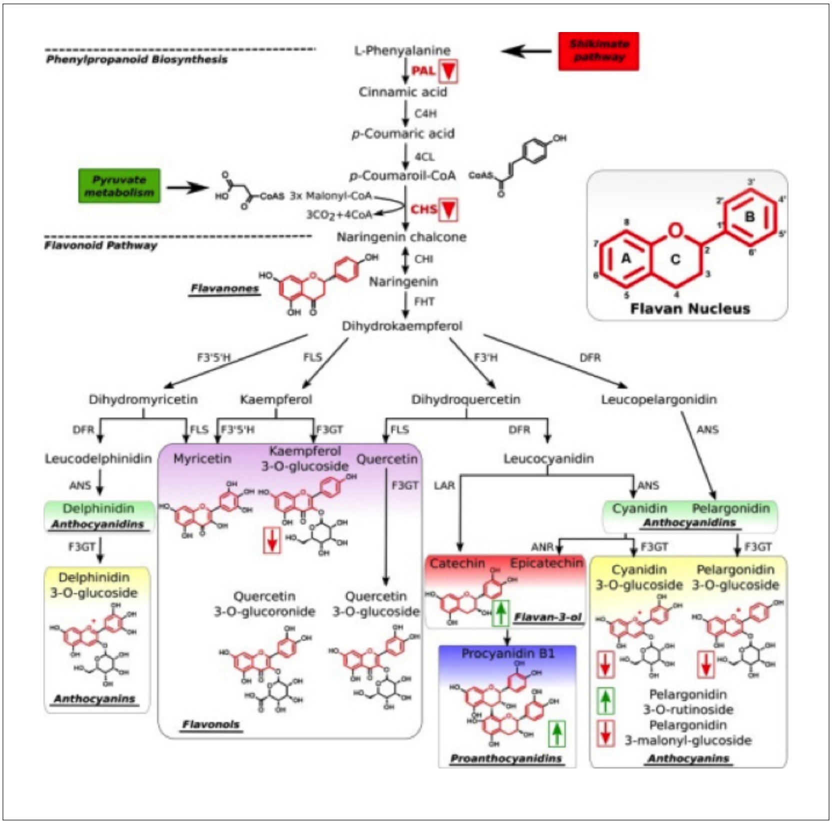


Figure 1.The flavonoid biosynthesis pathway (Casanal et al., 2013)

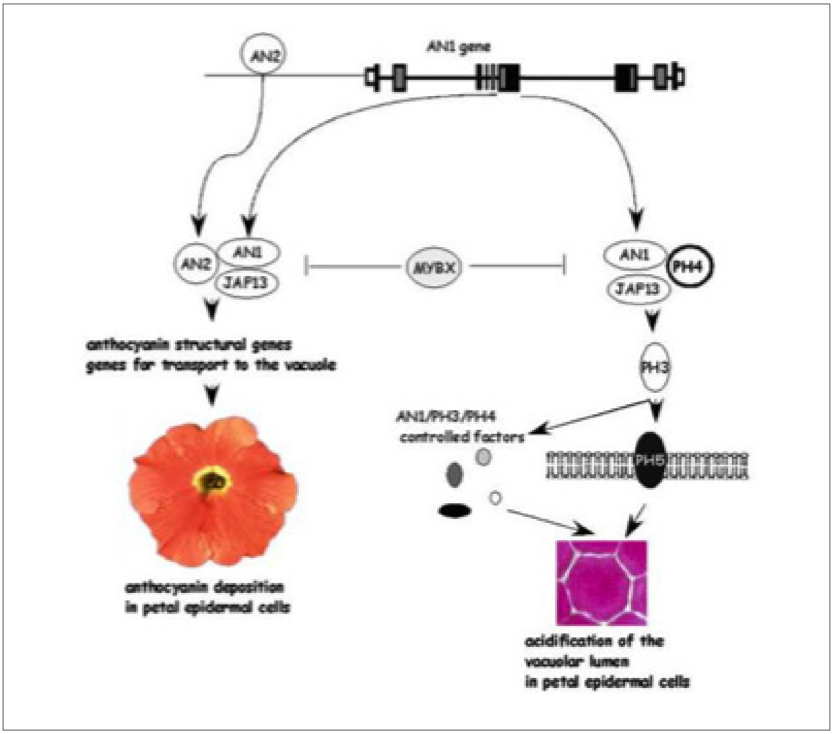


Figure 2. Transcription factors involved in regulation of anthocyanin deposition and acidification of vacuole (Gerats & Strommer, 2009)

The study of the target genes of AN1, PH3, and PH4 will shed light on the mechanism of vacuolino formation and the regulation of the protein sorting pathway that makes use of these compartments. For this reason, an RNAseq experiment has been performed to identify the genes differentially expressed in mutants for AN1, PH3, and PH4, and here I will present an analysis of these genes.

3. DATA & METHODS

3.1 OBTAIN RNA EXPRESSION DATA

During the analysis, I have looked at mRNAs that are significantly higher or lower expressed in an1-, ph3-, or ph4-mutant petals compared to a wild-type control. Petunia hybrida originates from crosses of the wild species Petunia inflata and Petunia axillaris. Within P. hybrida, there are several different lines, which are well characterized for their genetic background and are kept as pure lines in the collection of our group. For example, some have a non-functional AN1-gene due to a transposon insertion and others have a mutant allele of PH4, whereas others have a mutant PH3.

The mRNA sequences used in my analysis were obtained by the following experiment (which was not done by me, but by Yanbang Li). Six RNA samples were prepared: samples (1), (2), and (3) consisted of wild-type Petunia hybrida petals (line Rev W138, wt(ph4/ph7) and (wt)R167, respectively); sample (4) consisted of petals from a Petunia plant that had a mutant allele of AN1 due to the footprint left by a transposon insertion (strain W225); sample (5) petals had a ph4-mutant allele due to a transposon insertion (strain R162); sample (6) petals had a ph3-mutant allele due to a transposon insertion (R167). The first three samples all had red flowers, the fourth sample had white flowers, and the last two samples had purple/bluish flowers, which is characteristic of mutations affecting vacuolar pH in petals (see Figure 3). Sample (4) is genetically the same as sample (1) (isogenic), except for the non-functional AN1-gene, so sample (4) was compared to sample (1). Sample (2) is another wild-type, different from sample (1), but isogenic to sample (5), apart from the mutated PH4-gene. The same goes for sample (3) and (6). Due to their isogenic genetic background, these samples are the best material for comparison of gene expression in relation to the specific mutations we want to study. The RNAseq data was tested by looking for expression levels of known AN1-regulated mRNAs and by comparing the data to protein expression levels of an experiment in which an an1-mutant was compared to a wild-type.

3.2 RNA ANALYSIS

In a Petunia in which AN1 is not expressed, genes that are regulated by AN1 will be either more or less expressed, depending on whether they were negatively or positively affected. The same holds for ph3- and ph4-mutants. Expression levels of 35.237 genes were compared in the six samples. With the EDGE test, the P-value of the tagwise dispersion between the mutant and the wild-type was calculated, and the fold change was calculated as well. This was done for all three mutant groups. To make sure that only those mRNAs are analysed whose expression is significantly changed when either AN1, PH3, or PH4uj

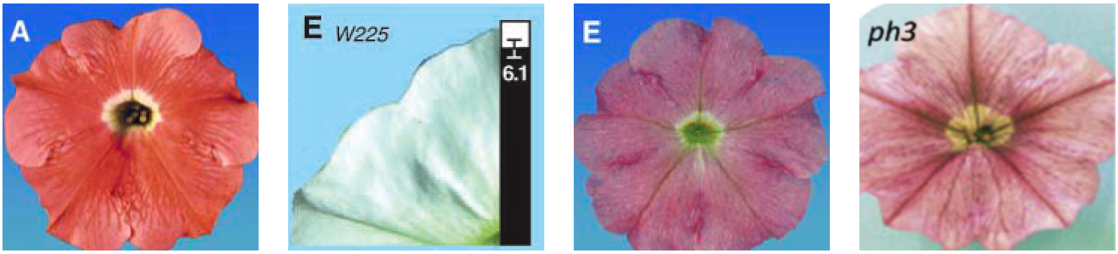


Figure 3A Figure 3B Figure 3C Figure 3D

Figure 3. Petunia flowers, wild-type and mutants

A. Wild-type; red-coloured, sample 1 - 3
B. an1-mutant; white-coloured, sample 4
C. ph4-mutant; purplish-coloured, sample 5
D. ph3-mutant; purplish-coloured, sample 6

are non-functional, a limit of $P < 0.05$ was set. On top of that, I have chosen to consider that the expression-level should be at least 3-fold increased or decreased compared to the petals where all transcription factors are normally expressed.

The mRNA sequences were divided into several groups, based on the mutants in which they are differentially expressed. There are thus groups consisting of mRNAs that are differentially expressed in all three mutants, in only two of the mutants, and in only one of the mutants. If a gene is regulated by AN1, and not by the other two transcription factors, significantly changed expression-levels will only be found in an an1-mutant; that mRNA will then be placed in the appropriate group. However, if a gene is regulated by all three transcription factors, its expression-levels will be changed in all mutants. Per group, one would like to know how many genes are up- or down-regulated and what the function of those genes is. The group to be considered most in-depth is the group consisting of genes that are regulated by all three transcription factors. In Appendix A, a short overview table of the other groups is given.

The Blast2GO program is used to get a first impression of the possible functions of the sequences (Conesa et al., 2005), and WEGO is used to visualize this first impression (Ye et al., 2006). For all the individual genes that were looked at, the Petunia gene identifier was used to search the Petunia database (Petunia SGN Cornell) for an mRNA transcript sequence. Subsequently, the mRNA transcript was blasted against the Petunia scaffold database to be able to see where exons and introns are located and to make sure there are no mistakes in the mRNA sequence. Occasionally, some nucleotides are missing in the mRNA sequence, or the start- or stopcodon is in the wrong place. The given mRNA sequence then translates into a shortened protein. The sequence obtained from comparing the mRNA with the scaffold is then used to translate to a protein sequence (ExPASY). The protein sequence is blasted against the NCBI database, to identify the closest sequence in that database (NCBI database)[2]. By using the close BLAST-hits, the protein sequence was further modified. For example, when there was a gap in the query sequence at a particular place for all hits, I looked for a missed exon in the genome of Petunia. All adapted sequences can be found in Appendix B. BLAST (Basic Local Alignment Search Tool) (Altschul et al., 1997; Altschul et al., 2005)[2] also looks for conserved domains in the protein, which can give more insight in the function and indicate the protein family to which a certain protein belongs (Domains NCBI)[1]. For further analysis, several phylogenetic trees were made ('One click' mode). For more information on genes, UniProt was often used (UniProt Consortium, 2015). Finally, the information obtained from the RNA analysis and from related literature is combined to get an impression of the pathways regulated by the transcription factors AN1, PH3, and PH4.

3.3 PCR

Primers were designed for Phy3, Phy4, Rab18b, and for the WAT1-related gene. For the WAT1-related gene, it was uncertain whether the coding sequence I found in the petunia database consisted of a single encoded protein or was resulting from an annotation mistake and consisted of two proteins. Of those, gene expression levels were checked with a PCR reaction. The primers used can be found in Appendix C.

To be able to perform a PCR reaction, cDNA has to be synthesized from the mRNA one wants to investigate. I had 24 mRNA samples (4 different stages per AN1 wild-type, an1-mutant, PH3 wild-type, ph3-mutant, PH4 wild-type, and ph4-mutant). First, the RNA had to be diluted to 2.5 µg in 11.5 µl sterile water. This had to be boiled for 1 minute and cooled down on ice. Then, for each sample, a mixture of 10 µl 5x RT Buffer, 5 µl water, 20 µl dNTPs, 3 µl oligo DT race primer, and 0,5 µl Reverse Transcriptase had to be added to the RNA. This had to incubate for 1 hour at 42°C. Last, 200 µl of sterile water was added.
With the cDNA samples, qPCR and PCR were performed. For a qPCR reaction, first 10 µl of forward and reverse primer have to be mixed with 80 µl water. Then 1 µl of this mixture has to be added to 5 µl of the qPCR mix and 4 µl of the template cDNA. This has to be pipetted in the PCR plate, which can then run. Via a fluorescent marker, it is possible to see at which cycle the cDNA starts to be duplicated, and thus how

---

1) I have not cited NCBI, BLAST, NCBI domains every single time in my text when I used BLAST to search through NCBI database for close BLAST-hits and domains in the protein, and information on those domains. They are cited in the text above, and appear in References.

much cDNA was originally present in the sample. With normal PCR and subsequent gel electrophoresis one can detect the presence of a particular gene in a sample based on the read-out of a gel. First, a particular gene fragment is amplified with PCR, and then the resulting product is run on a gel to check the presence of a product and its length.

4. RESULTS

4.1 TEST RESULTS

At the start of my work, I tested whether the results obtained from the RNAseq experiment described in 'data & methods' could reproduce earlier findings and already known effects on expression patterns. Several genes are known to be transcriptionally regulated by AN1, so in the experiment they are expected to be down in the an1-mutant, compared to the control. DFR, one of the key genes in the anthocyanin pathway is indeed lower, as well as PH3 and PH5. Several other genes involved in the anthocyanin-pathway, such as HF1 (flavonoid-3',5'-hydroxylase 1), HF2, AN9, 3GT, 5GT, and MYB27 are down as well (see Table 1A), as previously described in other work. Some genes involved in the anthocyanin pathway that were not significantly down in the RNAseq experiment can be found in Table 1B. As can be seen, AN1 itself is not significantly down in an an1-mutant. This is possible because RNAseq data looks at RNA-levels. Since the mutation in AN1 is caused by a transposon insertion, the mRNA can still be formed; however, it is not possible to form a functional AN1-protein.

Table 1A. Genes associated with anthocyanin pathway down-regulated in an1-mutant

Name	Gene ID	Function
3GT	Peaxi162Scf00163g08001.1	UDP-glucose: flavonoid 3-O-glucosyltransferase
5GT	Peaxi162Scf00378g01015.1	UDP-glucose: anthocyanin 5-O-glucosyltransferase
AN9	Peaxi162Scf00713g03008.1	Glutathione S-transferase; ANTHOCYANIN9
ANT17	Peaxi162Scf00620g05015.1	Leucoanthocyanidin dioxygenase
CHSJ	Peaxi162Scf00536g09002.1	Chalcone synthase J
DFR	Peaxi162Scf00366g06022.1	Dihydroflavonol 4-reductase
AAT	Peaxi162Scf00160g08037.1	Anthocyanin acetyltransferase
HF1	Peaxi162Scf00150g02019.1	Flavonoid 3',5'-hydroxylase 1
HF2	Peaxi162Scf00108g04001.1	Flavonoid 3',5'-hydroxylase 2
MF1	Peaxi162Scf00316g05019.1	Caffeoyl-CoA 3-O-methyltransferase 1
MF2	Peaxi162Scf00518g04034.1	Caffeoyl-CoA 3-O-methyltransferase 2
RT	Peaxi162Scf00487g06003.1	Anthocyanin rhamnosyltransferase
PH3	Peaxi162Scf00472g07010.1	WRKY family transcription factor
PH5	Peaxi162Scf00177g06021.1	Plasma membrane proton ATPase
MYB27	Peaxi162Scf03779g00019.1	MYB-domain protein (regulatory)
4CL_B	Peaxi162Scf00195g12023.1	4-coumarate:CoA ligase 2
GL2	Peaxi162Scf00015g00049.1	Homeobox-leucine zipper protein GLABRA 2
MybX	Peaxi162Scf00521g08011.1	MYB-domain protein (regulatory)

Table 1B. Genes associated with anthocyanin pathway not down-regulated in an1-mutant

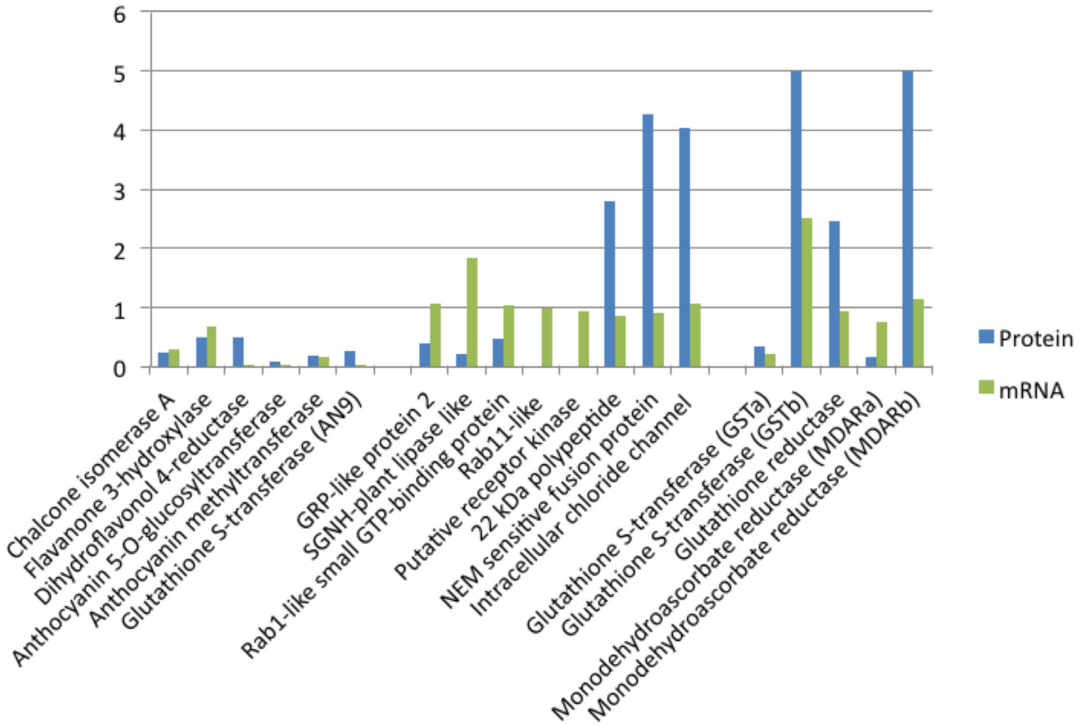
Name	Gene ID	Function	Significance*
AN1	Peaxi162Scf00338g09013.1	bHLH transcription factor; ANTHOCYANIN1	P = 3*10-4, f = -2.1
AN11	Peaxi162Scf00912g01010.1	WD40 transcription factor; ANTHOCYANIN11	P = 0.21, f = - 1.3
CHSA	Peaxi162Scf00047g12030.1	Chalcone synthase A	P = 0.21, f = - 1.4
F3H	Peaxi162Scf00328g12015.1	Naringenin, 2-oxoglutarate 3-dioxygenase	P = 0.06, f = - 1.5
JAF13	Peaxi162Scf00119g09045.1	bHLH transcription factor	P = 0.05, f = 1.6
HT1	Peaxi162Scf00201g02040.1	Flavonoid 3',5'-hydroxylase	P = 0.65, f = 1.2
MT	Peaxi162Scf00450g00026.1	Caffeoyl-CoA 3-O-methyltransferase	P = 0.38, f = 1.4
AN2	Peaxi162Scf00118g03011.1	ANTHOCYANIN2	P = 0.003, f = 1.4
MybB	Peaxi162Scf01210g00010.1	MYB-domain protein, ANTHOCYANIN4 (?)	P = 0.43, f = - 1.5
AN4-I	Peaxi162Scf00578g00006.1	ANTHOCYANIN4-I	P = 0.63, f = 1.7
AN4-II	Peaxi162Scf00578g00005.1	ANTHOCYANIN4-II	P = 1, f = - 1.1
PH1	No good hit in Petunia database		
PH4	Peaxi162Scf00349g05008.1	MYB-domain protein 5	P = 0.16, f = 1.3
ODO1	Peaxi162Scf00002g03010.1	MYB-domain protein 42	P = 0.02, f = - 2.2
FLS	Peaxi162Scf00927g03005.1	Flavonol synthase 1	P = 0.08, f = 1.7
C4H_A	Peaxi162Scf00556g03005.1	Trans-cinnamate 4-monooxygenase A	P = 0.09, f = 1.7
C4H_B	Peaxi162Scf00390g02028.1	Trans-cinnamate 4-monooxygenase B	P = 8*10-4, f = -1.9
4CL_A	Peaxi162Scf00314g08039.1	4-coumarate:CoA ligase A	P = 0.55, f = 1.2
4CL_C	Peaxi162Scf00610g03049.1	4-coumarate:CoA ligase C	P = 0.13, f = 1.6
4CL_D	Peaxi162Scf00207g03041.1	4-coumarate:CoA ligase D	P = 0.009, f = 1.9
PAL_A	Peaxi162Scf00858g02017.1	Phenylalanine ammonia-lyase A	P = 0.66, f = - 1.1
PAL_B	Peaxi162Scf00488g07004.1	Phenylalanine ammonia-lyase B	P = 0.06, f = 1.7
PAL_C	Peaxi162Scf00123g09008.1	Phenylalanine ammonia-lyase C	P = 0.12, f = 1.5
TT12	Peaxi162Scf00047g23038.1	MATE efflux family protein	P = 0.72, f = - 5.6
ANL2	Peaxi162Scf00021g09019.1	Homeobox-leucine zipper protein ANTHOCYANINLESS 2	P = 0.72, f = - 1.1
ANL2	Peaxi162Scf00233g06009.1	Homeobox-leucine zipper protein ANTHOCYANINLESS 2	P = 0.02, f = - 1.8
MybX	Peaxi162Scf00521g08011.1	MYB-domain protein (regulatory)	

*f = fold change

In a paper by Prinsi et al. (unpublished data), the authors examined the protein levels for several genes in a Petunia an1-mutant and compared them to an isogenic control wild-type. Not all of the proteins they found to be down were known in Petunia yet, but the paper gives the accession numbers of genes in other species. I looked for a similar pattern in our mRNA expression levels. Obviously, it is possible that post-transcriptional or post-translational regulation has occurred, so it is not always true that low protein expression levels have to be caused by low mRNA expression levels. Additionally, it is not true that high protein levels have to be caused by high mRNA levels because it could be a long-living protein.

Prinsi et al. used $\Delta = \frac{\text{(protein expression levels an1)}}{\text{(protein expression levels WT)}}$ to show the change in expression levels of an an1-mutant versus wild-type. My data used fold change $= \frac{\text{(mRNA expression levels WT)}}{\text{(mRNA expression levels an1)}}$ to show the change. So, to compare our results, I compared their Δ with our $\frac{1}{\text{(fold change)}}$. In Graph 1, the results of the comparison can be found. Some of the genes have a similar expression pattern, so low protein expression levels are also found in the mRNA expression patterns, whereas others do not seem to have a similar distribution. The first group of genes is involved in anthocyanin metabolism, and all these genes were known in Petunia, which made it easier to find the corresponding mRNA sequence in our experiment. The second group of genes is involved in the endomembrane system, and the third group is involved in redox status. For the second and third group, most proteins had an accession number of another species than Petunia, which made it difficult to find the corresponding gene in Petunia. This could have caused comparison of protein levels with mRNA levels of another protein, which is quite useless and leads to a wrong impression.

Graph 1: Comparison of protein (blue) and mRNA (green) expression levels in an an1-mutant.

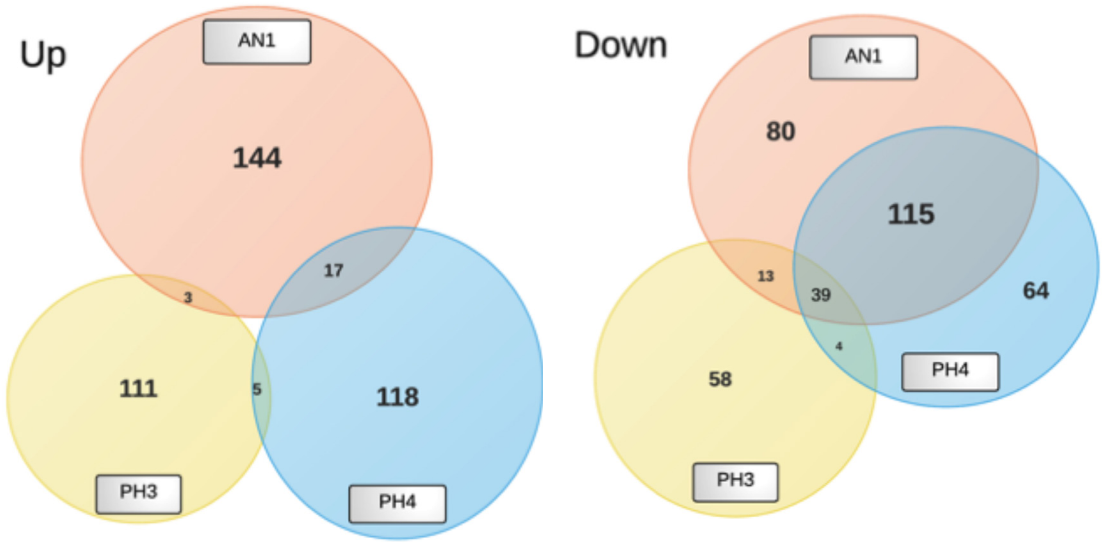


4.2 OVERVIEW

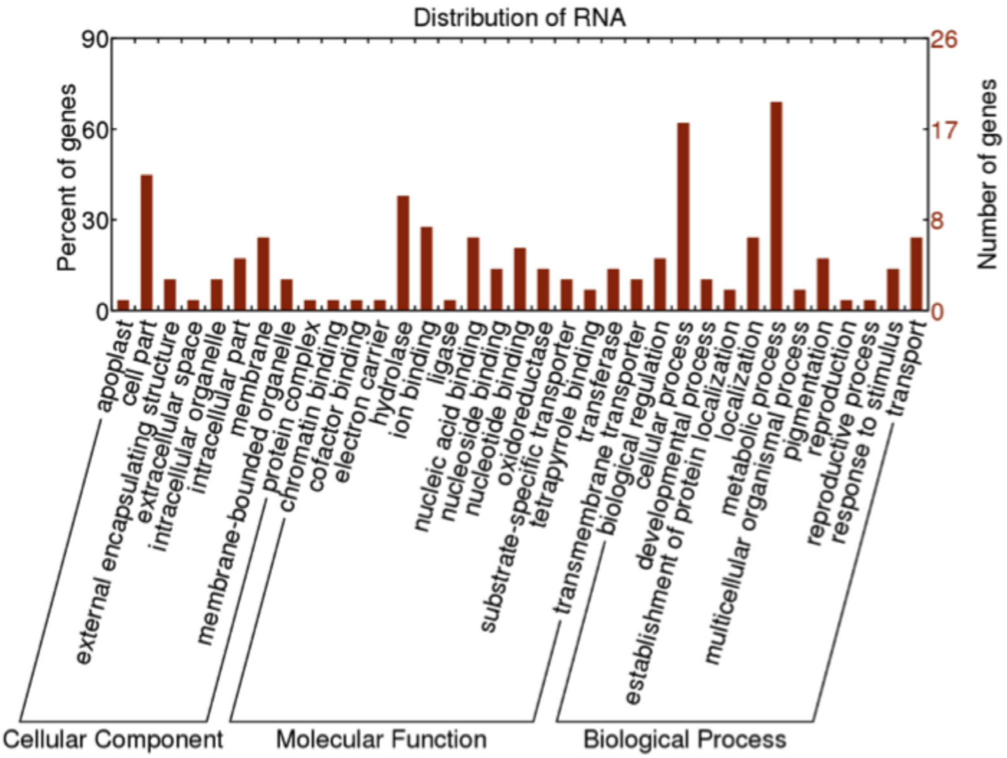
The RNAseq experiment described in 'data & methods' compared the expression levels of 35.237 genes in WT and mutant petals. Of those genes, 411 genes were identified in the an1-mutants that were significantly lower or higher expressed as compared to the wild-type. In ph3, 233 genes were identified, and in ph4, 362 genes were identified. Some genes were higher or lower in only one mutant, some were differentially expressed in only two mutants (but not in the third), and some genes were differentially expressed in all three mutants (see Graph 2A and 2B). If a gene is regulated by AN1 and not by the other two transcription factors, significantly changed expression-levels will only be found in the an1-mutant. However, if a gene is regulated by all three transcription factors, its expression-levels will be changed in all mutants. So if a gene is down in an1, ph3, and ph4, this could indicate that the three transcription factors regulate the gene together. A gene that is e.g. higher in an1, but lower in ph3, will be counted both in the non-overlapping part of the UP-AN1-circle and in the non-overlapping part of the DOWN-PH3-circle. This was analysed because if a gene is down due to one transcription factor and up due to another, those two transcription factors won't be controlling that gene at the same time.

Graph 2 shows two graphs. In the first graph, all genes are shown whose expression is higher in one of the mutants. When expression is higher in the mutant, the gene would have been inhibited in the wild-type. If there is a gene that was higher expressed in an1 and ph3, but not in ph4, it is found in the group where the AN1-circle overlaps the PH3-circle. There is not much overlap between the circles in Graph 2A. There is not even one single gene that is higher expressed in all three mutants. This indicates that all three transcription factors on their own can be involved in the inhibition of certain genes, but they do not seem to work together to inhibit genes. When a gene is influenced by several of the transcription factors it is usually upregulated. This was expected as most genes isolated in the past as targets of AN1, PH3, and PH4 where always up-regulated by these genes.

In Graph 2B, all genes with a lower gene expression are shown. PH4 and AN1 regulate quite a lot of the same genes (115 genes), whereas PH3 and PH4 only show very little overlap. There are just 4 genes that are regulated by PH3 and PH4, and not by AN1. The overlap between PH3 and AN1 is also not very large (13 genes). 39 genes are down in all three mutants. Those genes will be the main focus of the rest of this thesis because they might be involved in the pathway with the vacuolinos and vacuole acidification.



Graph 2. Venn-diagram of up- or downregulation genes
A: The amount of genes that have a higher expression level, divided per transcription factor.
B: The amount of genes that have a lower expression level, divided per transcription factor.
AN1 = an1-mutant; PH3 = ph3-mutant; PH4 = ph4-mutant



Graph 3. An overview of the distribution of the mRNA-sequences. The sequences are divided based on their GO-id's, which relate to the cellular component in which they are active, their molecular function and the biological process in which they are involved (Ye et al., 2006).

The first group that was examined is the group consisting of genes that are regulated by all three transcription factors. All these 39 genes had lower expression levels in the mutant, which indicates that they are normally upregulated by AN1, PH3, and PH4. Blast2GO was used to obtain GO-identities for all the sequences. Graph 3 shows a graph (which was produced with WEGO) in which an overview of all the GO-id's of the genes is given. A gene can have several GO-id's, and three types exist: 'cellular component', 'molecular function', and 'biological process'. It can be seen that a significant proportion of gene-products can be found in a cell part or in the membrane. Additionally, many genes appear to encode for a hydrolase, or are involved in binding of ions, nucleic acids, nucleotides, or nucleosides. Most of the genes are involved in cellular processes and metabolic processes, and a significant proportion of genes is involved in transport or protein localization.

In Table 2, the results of Blast2GO on the 39 genes that are lower expressed in all three transcription factor-mutants can be found. See Appendix A for tables of the results of the other groups.

4.3 IDENTIFIED GENES

A BLAST search in the NCBI database was conducted for all sequences found to be regulated by AN1, PH3, and PH4 to define the protein sequence encoded by the mRNA and to extrapolate the most probable function of the protein on basis of information existing in the database (for example, the database identifies specific conserved domains in the proteins and provides information about them). Subsequently, phylogenetic trees were made for potentially interesting proteins to see how they are related to similar proteins of other species. That aids in determining the specific subgroup to which a protein belongs. For this, a 'one-click' program of LIRMM (with MUSCLE alignment) was used. For more information on genes, UniProt (UniProt Consortium, 2015) was often used and complemented by literature.

Table 2. Blast2GO results – an1, ph3, ph4

#	Gene ID	Probable function	Process
GTPase			
1	Peaxi162Scf01105g00009.1	Ras-related protein Rab18b	GTPase (membrane traffic)
2	Peaxi162Scf00560g01026.1	Ras-related protein Rab5a	GTPase (membrane traffic)
3	Peaxi162Scf00295g11036.1	Ras-related protein Rab6a	GTPase (membrane traffic)
Exocyst			
4	Peaxi162Scf00420g06013.1	EXO70C2	Exocytosis
Transmembrane transport			
5	Peaxi162Scf00008g46011.1	Protein spinster homolog	TM lipid transport
6	Peaxi162Scf00907g01023.1	Probable transporter MCH1-like	TM carboxylate transport
7	Peaxi162Scf00110g10005.1	Sulphite exporter	TM sulphite transport
8	Peaxi162Scf00174g03014.1	Wat1-related/zinc finger protein	TM metabolite transport
9	Peaxi162Scf00089g07002.1	Adenine guanine permease AZG2	TM purine transport
10	Peaxi162Scf00177g06021.1	H+-transporting ATPase (PH5)	TM proton transport
11	Peaxi162Scf00029g28018.1	Copper-transporting ATPase HMA5	TM copper ion transport
RNase			
12	Peaxi162Scf02151g00010.1	Ribonuclease Phy4	RNase activity
13	Peaxi162Scf00746g02005.1	Ribonuclease Phy4-2	RNase activity
14	Peaxi162Scf01075g01010.1	Ribonuclease Phy3	RNase activity
Cytochrome P450			
15	Peaxi162Scf00108g04001.1	Flavonoid 3’5’ hydroxylase	Hydroxylase, ER, anthocyanins
16	Peaxi162Scf00003g18023.1	Cytochrome p450 94a2-like	Oxidation-reduction, iron-binding
DNA binding			
17	Peaxi162Scf00427g06019.1	Zinc finger protein nutcracker-like	Transcription factor
18	Peaxi162Scf00429g05002.1	DNA helicase exonuclease	DNA binding
19	Peaxi162Scf00786g04042.1	MYB-like transcription factor ETC3/TRY	Transcription factor
20	Peaxi162Scf00444g04024.1	ATP-binding / DNA-binding	Binding (ATP, DNA)
Protein binding			
21	Peaxi162Scf00141g11016.1	RING superfamily, E3 ligase	Protein-protein binding
22	Peaxi162Scf00390g03012.1	RING-H2 finger protein ATL72-like	Protein-protein binding
23	Peaxi162Scf00746g02007.1	AMR1/F-box protein	Protein-protein binding

4.3.1 RAB PROTEINS

Two members of the Rab-family were found to be lower expressed in all three mutants. A third one was almost significantly lower expressed in all three mutants; qPCR should be performed to confirm that it is indeed regulated by all three. I did PCR; however, the results were not fully conclusive since the expression levels in the wild-type were also quite low, which caused primer-dimers to form. When there was a peak in the wild-type where the gene-product should have made a peak, there was no more peak in the mutants, so the results seemed to confirm the hypothesis that that Rab-gene (Rab18) is also regulated by all three transcription factors.

Rab proteins are small GTP-binding proteins involved in several processes such as membrane trafficking, signal transduction, cell proliferation, and cytoskeletal organization. They are activated by the binding of GTP, and they are inactivated by the subsequent hydrolysis of GTP to GDP (Vernoud et al., 2003). The small GTP-binding protein family is very large, consisting of five subfamilies, namely Ras, Rab, Rho, Arf and Ran. Rab-proteins are mostly involved in membrane trafficking. Arabidopsis thaliana has 57 Rab-members. Figure 4B shows a phylogenetic tree of this Rab-family (Vernoud et al., 2003) and the regions

#	Gene ID	Probable function	Process
Phosphorylation and dephosphorylation			
24	Peaxi162Scf00046g01019.1	Nucleotide pyrophosphatase phosphodiesterase-like	Metal ion binding, dephosphorylation
25	Peaxi162Scf00575g04008.1	Probable inactive leucine-rich repeat receptor-like protein kinase	Phosphorylation
26	Peaxi162Scf01114g01017.1	Probable serine threonine-protein kinase NAK	Phosphorylation
27	Peaxi162Scf00330g01002.1	Cyclin-dependent protein kinase inhibitor SMR3-like	Kinase inhibitor, cell cycle
28	Peaxi162Scf01132g02011.1	PEBP family protein	Kinase inhibitor
Metabolic processes			
29	Peaxi162Scf00940g04011.1	Low-temperature-induced cysteine proteinase-like	Proteolysis
30	Peaxi162Scf00056g03013.1	Acetyl-CoA carboxylase carboxyltransferase subunit	Fatty acid biosynthesis
31	Peaxi162Scf00001g56015.1	3-hydroxyisobutyryl-hydrolase-like, mitochondrial	Hydrolysis, amino acid degradation
32	Peaxi162Scf00118g18013.1	N-acetyl-glutamate synthase / kinase	Amino acid synthesis / phosphorylation
33	Peaxi162Scf00038g02024.1	Exostosin-containing glycosyltransferase	Glycosyltransferase, exocytosis
34	Peaxi162Scf00198g09011.1	A/b hydrolase family protein	Hydrolysis
Oxidation-reduction processes			
35	Peaxi162Scf00895g03011.1	Probable 2-oxoglutarate-dependent dioxygenase aop1	Oxidation-reduction
36	Peaxi162Scf00038g02003.1	Cinnamoyl-CoA reductase	Oxidation-reduction, steroid dehydrogenase
37	Peaxi162Scf00616g03029.1	Probable trans-2-enoyl, mitochondrial	Oxidation-reduction, zinc ion binding
Rest			
38	Peaxi162Scf00058g21017.1	Protein rolling stone-like	Membrane, myoblast fusion (in fruit fly)
39	Peaxi162Scf00945g01018.1	Intracellular movement?	Acetyltransferase?
40	Peaxi162Scf00550g04031.1	Chitin synthase?	Defence?
41	Peaxi162Scf00328g01019.1	Beta-1,6-glucan production	Cell wall?

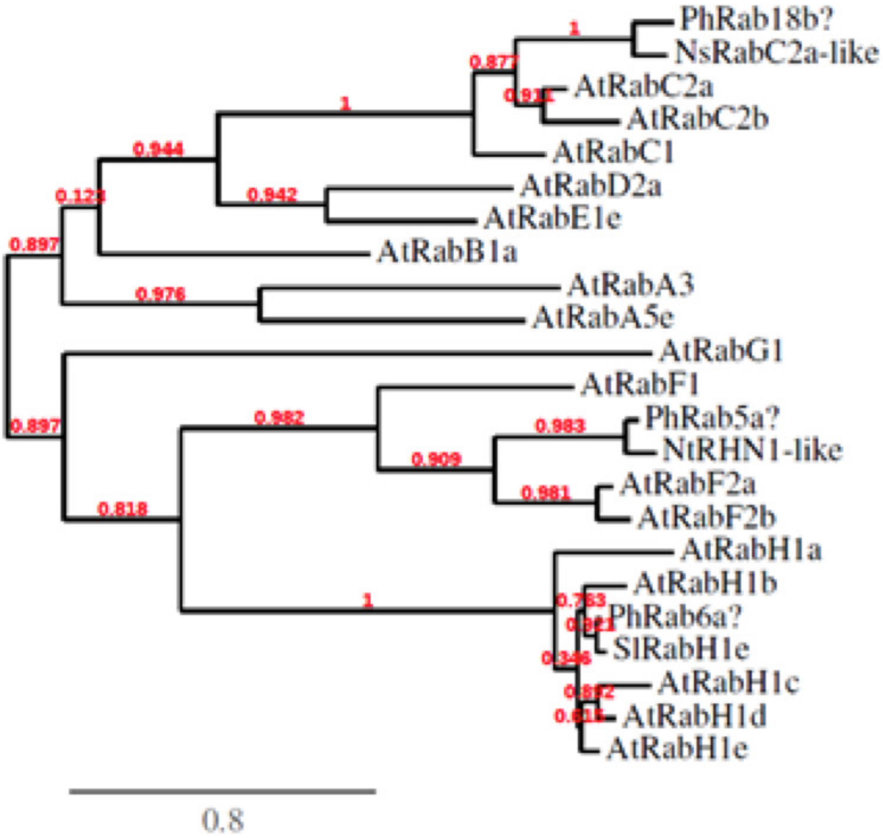
where the proteins are most likely active. The regions were identified by comparing them to research done in humans and yeast. It is found that homologous proteins often still work in the same regions in different species.

We identified three sequences. For sequence 1, the mRNA sequence was called 'Ras-related protein Rab-18-b' in the Petunia database, so this sequence is referred to as PhRab18b?. In the NCBI database, the closest hit obtained using this sequence as a query was a RABC2a-like protein from Nicotiana sylvestris. For the second sequence, 'Ras-related protein Rab-5a' was the name in the Petunia database, so the sequence was referred to as PhRab5a?. The closest BLAST-hit was a RHN1-like isoform X1 protein of Nicotiana tomentosiformis. According to Terryn et al (1992), RHN is a protein found in Nicotiana that is closely related to mammalian Rab5 proteins. Similarly, for sequence 3, 'Ras-related protein Rab-6a' was the name in the Petunia database, so the sequence was referred to as PhRab6a?. The closest BLAST-hit is a ras-related protein RABH1e in Solanum lycopersicum. In the research of Vernoud et al.,X0 (2003), a table can be found that shows that Rab18b is the same as RABC2a, Rab5a is the same as RABF2a, and Rab6a is the same as RABH1b.

Figure 4A shows a phylogenetic tree with some of the A. thaliana members, the three sequences that were identified in my analysis and the sequences that were the closest match in a BLAST-search in the NCBI database. As can be seen in this phylogenetic tree, PhRab18b? is a part of the C-branch of Arabidopsis, closer to C2 than to C1; however, it is not possible to see if the sequence is more closely related to C2a or C2b. PhRab5a? is closely related to F2a and F2b, but again, it is not possible to see if it is closer to F2a or F2b. PhRab6a? is most closely related to AtRABH1b, as is indicated by the name Rab6a. However, most of the close BLAST-hits were H1e-sequences from other species than Arabidopsis.

As can be seen in Figure 4B, PhRab18b?, PhRab5a?, and PhRab6a? could be involved in cell polarization, endosomes, and Golgi-ER retrograde transport, respectively. In mammals, Rab18 was shown to be involved in cell polarization (Lütcke et al., 1994), neuroendocrine secretion (Vasquez-Martinez et al., 2007), lipid droplets (Martin et al., 2005), and ER-Golgi trafficking (Dejgaard et al., 2008), but in plants, not much is known about this protein. In Arabidopsis, Rab18 is found to localize in peroxisomes (Hashimoto et al., 2008). Rab5a localizes to early endosomes and clathrin-coated vesicles in mammals, and it regulates endosome budding and fusion (Vernoud et al., 2003). In Arabidopsis, both Rab5a and 5b are found to be present at the prevacuolar compartment as well (Lee et al., 2004). Rab6 (in mammals) is involved in recycling of ER-resident proteins when they have to be transported in vesicles and brought back to the ER. In Arabidopsis, Rab6 is highly expressed in roots, but also in leaf, stem, and flower tissues. It is able to complement the function of Rab6 in yeast mutants (Bednarek et al., 1994).

Figure 4. Phylogenetic trees of RAB
A: phylogenetic tree of RAB with Arabidopsis (At), humans (Hs) and yeast (Sc) (Vernoud et al., 2003)



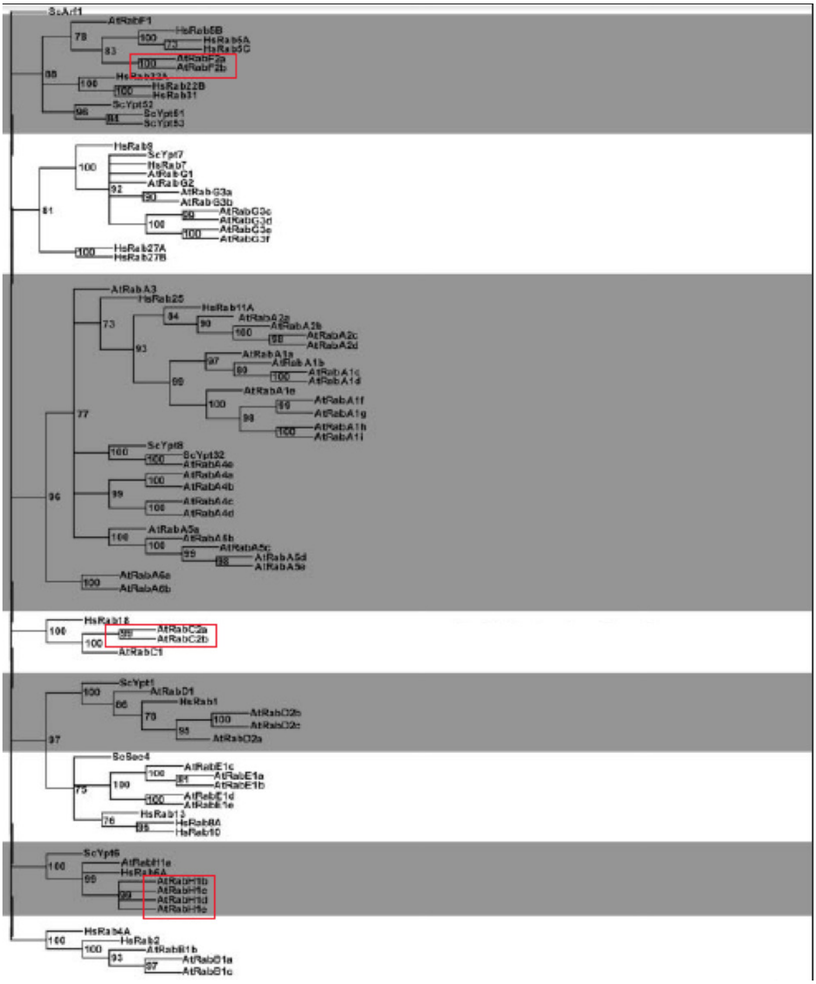


Figure 4. Phylogenetic trees of RAB

B: phylogenetic tree including Petunia sequences (made with 'One-click')

4.3.2 EXOCYST SUBUNIT

One gene the lab was already looking into is called Exo70. It was not identified as one of the 39 genes during the data analysis because its expression was not significantly down in ph3 (p = 0.35), although the fold change in ph3 was higher than 3 (fold change = -5.85). Other researchers in the lab have confirmed experimentally that the gene is influenced by both AN1, PH3, and PH4, so I decided to include it here as well. Exo70 is one of the eight subunits of the exocyst complex, which is involved in tethering vesicles to the membrane and exocytosis. Some of the Exo70-isoforms are involved in morphogenesis or cell polarization, while others are involved in defence against pathogens, stress reactions, or autophagy (Zarsky et al., 2013). The exocyst complex often interacts with Rab-proteins, amongst others Rab11 and Rab8 (Das & Guo, 2011).

In the NCBI database, the best alignment found was with Exo70B1 of Nicotiana tomentosiformis. The B1-protein is found to be involved in autophagy-related transport (Kulich et al., 2013). However, there seems to be a naming inconsistency in the NCBI database. If the Arabidopsis sequences for Exo70B1, Exo70C1, and Exo70C2 (Arabidopsis database; Synek et al., 2006) are blasted against the Nicotiana tabacum genome (TN90, solgenomics database), and if the found Nicotiana sequences are then blasted against the NCBI database, the closest sequences in N. tomentosiformis are all called either B1 or B1-like, even if the query sequence was C1 or C2 from N. tabacum.

The phylogenetic tree containing some of the Arabidopsis Exo70-genes, the Petunia sequence, and some of the Nicotiana tomentosiformis sequences (obtained as mentioned above, by blasting for B1, C1, C2 of N. tabacum), is shown in Figure 5. The overall topology of the tree is very similar to the phylogenetic tree as can be found in Synek et al., (2006). The Petunia gene (PhEXO70?) is most closely related to AtEXO70C2, and not to B1. The C2-subunit is mostly expressed in the stamen, pollen, and anthers, as well as in root hair-producing cells during root hair development (Li et al., 2010). Little is known about the C2-isoform.

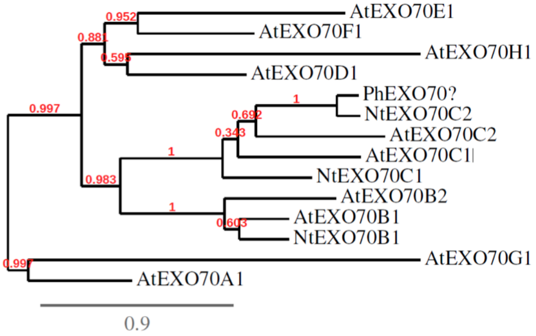


Figure 5. Phylogenetic tree for Exocyst subunits ('one click' mode) At = Arabidopsis thaliana; Ph = Petunia hybrida; Nt = Nicotiana tomentosiformis

4.3.3 TRANSMEMBRANE PROTEINS

Our list contains several proteins involved in transmembrane transport. One is most likely involved in the transport of protons, while others are thought to be involved in transporting sulphite, copper, lipids, metabolites, and purines.

Two proteins containing a major facilitator superfamily (MFS) domain were identified. These domains are associated with membrane transport of a variety of substrates, including sugars, ions, peptides and neurotransmitters, for example. The first protein has two MFS-domains and a Multidrug Resistance Protein domain. The closest BLAST-hit is a protein spinster homolog-1 like found in Solanum tuberosum (90% identical). Spinster homologs in other species are good BLAST-hits as well. Spinster proteins are involved in lipid transport. The Drosophila and human homologues seem to be involved in autophagy and cell death (Yanagisawa et al., 2003). In the sequence, one can recognize a repeated sequence. The first half of the protein is very similar to the second half. The alignment shows that the S. tuberosum sequence is much shorter and that it can be aligned twice against our sequence: once against the MSF-domain in the first half of our sequence, and once against the second domain in the other half.

The other MFS-domain protein has an MFS-domain of the oxalate/formate antiporter subfamily. An antiporter transports two substrates in opposite directions. In bacteria (Oxalabacter formigenes), one substrate (oxalate) is decarboxylated into another (formate) in the cytosol, and the oxalate/formate antiporter transports oxalate into the cell and at the same time removes formate by transporting it out of the cell. This consumes protons and drives an ion gradient, which facilitates proton entry and ATP synthesis (Sakhaee, 2009). Most of the closest BLAST-hits are uncharacterized predicted proteins; the third hit, however, is a MCH1-like transporter protein found in Solanum tuberosum. MCH1 might stand for 'monocarboxylate homologue 1' which is a homologue of the monocarboxylate membrane transporter (Reihl & Stolz, 2005). A monocarboxylate transporter (in humans) transports substrates such as lactate, pyruvate, and ketone bodies across the membrane, usually dependent on protons. However, in yeast, the MCH proteins are found to be involved in other processes, such as mitochondrial respiration and riboflavin transport (Makuc et al., 2001; Reihl & Stolz, 2005). Oxalate is not a monocarboxylate, but formate is.

Another protein has a TauE/SafE domain, which might be involved in the export of sulphite. The six closest BLAST-hits are all 'predicted uncharacterized proteins' with nothing known about them, and they are all around 80% identical. In the top ten of BLAST-hits is a 'sulphite exporter TauE/SafE family protein isoform 2' found in Theobroma cacao.

The fourth protein has domains involved in sulphate, and uracil/xanthine transport. Uracil is a pyrimidine base, whereas xanthine is a purine base. The closest BLAST-hits are adenine/guanine transporters AZG2 in several species. AZG2 belongs to a family of xanthine/uracil/vitamin C permeases. Adenine and guanine are also purine bases, so our protein is most likely involved in purine transport. A phylogenetic tree with AZG1 and AZG2 from Arabidopsis thaliana and Solanum lycopersicum indicates that our protein is closest to AZG2 (see Figure 6A).

The next protein has several ATPase domains involved in transporting cations, possibly magnesium and/or protons, a phosphoserine phosphatase, and a hydrolase. The closest BLAST-hit is a P-type ATPase

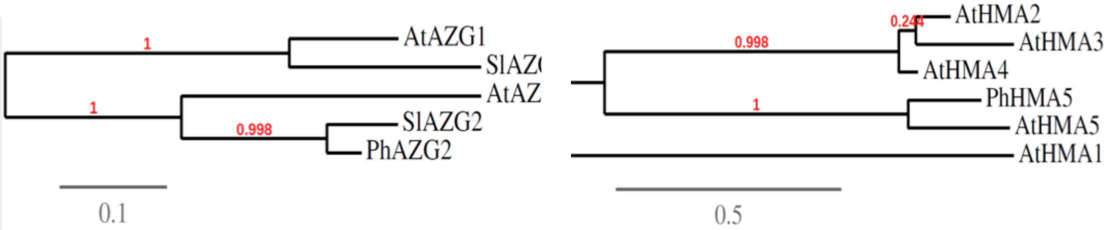


Figure 6A

Figure 6B

Figure 6: Phylogenetic tree AZG & HMA
A. Phylogenetic tree of AZG-proteins ('one-click')
B. Phylogenetic tree of HMA-proteins ('one-click')
At = Arabidopsis thaliana, Sl = Solanum lycopersicum, Ph = Petunia hybrida

---

of Petunia hybrida called PH5. PH5 is a H^+-transporting P-ATPase located on the tonoplast and is involved in regulating vacuolar pH; it thereby determines flower colour. Of this protein, it is already known that it is important for the acidification of the vacuole in Petunia (Verweij et al., 2008).

The last protein has a HMA-domain, which stands for heavy-metal-associated domain. It has a CopZ domain, which stands for a copper chaperone, so the heavy metal it is associated with is most likely copper. Additionally, an ATPase, hydrolase, and phosphatase domain are present. The closest BLAST-hit is a predicted protein called 'probable copper-transporting ATPase HMA5' found in Nicotiana tomentosiformis. Other close hits are homologous proteins from other species. A phylogenetic tree indicates that our protein is indeed closest to HMA5 and not to other HMA-containing proteins (see Figure 6B). HMA5 is a protein involved in copper detoxification. Heavy metal transporters are interesting for flower colouration since the presence of heavy metals in the vacuole can have an influence on the colour of the flower due to the formation of high order complexes of anthocyanins or anthocyanins in combination with co-pigments.

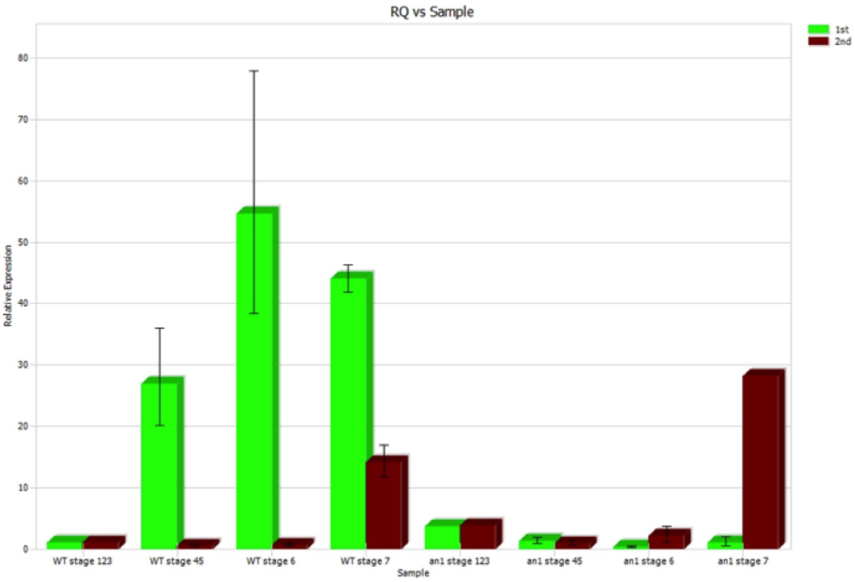
4.3.4 'DOUBLE' PROTEIN

One protein had some transmembrane proteins-associated domains, but also transcription regulation-associated domains. It has two EamA domains (which belong to the drug/ metabolite transporter superfamily (DMT)) in the first half of the protein, which are associated with transmembrane proteins, and two zinc finger domains in the second half of the protein, which are usually involved in DNA binding. However, zinc fingers can bind proteins or lipids as well. Additionally, there are other domains present in the first half of the protein involved in transport, such as a domain of the subfamily of the multidrug and toxic compound extrusion (MATE_tuaB_like) proteins, which are membrane proteins involved in exporting metabolites or products of the biosynthesis of cell wall components. Furthermore, an uncharacterized inner membrane protein domain was found, as well as a Nodulin MtN21 family protein.

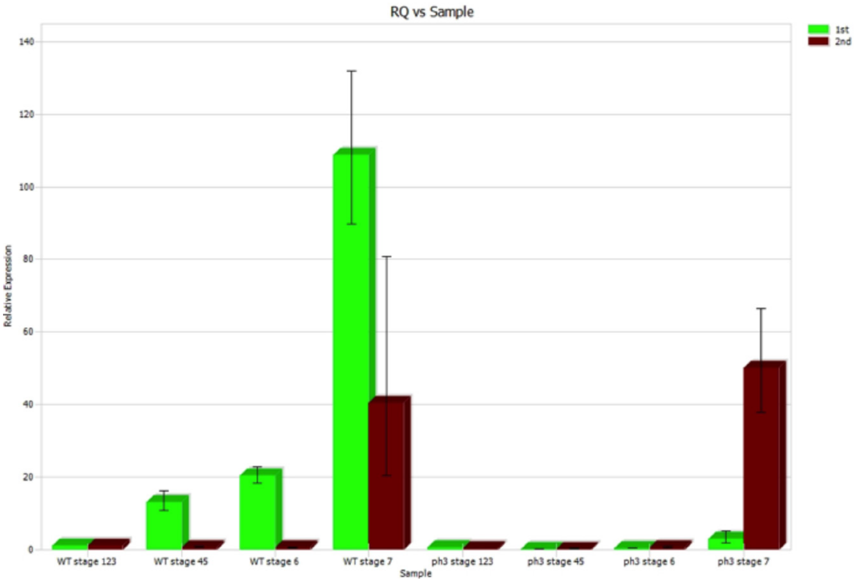
The closest BLAST-hits are PEThy;ZPT3-1 (94%) in Petunia hybrida, several ZAT9(-like) zinc finger proteins from Nicotiana species, and several WAT1-related protein At5g47470-like isoforms found in Nicotiana and Solanum species. The query cover for all of those sequences is approximately 50%, with similarity about 80%. The Petunia protein is a zinc-finger-containing transcription factor. It is similar to the second half of our protein, where the zinc-finger domains are located. The WAT1-related protein is similar to the first half of the protein. WAT1 stands for 'walls are thin' since it is required for cell wall deposition (in Arabidopsis) and a mutant has thin cell walls. It might also be involved in auxin transport to the vacuole and it is located on the tonoplast (Ranocha et al., 2013). Nodulin21 of Medicago truncatula is similar to WAT1 of Arabidopsis.

It looks as though the mRNA is actually coding for two proteins, one with a membrane transport-related function, and one with a transcription regulatory function. In the genomic DNA, exon 1 to 7 are close together, and they are all relatively small exons (around 150 bp). Exon 8 is 2000 bp away and is relatively large (1252 bp). The transmembrane domains are located in exon 1 to 7, whereas the zinc finger domains are located in exon 8. In related species, a similar orientation was sometimes found, however, in some cases the two parts were not close together. To check whether it codes for one or two proteins, qPCR was performed with primers for the first half of the protein and with primers for the second half of the protein. There was a significant difference between expression levels of the two halves (see Graph 4A, 4B, 4C). If it would be one protein, one would expect the expression levels of the mRNA to be the same or at least very similar. Levels for the first half did change in the wild-type compared to the mutants, whereas levels did not change (were very low) for the second half of the protein. Therefore, it appears to be a mistake in the annotation, as there are in fact two proteins encoded by this region and only the first protein (Wat1-related transmembrane protein) is regulated by the transcription factors AN1, PH3, and PH4. We also performed PCR and subsequent gel electrophoresis for this protein, the results of which are added in Appendix D.

Graph 4A



Graph 4B

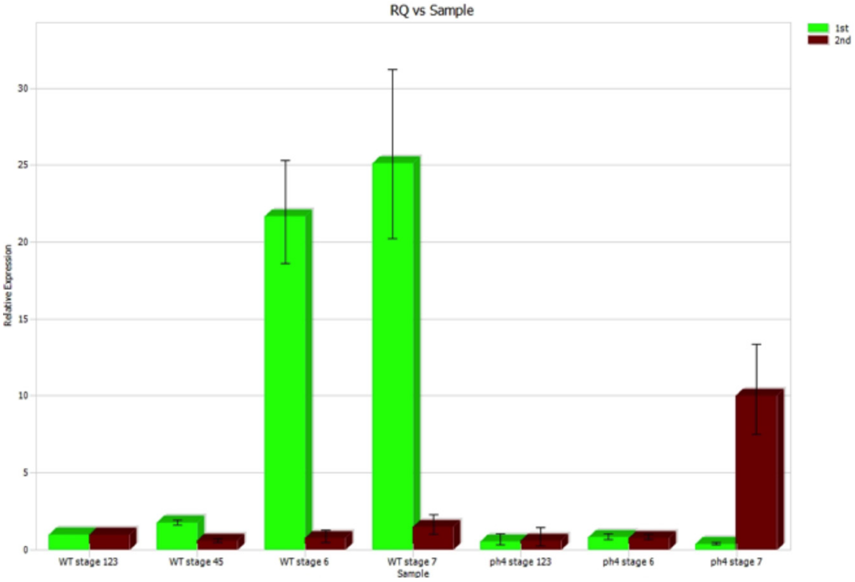


4.3.5 RNASES

Three RNases were found to be down-regulated in all three mutants. Two of the three mRNA sequences appeared to be coding for exactly the same protein, namely RNase Phy4 (which is found in Petunia hybrida), in two different locations on the genome. The mRNA annotation of one gene was wrong, which made it look like the gene was shorter than the other. The third RNase identified is Phy3. RNase Phy3 and Phy4 are members of the T2-family, which are enriched in flower tissues and may have a defensive role there (Hillwig et al., 2010). RNases of the T2-family usually have a pH-optimum between 4 and 5.5 and are either functioning extracellularly or targeted to secretory pathways (e.g. lysosomes, vacuoles) (MacIntosh, 2011; Luhtala & Parker, 2010). Within the T2-family, they are in between the S and S-like RNases. With a phylogenetic analysis they are placed

---

Graph 4C



Graph 4. qPCR results of 'Wat1-related / zinc finger protein'
Green bars are for the expression-levels of the first half of the protein, red bars for the second half of the protein. The first four samples of every graph are expression-levels in the wild-type (WT), divided per stage (stage 123 = first three stages of Petunia, closed flower; 45 = next two stages, after opening flower, 6 = later stage, 7 = senescence started). First stage in wild-type is used as reference, so there expression-levels are set to 1 for both 1st and 2nd half of the protein. The expression levels for the 2nd half of the protein were lower (approximately 30 cycles compared to 25 in absolute values) than the 1st half, for all graphs.
A. Results for samples of an1-mutant.
B. Results for samples of ph3-mutant.
C. Results for samples of ph4-mutant.

within the S-RNases (see phylogenetic tree in Hillwig et al., 2010), and they are basic, as are the S-RNases. Their expression pattern (in ovaries, petals, nectaries, pistil), however, matches S-like RNases. Amino acid patterns normally used to distinguish between the classes did not help either. This makes them unique because no others in that group have currently been discovered (Hillwig et al., 2010).

S-RNases are involved in gametophytic self-incompatibility. S-RNase is expressed in pistils, and a S-locus-F-box (SLF) protein is expressed in pollen. The S-RNase and SLF protein are genetically linked in the S-locus. When the S-RNase present in the pistils matches the SLF protein in the pollen (the pollen is thus recognized as self-pollen), pollen tube growth is inhibited and fertilization is prevented (MacIntosh, 2011). S-like RNases are mainly involved in nutrition through recycling of inorganic phosphate during periods of phosphate starvation, senescence, autophagy, or cell-death, and in defense against pathogens. RNases can also have a housekeeping role by turning over rRNA and regulating gene expression (MacIntosh, 2011; Luhtala & Parker, 2010). According to Hillwig et al., Phy3 and Phy4 might play an antimicrobial role in nectar in Petunia flowers (2010).

4.3.6 CYTOCHROME P450

Two members of the cytochrome P450 superfamily were identified. This family is involved in secondary metabolites biosynthesis, transport, and catabolism. It is also involved in the oxidative degradation of various compounds (e.g. environmental toxins). One of the two proteins is known as a flavonoid-3',5'-hydroxylase (F3'5'H). F3'5'H is involved in flower colour because it hydroxylates flavonoids to form anthocyanidins. A difference in hydroxylation pattern can change the colour of flowers from blue and violet to red and pink (Holton et al., 1993; Koes et al., 2005). The molecule dihydrokaempferol can be hydroxylated by F3'H to form cyanidins (red, pink) and it can stay unhydroxylated to form pelargonidins (red, pink). It can also be hydroxylated by F3'5'H to obtain delphinidins (blue, violet)(see Figure 2). More hydroxyl groups usually leads to a bluer colour (Holton et al., 1993).

In Petunia hybrida, two loci are known to code for F3'5'H, namely Hf1 and Hf2. Hf1 is more widely expressed and has more influence than Hf2 (Chen et al., 2007). Our sequence seems to be more closely related to Hf2.

The other cytochrome P450 protein was most closely related to 'Predicted: cytochrome P450 94A2-like' found in Nicotiana sylvestris and other cytochrome P450 94A1 or 94A2 proteins. Similarity is only around 50%. It has a cytochrome P450 domain, and one domain seems to be of the C-subfamily, which contradicts the closest BLAST-hits (all A-subfamily). It was difficult to make a phylogenetic tree for this protein because there is a large variety of cytochrome proteins present in every species and because the trees I made did not always group the several families together, making it difficult to conclude anything based on the phylogenetic trees.

4.3.7 DNA BINDING PROTEINS

Four probable nucleic acid binding proteins were identified among those encoded by the differentially expressed RNAs. The first protein has a C2H2-zinc finger domain that is usually involved in DNA binding. The sequence motif of C2H2 contains 2 cysteine amino acids and 2 histidine amino acids. The closest BLAST-hit is an unknown protein in Petunia integrifolia subsp inflata (98% identical). The second BLAST-hit is a zinc-finger Nutcracker protein in Solanum lycopersicum (82% identical). In Arabidopsis, this protein is involved in regulation of photoperiodic flowering by modulating sugar transport and metabolism (Uniprot Consortium, 2015).

The second protein contains a WRN_exo domain of the DnaQ_like_exo superfamily. It is a DEDDy 3'-5' exonuclease, which is a DNA helicase that exhibits exonuclease activity. It is involved in DNA replication, repair, and recombination and it maintains genomic stability and life span. The protein is quite short and the closest hits found were not very accurate (around 50%). Additionally, the closest hits are all unknown proteins. This protein was already identified as CAC12.3.

The third protein is very short, and it contains a SANT-domain. SANT stands for Sw13, ADA2, N-CoR, TFIIIB; this domain is involved in (telomeric) DNA-binding. It might be involved in capping, or it can be a repressor transcription factor. The SANT-domain is similar to a MYB5-domain, which is a transcriptional repressor. AN2 and PH4 are both MYB-domain containing proteins. The closest BLAST-hits are all transcription factors, e.g. TRY in Solanum tuberosum and ETC3 in Nicotiana tomentosiformis. TRY (Triptychon) in Arabidopsis is involved in trichome development, the development of hair-developing cells in the root epidermis. ETC3 (Enhancer of Try and CPC 3) is also found to be involved in trichome development. TRY and ETC3 seem to have a similar function in inhibiting trichome development (Wester et al., 2009). It is difficult to produce a phylogenetic tree for the ETC or TRY proteins; when I made an alignment, I was not able to reproduce the alignment made in the paper of Wang et al (2009) (see Figure 7A). In my phylogenetic tree, our protein was closer to the TRY protein of Arabidopsis than to ETC3 (see Figure 7B). It is interesting that the genes regulating the trichome development in Arabidopsis have very similar domains as the genes regulating the anthocyanin pathway in Petunia (namely a WD40-protein (AN11; TTG1), bHLH-domain containing protein (AN1, JAF13; Glabra3, EGL3), and MYB-domain containing protein (AN2; Glabra1, MYB23, WEREWOLF). However, our ETC3/TRY protein is not the same as MybX, which is a protein involved in flower development in Petunia known to be related to the trichome development proteins in Arabidopsis.

The fourth protein contains a Structural Maintenance of Chromosomes (SMC) domain that binds to DNA and is involved in organizing and segregating chromosomes for partition. There is also a domain of the myosin heavy chain and a provisional type I restriction enzyme EcoKI subunit R. The closest BLAST-hits are all unknown or uncharacterized proteins, and the first protein with a known function is a putative ATP-binding protein from Theobroma cacao, with only 49% similarity.

4.3.8 PROTEIN BINDING PROTEINS

Several proteins were identified within my list that were involved with the binding of other proteins, metal ions or ATP. The first protein contained a zinc-finger domain from the RING (Really Interesting New Gene) super family that binds to two zinc atoms. This domain exists in several types; our protein is of the C3HC4-type (so 3 times a cysteine, a histidine, 4 times a cysteine). Such a domain is probably involved in protein-protein interactions and is identified in proteins with a wide range of functions. This RING-finger also contains an E3 ubiquitin ligase, which is involved in posttranslational modification, protein turnover, and chaperone proteins. E3 ligases are ligases that mediate the transfer of ubiquitin to proteins and targets the proteins to the proteasome (Serrano et al., 2006). The first nine close BLAST-hits are 'Predicted uncharacterized proteins' with approximately 80% similarity to our protein.

The second protein is a short protein (155 aa) also containing a RING domain and a ubiquitin ligase domain. Here the RING-finger is from the H2 type. It is also a member of the rad18 DNA repair protein family (involved in nucleotide excision). The closest BLAST-hit is a RING-H2 finger protein ATL72-like protein from Nicotiana sylvestris with 70% similarity and a query cover of 79%. Other close hits were ATL72, ATL73, and ATL74 in several species. ATL72, ATL73, and ATL74 in Arabidopsis are involved in protein modification and ubiquitination, and zinc-ion binding, and they are

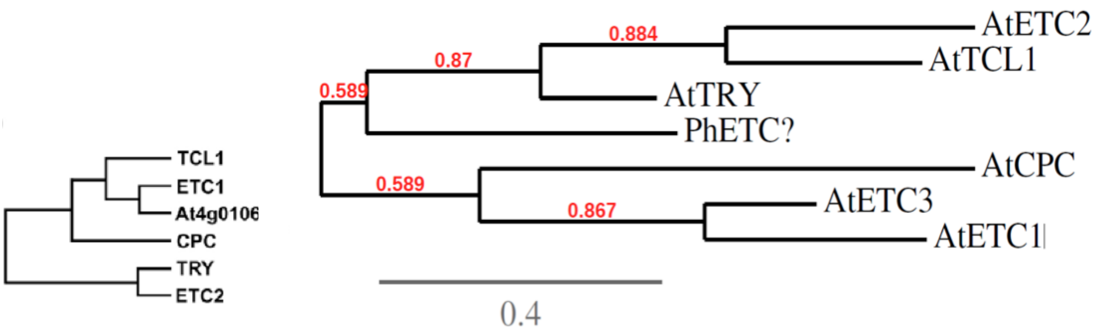


Figure 7A

Figure 7B

Figure 7. Phylogenetic tree of related proteins involved in trichome development
A. Phylogenetic tree from Arabidopsis sequences (Wang et al., 2009). At4901060 = ETC3
B. Phylogenetic tree for family of ETC and TRY proteins ('one-click')
At = Arabidopsis thaliana. PhETC? = Petunia protein

located in the membrane. ATL73 is also involved in stamen and pollen development, anther dehiscence, and jasmonic acid biosynthesis (UniProt Consortium, 2015). A phylogenetic tree was made, which contains several ATL-proteins from Nicotiana sylvestris, Nicotiana tomentosiformis and Arabidopsis thaliana. The Arabidopsis proteins seemed to be more closely related to each other than to proteins in the other species. This might indicate that the Arabidopsis lineage originated from the duplication of a common ancestor. The tree was not conclusive for our protein, but it seems to be closer to ATL72 or ATL73 than to ATL74 (see Figure 8).

The third protein contains two DUF295 domains (domain of unknown function), which are present in plants but have an unknown function. It also contains two AMR1 domains, which stands for ascorbic acid mannose regulator 1. AMR1 has an F-box (Zhang et al., 2009). AMR1 in Arabidopsis regulates the mannose/1-galactose ascorbic acid biosynthetic pathway (Zhang et al., 2009). F-box proteins are involved in signal transduction, regulation of the cell cycle, and protein ubiquitination. The closest BLAST-hit is an F-box protein At4g35733-like isoform X1 from Nicotiana tomentosiformis with only 56% similarity. In the alignment, it can be seen that the other protein is much shorter and it can be aligned twice to our protein: once to the first domain in the first half of our protein and once to the second domain in the second half.

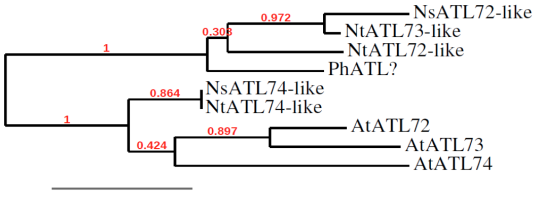


Figure 8. Phylogenetic tree ATL ('one-click' mode)
At = Arabidopsis thaliana, Nt = Nicotiana tomentisoformis,
Ns = N. sylvestris, Ph = Petunia hybrida

4.3.9 PHOSPHORYLATION AND DEPHOSPHORYLATION

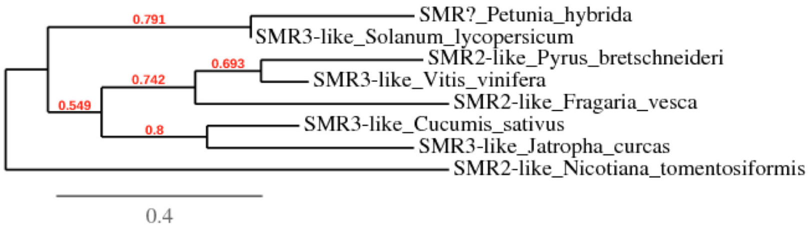
Phosphorylation and dephosphorylation are important processes that add or remove a phosphate group from a protein, respectively. Most often, these processes of protein modification are involved in the activation and inactivation of enzymes. The first protein is a purple acid phosphatase (PAP) of the metallophosphatase superfamily (MPP). It is a diverse, wide-spread family. Since they are insensitive to tartrate inhibition, they are also known as 'tartrate resistant acid phosphatases (TRAP)'. The protein contains a binuclear metal center and catalyzes activated phosphoric acid mono- and di-esters and anhydrides. The metals in our protein are most likely iron and/or zinc since there is an 'Iron/zinc purple acid phosphatase-like protein C'-domain present. The closest BLAST-hit is a nucleotide pyrophosphatase/ phosphodiesterase (NPP)-like protein in Nicotiana tomentosiformis. NPPs hydrolyze phosphodiester or pyrophosphate bonds (e.g. from ATP) to generate PPi. Some NPPs are located at the membrane where they hydrolyze bonds from extracellular nucleotides. They are involved in processes as nucleotide recycling, modulation of purinergic receptor signaling, regulation of extracellular pyro-phosphate levels, stimulation of cell motility, and activity of ecto-kinases (Goding et al., 2003).

The next protein shows a Leucine-Rich Repeat (LRR) domain. A LRR is often involved in protein-protein interactions, and it is present in many functionally unrelated proteins. This protein also contains a Serine/Threonine Kinase domain. This is a domain that phosphorylates serine or threonine. The closest BLAST-hit is a 'Predicted: probable inactive leucine-rich repeat receptor-like protein kinase At3g03770' found in Nicotiana tomentosiformis.

The third protein also contains a Serine/ Threonine kinase domain, with an additional tyrosine kinase domain. The highest BLAST-hits are 'Predicted: probable serine/threonine-protein kinase NAK isoform X1, X2 or X3' from several species. Those hits are all approximately 70% similarity. In Arabidopsis, this protein is involved in plant growth and development and it is located in the cell membrane (Uniprot Consortium, 2015).

The fourth protein does not contain a known domain and is only 156 amino acids. Many of the best BLAST-hits are unknown proteins. Some of the close hits (approximately 50% similarity or less) are cyclin-dependent protein kinase inhibitor SMR2-like, or SMR3-like. According to my phylogenetic tree, this protein is most likely a bit closer to a SMR3-like protein (see Figure 9). This protein is thus most likely involved in inhibiting cyclin-dependent protein kinases. A cyclin-dependent kinase is activated by a

Figure 9. Phylogenetic tree ATL ('one-click' mode)
At = Arabidopsis thaliana, Nt = Nicotiana tomentisoformis, Ns = N. sylvestris, Ph = Petunia hybrida



regulatory protein called a cyclin; they are found to be involved in processes such as regulation of the cell cycle, translation, and mRNA processing.

The fifth protein is quite short (168 aa) and most of the close BLAST-hits are unknown, uncharacterized, or hypothetical proteins. One hit is a UPF0098 protein MTH_273-like, and another is a PEBP (phosphatidylethanolamine-binding protein) family protein. The protein has a PEBP domain that regulates protein phosphorylation as a kinase inhibitor of the YbhB / YbcL family.

4.3.10 METABOLIC PROCESSES

Among the list of target genes of AN1, PH3, and PH4, several proteins were identified that are involved in metabolic processes such as protein degradation, lipid metabolism, and glycosyltransferase activity. The first protein contains a cysteine endopeptidase domain from the C1a subfamily that is similar to papain and mammalian cathepsins. In plants, they mobilize storage proteins in seeds. It also contains a cathepsin propeptide inhibitor domain that is cleaved from the translated protein to activate it. Additionally, a granulin domain is present. Granulins are involved in cell growth. It is closely related to low-temperature-induced cysteine proteinase proteins found in Nicotiana species, for example. This protein has been previously identified to be regulated by AN1, PH3, and PH4. It was then known as CAC16.5.

The second protein contains a domain of an acetyl-CoA carboxylase carboxyltransferase alpha subunit. The closest BLAST-hits are also carboxylase carboxyltransferases, which are involved in lipid metabolism. A carboxyl group is transferred to acetyl-CoA to form malonyl-CoA (see Figure 10A). Malonyl-CoA is an important substrate for the biosynthesis of fatty acids.

The third protein might also be involved in lipid metabolism, but now catabolism instead of anabolism. It is a member of the crotonase/enoyl-coenzyme A hydratase superfamily. Many of the members of this family, including enoyl-CoA hydratases (ECH, sometimes called crotonase) and 3-hydroxybutyryl-CoA dehydratases, are involved in fatty acid metabolism. Our protein has a C-terminal from an ECH, and other domains similar to ECHs as well. Enoyl-CoA hydratase is involved in catabolizing fatty acids by performing the second step in the β-oxidation pathway (see Figure 10B). It hydrates a double bond on trans-2-enoyl-CoA (Agnihotri & Liu, 2003). It is also involved in valine degradation (see Figure 11B). However, the protein also seems to have a domain similar to a 3-hydroxyisobutyryl-CoA hydrolase-like protein (3-HIB-CoA), and the closest BLAST-hits are 'Predicted: 3-hydroxyisobutyryl-CoA hydrolase-like protein 1, mitochondrial' from several species. 3-HIB-CoA hydrolase is a mitochondrial protein, a member of the crotonase superfamily, although it does not have hydratase or isomerase activity, as have other members of that family and it is also not involved in lipid metabolism (Shimomura et al., 2000). 3-HIB-CoA hydrolase hydrolyses 3-hydroxy-2-methylpropanoyl-CoA to form CoA and 2-hydroxy-2-methylpropanoate. It plays a role in degradation of several branch-chained amino acids such as valine (see Figure 11B).

The fourth protein has an N-acetylglutamate kinase (NAGK) domain and a NAG synthase (NAGS) domain involved in amino acid transport and metabolism. The domains are involved in the arginine biosynthesis pathway. They are members of the Amino Acid Kinase (AAK) superfamily and the N-acetyltransferase (NAT) superfamily, respectively. The NAGS domain catalyzes the transfer of an acyl group to a substrate (the amino acid glutamate), which is than phosphorylated by a NAG kinase. The closest BLAST-hit is N-acetyl-glutamate synthase (NAGS) from Solanum lycopersicum; other close BLAST-hits are 'probable amino-acid acetyltransferase NAGS1'-

genes from several species. A NAGS is an enzyme that catalyzes the production of N-acetylglutamate (NAG) from glutamate and acetyl-CoA. NAGS1 (in Arabidopsis) is involved in the arginine biosynthesis pathway (UniProt Consortium, 2015)(see Figure 11A).

The fifth protein has two domains of the Exostosin superfamily. The exostosin gene in humans is a glycosyltransferase involved in the biosynthesis of heparan-sulfate (part of the extracellular matrix) and it might also be a tumour suppressor gene (Uniprot Consortium, 2015). In humans, a mutation in this gene causes Multiple Exostose, a disease characterized by the growth of multiple benign bone tumors. All the closest BLAST-hits seem to code for glycosyltransferases. Glycosyltransferases catalyse the transfer of a glycosyl-group to another molecule. The alignment shows that the other sequence is much shorter and that it can be aligned twice against our sequence, once against the exostosin domain in the first half of our sequence, and once against the second domain in the other half.

The last protein has only hypothetical and uncharacterized proteins as hits, but it has an alpha-beta hydrolase domain. It might be an acyl- or acetyltransferase, or a 2-succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylate (SHCHC) synthase. The latter forms SHCHC by eliminating pyruvate from 2-succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexene-1-carboxylate (SEPHCHC).

4.3.11 OXIDATION-REDUCTION PROCESSES

The first protein I characterized in the collection that is involved in oxidation-reduction has a 2OG-Fe(II) oxygenase superfamily domain. Within this superfamily, it is a part of the 2-oxoglutarate-dependent dioxygenases. Dioxygenases catalyse a reaction where two oxygen-atoms are incorporated into substrates. When one of the substrates is

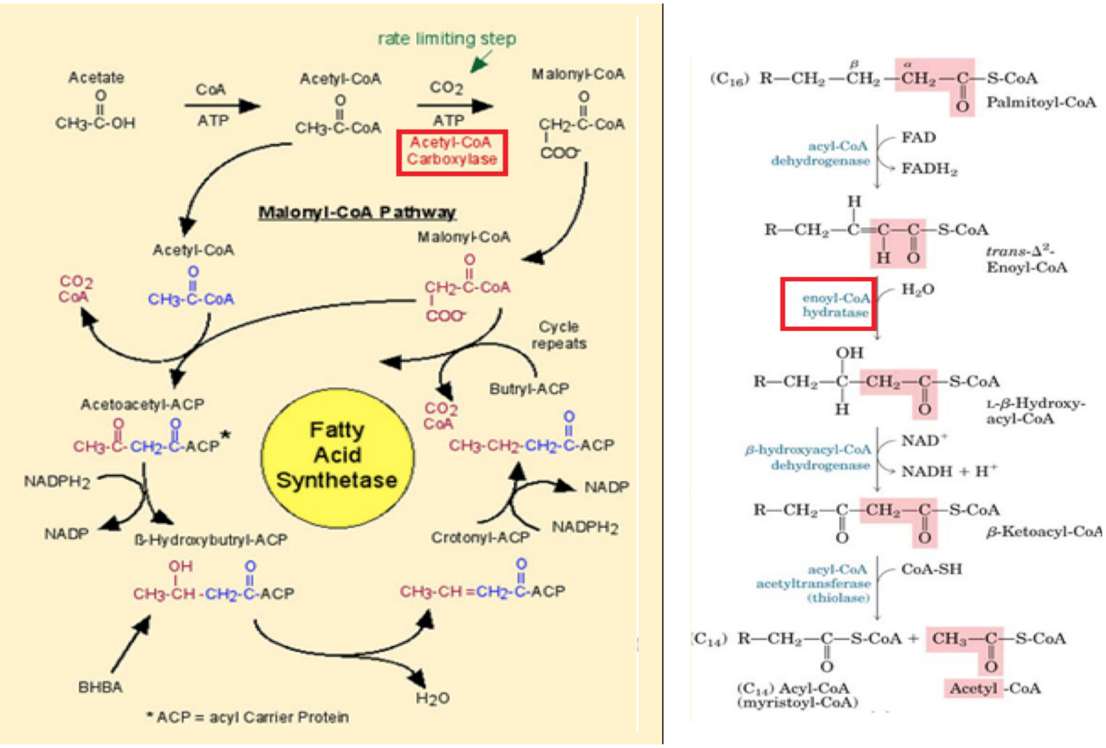


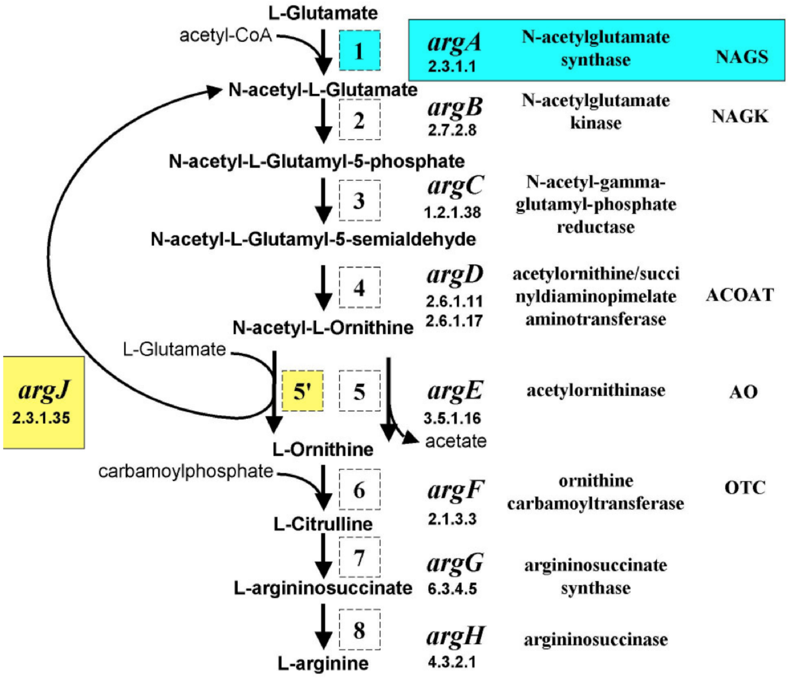
Figure 10A

Figure 10B

Figure 10. Some lipid metabolism pathways the proteins might be involved in
A. Fatty acid biosynthesis; Acetyl-CoA carboxylase is indicated by a red square (Walter, 2009)
B. Beta-oxidation pathway of lipid catabolism; ECH is indicated by a red square (UFhealth)

2-oxoglutarate, the enzyme is called a 2-oxoglutarate-dependent dioxygenase (De Carolis & De Luca, 1994). The protein binds to an iron-atom. The closest BLAST-hit is 'Predicted: probable 2-oxoglutarate-dependent dioxygenase AOP1.2-like' from Solanum tuberosum, with 65% similarity. However, most of the close BLAST-hits are more or less 125 amino acids longer than our protein, and when one blasts a protein like that to the Petunia database you can find the region coding for those amino acids (called 'exon 1' from now on). Exon 1 is almost 12000 basepairs away from the annotated gene in Petunia axillaris, whereas it is close to the other exons in Petunia inflata. The region in between exon 1 and the rest of the gene in Petunia axillaris is a sequence that can often be found in the rest of the genome, so it might be a transposon-like highly repeated region. It is possible that the gene is inactivated because of this transposon, but that is not certain. Expression levels in the wild-type RNAseq data were low. In exon 1, another domain can be found, namely a 'non-haem dioxygenase in morphine synthesis N-terminal' which is a highly conserved N-terminal region of proteins with 2-oxoglutarate/Fe(II)-dependent dioxygenase activity. In Arabidopsis, AOP1.2 might be involved in the production of secondary metabolites called glucosinolates (Kliebenstein et al., 2001). Glucosinolates are produced from amino acids and they are important for the taste of mustard, amongst others. They might also be involved in plant defence.

The second protein is very short. The four first BLAST-hits are hypothetical proteins. Some other hits consist of a bifunctional dihydroflavonol 4-reductase/flavanone 4-reductase (Gossypium arboreum; 67%) and several cinnamoyl-CoA reductase 1-like proteins (several species, around 60%). It has a cinnamoyl-CoA reductase (CCR) domain and a flavonoid reductase (FR) domain. The DFR-protein important



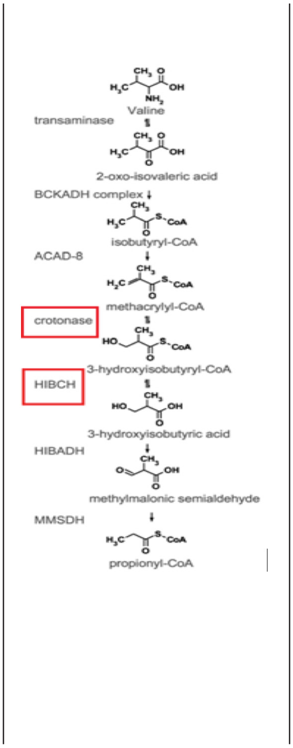


Figure 11A

Figure 11B

Figure 11. Two amino acid metabolic pathways the proteins might be involved in
A. Arginine biosynthesis pathway; NAGS is in the blue square, NAGK right below it (Xu et al., 2006)
B. Valine degradation; crotonase and HIBCH are indicated by a red square (Wanders et al., 2012)

in flavonoid biosynthesis also has a dihydroflavonol reductase domain. There is also a cinnamyl-alcohol dehydrogenase (CAD) family protein domain and a nucleoside-diphosphate-sugar epimerase (NDS-E) domain, which is involved in cell envelope biogenesis and carbohydrate transport and metabolism. All four domains are members of the short-chain dehydrogenase/reductase (SDR) superfamily, which are NAD(P)(H)-dependent proteins involved in many different reactions. CCR and CAD are both important for the biosynthesis of phenylpropanoids, amongst others lignin.

The third protein is either a 2-enoyl thioester reductase (ETR) or an alcohol dehydrogenase. These are both part of the medium-chain dehydrogenase/reductase family, but the alcohol dehydrogenases are dependent on zinc, whereas the ETR are dependent on NADPH. There seems to be a domain involved in zinc-binding, but other domains are indicating this protein is an ETR. The closest BLAST-hits are probable trans-2-enoyl-CoA reductases (so ETR) in several species. The enzyme is involved in fatty acid biosynthesis.

4.3.12 REMAINING DIFFERENTIAL TRANSCRIPTS

Four proteins could not be classified very well since they didn't have clear domains and a BLAST-search did not result in known proteins. The first protein has no known domain and the closest BLAST hits are all very low in homology with the query (around 30% similarity). Several hits are a protein 'rolling stone-like'. In Drosophila, this protein is known to be involved in muscles, namely in myoblast fusion and skeletal muscle development. It is located in the membrane. I blasted the closest BLAST-hit against the Petunia genome; this aligned in the same scaffold as our query, but there were quite some differences that seemed to be mostly due to a frame-shift. The annotated Petunia gene and the 'improved' gene (similarity approximately 60%) can both be found in Appendix B.

The second protein has a putative NT-C2 domain, which was originally found in EEIG1 and EHBP1 (mammals) proteins. These are involved in the uptake of dsRNA to induce RNAi silencing and to regulate endocytotic recycling. It is also found in plants, namely in RPG (Rhizobium-directed polar growth) and PMI1 (plastid movement impaired 1), which are important for intracellular movement of chloroplasts in response to blue light. Most of the BLAST-hits are unknown/ uncharacterized proteins. However, there is one hit (78% query cover, 57% similarity) with a phospholipid-serol O-acetyltransferase-like protein found in Gossypium arboreum. A protein like this transfers an acetyl-group to a phospholipid-serine O-acetyl molecule.

For the third and fourth protein, no domains were identified. Also, results of BLAST-hits were mostly unknown/uncharacterized proteins. For the third protein, there was a hit (91% query cover, 47% similarity) with a chitin synthase 1 protein of Gossypium arboreum. However, it is doubtful that this protein really is a chitin synthase, since chitin is not present in plants. In this protein, I could find a structure similar to AOP1.2 (see above). Here, comparison with high BLAST-hits appeared to indicate that a last exon was not annotated by the Petunia database; this sequence could be found in the database. However, in the P. axillaris-genome it was approximately 13000 basepairs away from the first (annotated) part of the protein. In P. inflata, they were only 2000 bp away from each other. The region in between was repeated in the genome hundreds of times, indicating that this might be a transposon element or another type of repetitive DNA sequence. For the fourth protein, a hit with big1 in Gossypium arboreum was found (74% query cover, 52% similarity). BIG1 (bad in glucose1) in yeast is an endomembrane protein involved in supplying beta-1,6-glucan for the cell wall (Azuma et al., 2002).

5. SUMMARY AND CONCLUSION

In short, at least 40 genes seem to be transcribed under the control of AN1, PH3, and PH4 in Petunia. Some of them have characteristics that make them suitable candidates for involvement in the vacuolino pathway, while others could play a role in defining the petal colour or could open the way to study other phenomena related to the differentiation of petal epidermal cells. Since the vacuolino pathway is part of the cellular membrane trafficking, the most interesting proteins are those potentially involved in membrane trafficking and lipid synthesis together with proteins located in the membranes.

Three Rab proteins were identified. Rab proteins are small GTPases involved in membrane trafficking, signal transduction, cell proliferation and cytoskeletal organization. The three Rabs identified were Rab18b, Rab5a and Rab6a, which are most likely located in peroxisomes, early endosomes and the ER, respectively. An Exocyst subunit, EXO70C2, is possibly involved in tethering vesicles to the membrane and exocytosis. The possible involvement of these proteins in the pathway of the synthesis and physiology of vacuolinos is obvious and therefore they are among the best candidates for further studies.

Seven transmembrane transporters were identified; they are involved in transporting lipids, (mono-)carboxylates, metabolites, sulphite, purines, protons, and copper. A lipid-transporter (Spinster protein) is interesting as a candidate gene of the vacuolino-pathway since lipids are important for membrane synthesis and vesicle production. This protein is also possibly involved in autophagy and cell-death. The proton-transporter PH5 is well-known to be involved in the acidification of the vacuolar lumen and the colouration of the flower; it changes the pH inside the vacuole where anthocyanins accumulate which influences pigment colour. The protein involved in copper-transport might be important since the presence of heavy metals in the vacuole can influence flower colour as well. Another promising transmembrane protein is a WAT1-related protein. It is involved in metabolite transport. In Arabidopsis, it is important for the correct formation of the cell wall, auxin transport, and it is located on the tonoplast, which is promising.

Three RNases of the T2-family are down-regulated in the mutants that might be involved in plant defence or self-incompatibility. It is not certain whether these RNases belong to the S-like RNases or the S-RNases, and so far, no RNases have been identified that belong to the same group. S-RNases are mainly involved in self-incompatibility, whereas S-like RNases are mainly involved in nutrition through recycling of inorganic phosphate during periods of phosphate starvation, senescence, autophagy, or cell-death. RNases can also regulate gene expression and they can be targeted to secretory pathways.

Four proteins were found to be involved in DNA binding. One contains a zinc-finger domain and might be involved in regulating sugar transport and flowering. Another appears to be an exonuclease, which is already found to be involved in the vacuolino pathway. The third is a MYB transcription factor resembling negative regulators ETC3 or TRY of trichome development in Arabidopsis, and the last has a Structural Maintenance of Chromosomes (SMC) domain that is involved in organizing and segregating chromosomes for partition. The ETC3/TRY transcription factor might be interesting to look into because trichome development in Arabidopsis is regulated by the same set of transcription factors as the anthocyanin pathway and vacuolar acidification in Petunia.

I identified several proteins possibly involved in lipid metabolism, namely a carboxyltransferase which carboxylates acetyl-CoA to produce malonyl-CoA, a trans-2-enoyl-CoA reductase which is involved in fatty acid biosynthesis and an Enoyl-CoA hydratase (ECH) which breaks down lipids. However, the ECH might also be closer to another member of the same family (3-HIB-CoA), which is involved in the degradation of valine and other branch-chained amino acids. Another protein (NAGS) was found to be involved in the biosynthesis of the amino acid arginine.

A protein possibly involved in the production of glucosinolates was found; it is a 2-oxoglutarate-dependent dioxygenase, similar to AOP1.2. One protein of the short chain dehydrogenase/reductase family was found, most likely a cinnamoyl-CoA reductase involved in synthesis of lignin for the cell wall. One protein has an exostosin domain. The exostosin gene itself is a glycosyltransferase involved in the production of heparan-sulfate, an important part of the extracellular matrix. Another protein found is an alpha beta hydrolase.

Two RING-domain containing proteins were identified; one has a RING zinc finger, and the other a RING-H2 domain. They both have an E3 ubiquitin ligase domain that is involved in targeting proteins to the proteasome for degradation. The second protein is similar to ATL72 or ATL73, which are located in the membrane. Another protein is an F-box that might also be involved in protein degradation. Another protein involved in protein degradation is a cysteine protease.

Two cytochrome P450 domain containing proteins were identified. One is flavonoid-3’5’-hydroxylase which hydroxylates flavonoids to form anthocyanidins, and is thus involved in flower colouration. The other might be involved in secondary metabolites biosynthesis, transport, and catabolism or oxidative degradation of various compounds, like environmental toxins.

One phosphatase was found, a nucleotide pyrophosphatase/ phosphodiesterase (NPP)-like protein which generates phosphate from (extracellular) nucleotides and which might be located at membranes. It is involved in processes like nucleotide recycling, modulation of purinergic receptor signalling, and regulation of extracellular pyrophosphate levels. It could be interesting that both this phosphatase and the purine transmembrane transporter were identified. Furthermore, the RNases mentioned earlier might be involved in phosphate recycling, similar to the NPP-like protein.

I also found two kinases, namely a Leucine Rich Repeat-domain containing kinase and a serine/threonine/tyrosine NAK kinase. Both are possibly involved in adding a phosphate group to serine and/or threonine. The latter might also phosphorylate tyrosine and is located in the cell membrane. Two kinase inhibitors were also present in the list, namely a cyclin-dependent kinase inhibitor and a PEBP kinase inhibitor.

Several proteins were difficult to identify. One might be a 'rolling stone-like' protein which is located in the membrane. Another might have a domain that induces RNAi silencing and is involved in endocytotic recycling and intracellular movements of chloroplasts. The last two proteins had no known domains and no good hits, so it is difficult to say anything about them. A possible hit for the one is a chitin synthase, and for the other an endomembrane protein involved in cell wall synthesis. The little that can be said about these last four proteins seems to have something to do with cell membranes or cell walls.

During this study, it became clear that the Petunia axillaris genome contains many repeats. This could partly be due to transposon insertions, but also sequencing errors. Several proteins were identified that contained a double domain, whereas the related sequence of other species only contained one domain (e.g. Spinster, AMR1-F-box). Also, some proteins missed an exon that could be found in the genome, but far away from the other exons. In other genomes, the exons were close together and the intron in P. axillaris was repeated in the genome several times. Additionally, when one blasts a sequence against the Petunia genome, the software often finds more than one relatively similar sequence.

What attracted my attention is that there were several proteins that might have something to do with plant defence signalling, including Phy4, Phy3, AOP1.2, the exocyst subunit, and ATL72. This could be an interesting subject to investigate further, as one could ask whether the defence mechanisms in the flower are decreased in an1-, ph4-, or ph3-mutants. However, in flowers, these proteins could also work in a different signalling pathway that could have to do with processes like the triggering of senescence. This is also supported by the finding of proteins like the RNases, three proteins involved in ubiquitination, the Spinster protein, a cysteine protease, which are all possibly involved in degradation and thus possibly senescence. Since these proteins are normally activated or upregulated by AN1, PH3, and PH4, they would induce senescence in the normal flower cells. However, the up-regulation of an exonuclease, which might be involved in stability and increasing the life span of cells, is contradictory in this sense. The proteomics study of an an1-mutant (Prinsi et al., unpublished) that was used to test the RNAseq data also suggested that senescence is induced later in an an1-mutant.

To find out which proteins are actually important for the vacuolino pathway, one should knock out the proteins mentioned in this study (one by one) and see how flower colour and the formation of the vacuolino pathway is affected. That would provide better insight in their function and in their involvement of the processes controlled by the AN1-PH3-PH4 complex. Better insight in those processes will aid in the understanding of plant functioning and development in general, which can be important for agricultural purposes and the flower industry.

REFERENCES

Agnihotri G, Liu H. 2003. Enoyl-CoA hydratase: reaction, mechanism, and inhibition. Bioorganic & Medicinal chemistry. 11:9-20.

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25:3389-3402.

Altschul SF, Wootton JC, Gertz EM, Agarwala R, Morgulis A, Schäffer AA, Yu Y. 2005. Protein database searches using compositionally adjusted substitution matrices. FEBS J. 272:5101-5109.

Azuma M, Levinson JN, Pagé N, Bussey H. 2002. Saccharomyces cerevisiae Big1p, a putative endoplasmic reticulum membrane protein required for normal levels of cell wall beta-1,6-glucan. Yeast. 19(9):783-793.

Bednarek SY, Reynolds T, Schroeder M, Crabowski R, Hengst L, Callwitz D, Raikhel NV. 1994. A small CTP-binding protein from Arabidopsis thaliana functionally complements the yeast YPT6 null mutant. Plant physiology. 104:591-596.

Casañal A, Zander U, Muñoz C, Dupeux F, Luque I, Botella MA, Schwab W, Valpuesta V, Marquez JA. 2013. The strawberry pathogenesis-related 10 (PR-10) Fra a proteins control flavonoid biosynthesis by binding to metabolic intermediates. Journal of biological chemistry. 288:35322-35332.

Chen S, Matsubara K, Kokubun H, Kodama H, Watanabe H, Marchesi E, Ando T. 2007. Reconstructing historical events that occurred in the Petunia Hf1 gene, which governs anthocyanin biosynthesis, and effects of artificial selection by breeding. Breeding Science. 57:203-211.

Conesa A, Götz S, Garcia-Gomez JM, Terol J, Talon M, Robles M. 2005. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics. 21:3674-3676.

Das A, Guo W. 2011. Rabs and the Exocyst in ciliogenesis, tubulogenesis, and beyond. Trends Cell Biology. 21(7):383-386. doi:10.1016/j.tcb.2011.03.006.

De Carolis, De Luca V. 1994. 2-oxoglutarate-dependent dioxygenase and related enzymes: biochemical characterization. Phytochemistry. 36(5):1093-1107.

Dejgaard SY, Murshid A, Erman A, Kizilay Ö, Verbich D, Lodge R, Dejgaard K, Ly-Hartig TBN, Pepperkok R, Simpson JC, Presley JF. 2008. Rab18 and Rab43 have key roles in ER-Golgi trafficking. Journal of cell science. 121:2768-2781. doi:10.1242/jcs.021808

De Vlaming P, Van Eekeres JE, Wiering H. 1982. A gene for flower colouring fading in Petunia hybrida. Theoretical and applied genetics. 61(1):41-46.

Domains NCBI.
1. Marchler-Bauer A et al. 2015. CDD: NCBI's conserved domain database. Nucleic Acids Res. 43(D)222-6.
2. Marchler-Bauer A et al. 2011. CDD: a Conserved Domain Database for the functional annotation of proteins. Nucleic Acids Res. 39(D)225-9.
3. Marchler-Bauer A et al. 2009. CDD: specific functional annotation with the Conserved Domain Database. Nucleic Acids Res. 37(D)205-10.
4. Marchler-Bauer A, Bryant SH. 2004. CD-Search: protein domain annotations on the fly. Nucleic Acids Res. 32(W)327-331.

ExPASY translate tool. Bioinformatics resource portal. Swiss institute of bioinformatics. Available from: http://web.expasy.org/translate/

Unpublished data (thesis): Faraco M. 2011. Vacuolino, a novel endomembrane compartment involved in the trafficking of proteins to the vacuole in epidermal petal cells of petunia flowers.

Gerats T, Strommer J. 2009. Petunia: evolutionary, developmental and physiological genetics. Springer.

Goding JW, Grobben B, Slegers H. 2003. Physiological and pathophysiological functions of the ecto-nucleotide pyrophosphatase/phosphodiesterase family. Molecular basis of disease. 1638(1):1-19. doi:10.1016/S0925-4439(03)00058-9

Hashimoto K, Igarashi H, Mano S, Takenaka C, Shiina T, Yamaguchi M, Demura T, Nishimura M, Shimmen T, Yokota E. 2008. An isoform of Arabidopsis myosin XI interacts with small GTPases in its C-terminal tail region. Journal of experimental botany. 59(13):3523-3531. doi:10.1093/jxb/ern202

Hillwig MS, Liu X, Liu G, Thornburg RW, MacIntosh GC. 2010. Petunia nectar proteins have ribonuclease activity. Journal of experimental botany. 61(11):2951-2965.

Holton TA, Brugliera F, Lester DR, Tanaka Y, Hyland CD, Menting JGT, Lu CY, Farcy E, Stevenson TW, Cornish EC. 1993. Cloning and expression of cytochrome P450 genes controlling flower colour. Nature. 366:276-279.

Kliebenstein DJ, Lambrix VM, Reichelt M, Gershenzon J, Mitchell-Olds T. 2001. Gene duplication in the diversification of secondary metabolism: tandem 2-oxoglutarate-dependent dioxygenases control glucosinolate biosynthesis in Arabidopsis. The Plant cell. 13:681-693.

Koes R, Verweij W, Quattrocchio F. 2005. Flavonoids, a colorful model for the regulation and evolution of biochemical pathways. Trends in plant science. 10(5):236-242.

Kulich I, Pecenkova T, Sekeres J, Smetana O, Fendrych M, Foissner I, Hoftberger M, Zarsky V. 2013. Arabidopsis Exocyst subcomplex containing subunit EXO70B1 is involved in autophagy-related transport to the vacuole. Traffic. 14:1155-1165.

Lee GJ, Sohn EJ, Lee MH, Hwang I. 2004. The Arabidopsis Rab5 homologs Rha1 and Ara7 localize to the prevacuolar compartment. Plant cell physiology. 45(9):1211-1220.

Li S, van Os GMA, Ren S, Yu D, Ketelaar T, Emons AMC, Liu CM. 2010. Expression and functional analyses of EXO70 genes in Arabidopsis implicate their roles in regulating cell type-specific exocytosis. Plant physiology. 154:1819-1830.

Luhtala N, Parker R. 2010. T2 family ribonucleases: ancient enzymes with diverse roles. Trends biochemical science. 35(5):253-259.

Lütcke A, Parton RG, Murphy C, Olkkonen VM, Dupree P, Valencia A, Simons K, Zerial M. 1994. Cloning and subcellular localization of novel rab proteins reveals polarized and cell-type specific expression. Journal of cell science. 107:3437-3448.

Makuc J, Paiva S, Schauen M, Krämer R, André B, Casal M, Leão C, Boles E. 2001. The putative monocarboxylate permeases of the yeast Saccharomyces cerevisiae do not transport monocarboxylic acids across the plasma membrane. Yeast. 18:1131-1143. DOI: 10.1002/yea.763

MacIntosh GC. 2011. Chapter 4. RNase T2 family: enzymatic properties, functional diversity, and evolution of ancient ribonucleases. From: Nicholson AW. Ribonucleases, nucleic acid and molecular biology. Springer-Verlag Berlin Heidelberg.

Martin S, Driessen K, Nixon SJ, Zerial M, Parton RG. 2005. Regulated localization of Rab18 to lipid droplets. Effects of lipolytic stimulation and inhibition of lipid droplet catabolism. Journal of biological chemistry. 280(51):42325-42335. DOI 10.1074/jbc.M506651200

Mol J, Grotewold E, Koes R. 1998. How genes paint flowers and seeds. Trends in plant science. 3(6):212-217. Doi:10.1016/S1360-1385(98)01242-4
NCBI database. National center for biotechnology information. US national library of medicine. National institutes of health. Available from: http://www.ncbi.nlm.nih.gov/

'One click' mode. Méthodes et algorithmes pour la bio-informatique LIRMM. Available from: http://phylogeny.lirmm.fr/phylo_cgi/simple_phylogeny.cgi
1. Dereeper A, Guignon V, Blanc G, Audic S, Buffet S, Chevenet F, Dufayard JF, Guindon S, Lefort V, Lescot M, Claverie JM, Gascuel O. 2008. Phylogeny.fr: robust phylogenetic analysis for the non-specialist. Nucleic Acids Res. 36(Web Server issue):W465-9. Epub 2008 Apr 19.
2. Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32(5):1792-7.

3. Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. 2010. New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. Syst Biol. 59(3):307-21.
4. Anisimova M, Gascuel O. 2006. Approximate likelihood ratio test for branchs: A fast, accurate and powerful alternative. Syst Biol. 55(4):539-52.
5. Chevenet F, Brun C, Banuls AL, Jacq B, Chisten R. 2006. TreeDyn: towards dynamic graphics and annotations for analyses of trees. BMC Bioinformatics. 7:439.

Petunia SGN Cornell. Genome of Petunia axillaris and Petunia inflata. Private site. Available from: http://petunia.sgn.cornell.edu/index.html

Unpublished data: Prinsi B, Negri AS, Quattrocchio FM, Koes RE, Espen L. Proteomics of red and white petals in petunia reveals novel function of the ANTHOCYANIN1 regulator in determining flower longevity.

Quattrocchio F, Verweij W, Kroon A, Spelt C, Mol J, Koes R. 2006. PH4 of Petunia is an R2R3 MYB protein that activates vacuolar acidification through interactions with Basic-Helix-Loop-Helix transcription factors of the anthocyanin pathway. Plant Cell. 18:1274-1291.

Ranocha P, Dima O, Nagy R, Felten J, Corratge-Faillie C, Novak O, Morreel K, Lacombe B, Martinez Y, Pfrunder S Jin X, Renou JP, Thibaud JB, Ljung K, Fischer Urs, Martinoia E, Boerjan W, Goffner D. 2013. Arabidopsis WAT1 is a vacuolar auxin transport facilitator required for auxin homoeostasis. Nature Communications. 4:2625. doi:10.1038/ncomms3625

Reihl P, Stolz J. 2005. The monocarboxylate transporter homologue Mch5p catalyzes riboflavin (vitamin B2) uptake in Saccharomyces cerevisiae. Journal of biological chemistry. 280:39809-39817.

Sakhaee K. 2009. Recent advances in the pathophysiology of nephrolithiasis. Kidney international, Medscape. 75(6):585-595.

Serrano M, Parra S, Alcaraz LD, Guzman P. 2006. The ATL gene family from Arabidopsis thaliana and Oryza sativa comprises a large number of putative ubiquitin ligases of the RING-H2 type. Journal of molecular evolution. 62:434-445. DOI: 10.1007/s00239-005-0038-y

Solgenomics database. Fernandez-Pozo N, Menda N, Edwards JD, Saha S, Tecle IY, Strickler SR, Bombarely A, Fisher-York T, Pujar A, Foerster H, Yan A, Mueller LA. 2014 (online). The Sol Genomics Network (SGN) – from genotype to phenotype to breeding. Nucleic Acids Research. 43(D1):D1036-D1041. Doi: 10.1093/nar/gku1195 Available from: solgenomics.net

Spelt C, Quattrocchio F, Mol J, Koes R. 2002. ANTHOCYANIN1 of Petunia controls pigment synthesis, vacuolar pH, and seed coat development by genetically distinct mechanisms. Plant Cell. 14(9):2121-2135. Doi: http://dx.doi.org/10.1105/tpc.003772

Synek L, Schlager N, Elias M, Quentin M, Hauser MT, Zarsky V. 2006. AtEXO70A1, a member of a family of putative exocyst subunits specifically expanded in land plants, is important for polar growth and plant development. The plant journal. 48:54-72.

Terryn N, Anuntalabhochai S, Van Montagu M, Inze D. 1992. Analysis of a Nicotiana plumbaginifolia cDNA encoding a novel small GTP-binding protein. FEBS. 299(3):287-290.

UFhealth. Shands children's hospital. University of Florida. Department of paediatrics. Division of genetics and metabolism. Available from: https://www.peds.ufl.edu/divisions/ genetics/_style/images/Metabolic%20diagrams/Beta-oxidation-dehydro-cycle.jpg

The UniProt Consortium. 2015. UniProt: a hub for protein information. Nucleic acids res. 43:D204-D212.

Vazquez-Martinez R, Cruz-Garcia D, Duran-Prado M, Peinado JR, Castano JP, Malagon MM. 2007. Rab18 inhibits secretory activity in neuroendocrine cells by interacting with secretory granules. Traffic. 8:867-882. doi: 10.1111/j.1600-0854.2007.00570.x

Vernoud V, Horton AC, Yang Z, Nielsen E. 2003. Analysis of the small GTPase gene superfamily of Arabidopsis. Plant physiology. 131:1191-1208.

Verweij W, Spelt C, Di Sansebastiano GP, Vermeer J, Reale L, Ferranti F, Koes R, Quattrocchio F. 2008. An H+ P-ATPase on the tonoplast determines vacuolar pH and flower colour. Nature Cell Biology. 10(12):1456-1462. Doi: 10.1038/ncb1805

Walter LH. 2009. University of Illinois. Milk Composition & Synthesis Resource Library. Available from: http://ansci.illinois.edu/static/ansc438/Milkcompsynth/milksynth_fatfasynthesis.html

Wanders RJ, Duran M, Loupatty FJ. 2012. Enzymology of the branched-chain amino acid oxidation disorders: the valine pathway. Journal of inherited metabolic disease. 35(1):5-12. doi: 10.1007/s10545-010-9236-x

Wang S, Kwak SH, Zeng Q, Ellis BE, Chen XY, Schiefelbein J, Chen JG. 2009. TRICHOMELESS1 regulates trichome patterning by suppressing GLABRA1 in Arabidopsis. Development. 134:3873-3882. Doi: 10.1242/dev.009597

Wester K, Digiuni S, Geier F, Timmer J, Fleck C, Hülskamp M. 2009. Functional diversity of R3 single-repeat genes in trichome development. Development. 136:1487-1496. doi:10.1242/ dev.021733

Xu Y, Glansdorff N, Labedan B. 2006. Bioinformatic analysis of an unusual gene-enzyme relationship in the arginine biosynthetic pathway among marine gamma proteobacteria: implications concerning the formation of N-acetylated intermediates in prokaryotes. BMC genomics. 7:4. Doi:10.1186/1471-2164-7-4

Yanagisawa H, Miyashita T, Nakano Y, Yamamoto D. 2003. HSpin1, a transmembrane protein interacting with Bcl-2/Bcl-X, induces a caspase-independent autophagic cell death. Cell death and differentiation. 10:798-807.

Ye J, Fang L, Zheng H, Zhang Y, Chen J, Zhang Z, Wang J, Li S, Li R, Bolund L, Wang J. 2006. WEGO: a web tool for plotting GO annotations. Nucleic Acids Res. 34(Web Server issue):W293-297.

Zarsky V, Kulich I, Fendrych M, Pecenkova T. 2013. Exocyst complexes multiple functions in plant cells secretory pathways. Current opinion in plant biology. 16:7260-733.

Zhang W, Lorence A, Gruszewski HA, Chevone BI, Nessler CL. 2009. AMR1, an Arabidopsis gene that coordinately and negatively regulates the mannose/L-galactose ascorbic acid biosynthetic pathway. Plant physiology. 150:942-950. doi/10.1104/pp.109.138453

APPENDIX A. GENE OVERVIEW PER GROUP

Table 3A. Blast2GO results – ph3, ph4

#	Gene ID	Probable function	Process
UP			
1	Peaxi162Scf00111g00024.1	Probable xyloglucan endotransglucosylase hydrolase protein 16	Hydrolase, glucan metabolic activity
2	Peaxi162Scf00610g02001.1	Cytochrome p450 94a1-like	Iron binding, redox
3	Peaxi162Scf00030g13003.1	Protein exordium-like	-
4	Peaxi162Scf00119g15009.1	Mediator of RNA polymerase II transcription subunit	Transcription regulation
5	Peaxi162Scf00332g07017.1	Predicted: uncharacterized protein LOC104090725	
DOWN			
6	Peaxi162Scf00045g01041.1	Probable inactive leucine-rich repeat receptor-like protein kinase	Transferase activity, phosphorylation
7	Peaxi162Scf00315g08019.1	Probable xyloglucan endo-transglucosylase hydrolase protein 23	Hydrolase, glucan metabolic activity
8	Peaxi162Scf00362g03053.1	Receptor-like cytosolic serine threonine-protein kinase rbk1	ATP-binding, phosphorylation
9	Peaxi162Scf00859g02007.1	Minor allergen alt a 7-like	Transcription regulation, plasmodesma

Table 3B. Blast2GO results – an1, ph3

#	Gene ID	Probable function	Process
UP			
1	Peaxi162Scf00167g13035.1	Peroxidase P7-like	Cell wall, peroxidase
2	Peaxi162Scf00610g02002.1	Cytochrome P450 monooxygenase / 94a2-like	Redox, iron-binding, monooxygenase
3	Peaxi162Scf01149g00028.1	Putative expansin-B14	Extracellular, sexual reproduction
DOWN			
4	Peaxi162Scf00128g15034.1	11-beta-hydroxysteroid dehydrogenase 1B	Redox, metabolic process
5	Peaxi162Scf00032g17001.1	Premnaspirodiene oxygenase-like	Redox, monooxygenase
6	Peaxi162Scf00041g04033.1	Probable beta-D-xylosidase 7	Hydrolase, carbohydrates
7	Peaxi162Scf00041g04035.1	Probable beta-D-xylosidase 7	Hydrolase, carbohydrates, cell wall, plasmodesma
8	Peaxi162Scf00140g11028.1	Beta-glucosidase 11-like	Hydrolase, carbohydrate
9	Peaxi162Scf00089g04031.1	Flavonoid 3',5'-methyltransferase	Methyltransferase
10	Peaxi162Scf00316g05019.1	Flavonoid 3',5'-methyltransferase	Methyltransferase
11	Peaxi162Scf01165g02016.1	Probable glycerophosphoryl diester phosphodiesterase 1-like	Lipid metabolism
12	Peaxi162Scf00975g01030.1	PREDICTED: uncharacterized protein	Response to salt stress
13	Peaxi162Scf00377g01008.1	Serine carboxypeptidase	Proteolysis
14	Peaxi162Scf00050g22023.1	Molybdopterin biosynthesis protein CNX1	Metabolic process
15	Peaxi16sScf00420g06003.1	Transcription factor tcp20	Regulation, DNA binding
16	Peaxi162Scf00149g05005.1	Protein exordium-like 7	-

Table 3C. Blast2GO results – an1, ph4 (only up)

#	Gene ID	Probable function	Process
UP			
1	Peaxi162Scf00003g33039.1	Isoleucine N-monooxygenase 1-like	Redox, monooxygenase
2	Peaxi162Scf00491g01012.1	Cytochrome P450 monooxygenase	Redox, monooxygenase
3	Peaxi162Scf00053g05030.1	Ubiquinol oxidase 3, mitochondrial-like	Membrane, respiratory chain
4	Peaxi162Scf01060g01046.1	ABC transporter G family member	Membrane, ATPase
5	Peaxi162Scf00103g07001.1	Mavicyanin-like / blue copper	Copper ion binding, electron carrier
6	Peaxi162Scf00130g03026.1	Heat stress transcription factor B-3	Transcription factor
7	Peaxi162Scf00134g18010.1	MYB-related transcription factor / ODORANT-1	Regulation secondary cell wall biosynthesis
8	Peaxi162Scf00152g02036.1	Phenylacetaldehyde synthase / tyrosine DOPA decarboxylase 1-like	Amino acid metabolism
9	Peaxi162Scf00628g00007.1	branched-chain-amino-acid aminotransferase 2, chloroplastic-like	Transaminase activity, amino acid metabolism
10	Peaxi162Scf00595g01003.1	Alanine-glyoxylate aminotransferase 2 homolog 3, mitochondrial-like	Transaminase activity, response to nitrogen
11	Peaxi162Scf00229g07006.1	3'-hydroxy-N-methyl-(S)-coclaurine 4'-O-methyltransferase-like / OMT3	Methyltransferase, protein dimerization
12	Peaxi162Scf00420g02005.1	Hop-interacting protein THI031	Intracellular
13	Peaxi162Scf00877g01036.1	MATE efflux family protein 5-like	Transmembrane transport
14 P	eaxi162Scf01075g00007.1	MATE efflux family protein 5-like	Transmembrane transport
15	Peaxi162Scf01023g01005.1	zeatin O-glucosyltransferase-like	Transferring hexosyl
16	Peaxi162Scf00640g05003.1	Saur family protein	-
17	Peaxi162Scf00299g09018.1	Predicted: uncharacterized protein LOC104113089	

APPENDIX B. PROTEIN SEQUENCES THAT DIFFER FROM THEIR ANNOTATED SEQUENCE

The next sequences were different in the Petunia database. Based on comparison with similar proteins in the NCBI database, I have adapted some sequences. Some changes had to be made because the given mRNA sequence would not start with a start-codon or would not end with a stop-codon. When there was a (large) gap in either the query sequence or the subject sequence in several comparisons, I looked in the genomic DNA whether it might be possible to adapt the query sequence in such a way it would fit the subject sequences better. I only adapted the sequence when the change would be at the end or start of an exon or when it would add an exon and when there was a good overlap in sequence with several subject sequences in the NCBI database. However, it might still be that the mRNA sequence was not wrong and that my change is not correct.

Green-coloured amino acids are added by me, so they were not present in the mRNA sequence given by the Petunia database, whereas red-coloured amino acids were present in the Petunia database sequence and which might be wrongly annotated.

>1. Rab18b
MMTMMDRRKNHNYGGMSSPSSYDYSFKILLVGDSGVGKSSLLLSFISHHLPQDLSPTIGVDFKIRMLTVG
GKRLKLTIWDTAGQERFGALTSSYYRGAHGIIL VYDVTRRDTFTSLSETWQKDIKFYSTNPDCIKMLVGNKV
DRDSERAVTREEGLAFAKEHNCLFLECSARTRENVQLCFKDLTLKILEEPSLLDKGSTVVKNQILKQKEFQKS
QRNKNCCCSS-

>5. Protein spinster homologue
[TRRIFGISLSLIIINLAAI] MERADENLLPSVYKEVSETFKAGPSDLGYLTFIRNIAKGITSPVAGILVINHDRPTV
LAIGIMFWSLSTAAVGASKYFLQIGIWRAVNGFGLGIVIPTLQSFIADSYMDEVRGTGFGFLHLIGTVGGIG
GGAIATVMAGYEFWGVPGWRFAFVTVATFSILTAFLVFTFVVDPRRKTNKEHDLAKNISDREDLIEKENTN
SLSIWTESWIAMKAVMKVQTFRFIVLQ GLVGSLPWTAMVFFTLWFELIGFDHNSAATLVGLFAAGCALGS
FLGGVIADRLCRRYPHSGRIMCAQFSSFMGIPFSWFLLRVVPQSVSSYYTFAVTLFLMGLTISWNATASNG
PMFAEVVPSKHRTMIYAFDRAFEVSFSSFAAPLVGILAEKLYGYDVKSIDPALGSAKEALALSKGLISMMSSS
QLQLDFVQAKGAVSPSKPKFCERQIFKEQKTRRFFGISLSLIIINMAAIMERADENLLPAVYKEVSETFKAGP
SDLGYLTFIRNFVQGIASPVAGVLVINYDRPTVLAIGILCWSLSTGAVGASKYFLQVGIWRAVNGFGLAIVIP
ALQSFIADSYMDETRGTGFGFLSLVGTVGGIGGGAIATVMAGYEFWGIPGWRFAFVMMATFSTLIAFLVF
TYVVDPRRRASIEHDIAKDTTDRENLIGKGNLNSISIWMESWIAMKAVMKVKTFQFIVLQGLVGSLPWTA
MVFFTLWFELIGAAAALVSLFAAGCAIGSFLGGVIGDRLSHIYPHSGRIMCAQFSSFMGIPFSWFLLRVVPQ
SVSSYYTFAVTLFLMGLTISWNATASNGPMFAEVVPSKHRTMIYAFDRAFEVSFSSFAAPVVGILAERLYGY
DAKSVDPLLGSAKEALALSKGLFSMMVVPFGLCCLFYTPLYWIFNQDRENARIGSIKETEMI-

>6. MCH1-like transporter
MEQSSKWMILIASTWIQAFTGTNFNFSSYSSELKSVLGISQVQLNYLSMASDMGKAFGWCSGVSLMYFP
LWVVLFIAAFMGLVGYGLQWLVIQKIIILPYFLVFLLCVLAGCSICWFNTVCYVLCIKHFPANRPFALSLSISF
NGASAALYNLIANAINSSDDTLYLLLNALVPLVTSVAALPILRQPHSQTHRPDSVHREYLIFLCLTILAVFTGLY
LLLLSSVSYNAPTARILLAGAIFLLVLPMIAPGVIRTEEWSQLFHPSCMPLEDNSDDFEMHKEMLWKEDSS
MTVWTENSHGSKEKESWTSSFLLRDRLLLLGEEHSANLLIRRLDFWLYYLAYFCGGTLVLVYSNNLGQISES
LGYSSEISFLVALYSACSFFGRLLSAAPDFLRDKMYYARTGWLAFALIPTPLAFFLLVLSGSKAALSAATTLVG
LSSGFVFSAAVAITSELFGPNSAGVNHNILITNIPLGSLLYGLLAALVYEANLENSNQVLVLDGSKVCMGRN
CYIQTFMWWGCISLLGVASSLLLFLRTKAAYDSQERNRNCMRLI-

>7. Sulphite exporter
MATKNFVVYLLTNFSVAVLSVFFINNFHNGGFEPRNSSNSNLLQLPIGSQTKVWPELKFSWRPVVAAVIGF
LGSACGTVGGVGGGGIFVPMLTLIVGFDSKSAAAISKCMIMGASASSVWYNLRVPHPCREVPIIDYDLALL
FQPMLMLGITIGVALSVVFPSWLITVLIIILFLGTSSRSFLKGIKMWKEETIFKFKLCYTSCVSVSIDTDYEPLVP
EEEKTTLQLIKFNLDLKRILVLFLVWIMFLLLQVLKNKLVVCSPLYWVLIILQFPVALAVFCYECVKLYKESKKR
RLAGNTISVCEAAIQWSVTNLAFCAFCGILGGTVGGLLGSGGGFILGPLLLEIGVIPQVASATATFVMMFSS
SLSVVEFYLLDRFPIPYAVFLMMVSILAGFWGQFFVRKVVTILRRASIIVFILSGVIFASALTMGVIGIDKSITM
IQNHEFMGFLDFCSSQ-

>8. Wat1-related / zinc finger
One protein:
MGSRKREILEDFAIIGGLIGVQFMYAGNSVVSSYLMWLGFKPSSLIILSSLATFLILSPFSFMFERSQWPRGV
SWKLLIQLFLISFGGVTLFQSLFMKGIKLTSPAMATAMPNLAPGLIFLIAWAFGLEKVEIRCKYSKAKIAGTLL
CVTGAVFMSLMQSSSNAQKDTPSSPPDHYHKDHFLNAEKIHGSLYLMAAIVVLSSNIVLQAATLGDFPAPI
SLCAITSLIGMVLTGIVQLIQQGSLEIGLPLLSIRDLIGYSLLAGIVSGACVSFNNWAMKKRGPVLVSVFSPVG
TVITVVLSAITLKDTITMGSLGGMFLMFTGLYFVLWAKRNEGFLNNTNSSESEVCKRGFNSAGALGGHMR
SHGVGDHNKNYGEEISEQRYKINNFRRDKPEGQKHSYNLRANTNRLLGNRASEDRDKKSSMWPPNDRG
KYAIDETLTLSSMSSPGSSDLERSTKPYNAKEGFYGNDKDKYASREEEEDLANCLVMLSNKSYVLSDNNEAT
YKAKEVEKGMFQCKACKKVFSSHQALGGHRASHKKVKGCYAAKVNADNDDNNDNNNNDTDIDEDSISP
SDLIFHQESNSFQSQSPSSSSSFSRKRSRVHQCSICHRVFSSGQALGGHKRCHWLSSSLPENTFMPTFQEIQ
YHTQEQGLFNKPMFTNFDQPLDLNFPAQLGNPAEFELKLHNPFEQEGPRSYLQLWTDQQINTNLHQNE
KCKDSTEDLRREENYKAKEAKLSNLKDVNLDGGSSWLQVGIGPTPDIVATL-

Two proteins:
Wat1-related:
MGSRKREILEDFAIIGGLIGVQFMYAGNSVVSSYLMWLGFKPSSLIILSSLATFLILSPFSFMFERSQWPRGV
SWKLLIQLFLISFGGVTLFQSLFMKGIKLTSPAMATAMPNLAPGLIFLIAWAFGLEKVEIRCKYSKAKIAGTLL
CVTGAVFMSLMQSSSNAQKDTPSSPPDHYHKDHFLNAEKIHGSLYLMAAIVVLSSNIVLQAATLGDFPAPI
SLCAITSLIGMVLTGIVQLIQQGSLEIGLPLLSIRDLIGYSLLAGIVSGACVSFNNWAMKKRGPVLVSVFSPVG
TVITVVLSAITLKDTITMGSLGGMFLMFTGLYFVLWAKRNEGFLNNTNSSESEYDVEKPLLH-

Zinc finger:
MVDNSQKNEPSTVIHYCRVCKRGFNSAGALGGHMRSHGVGDHNKNYGEEISEQRYKINNFRRDKPEGQ
KHSYNLRANTNRLLGNRASEDRDKKSSMWPPNDRGKYAIDETLTLSSMSSPGSSDLERSTKPYNAKEGFY
GNDKDKYASREEEEDLANCLVMLSNKSYVLSDNNEATYKAKEVEKGMFQCKACKKVFSSHQALGGHRAS
HKKVKGCYAAKVNADNDDNNDNNNNDTDIDEDSISPSDLIFHQESNSFQSQSPSSSSSFSRKRSRVHQCSI
CHRVFSSGQALGGHKRCHWLSSSLPENTFMPTFQEIQYHTQEQGLFNKPMFTNFDQPLDLNFPAQLGN
PAEFELKLHNPFEQEGPRSYLQLWTDQQINTNLHQNEKCKDSTEDLRREENYKAKEAKLSNLKDVNLDGG
SSWLQVGIGPTPDIVATL-

>9. AZG2
MGGELCTTGSFTKLGKSLKKMEMKVNESFSKSRIGKYFKLEARKSCFTKEFRAGTATFLTMAYIISVNASILS
DSGGTCSVSDCSFPVNQTHATSDCLFKPNEGYEKCVAKVRSDLIVATALASMIGSFAMGLLANLPLGLAPG
MGPNAYLAYNLVGFHGSGSMSYQTVMAIFLVEGCAFLAIAVFGLRGRIARFIPQPVRLACAAGIGLFIAFV
GLQAHQGVGLVGPDASTLVTLTACSSTNPVTGECTGGKMQSPTFWLSMVGFIIMCYGLMKNIKGSMIY
GILFVTLISWIRNTAVTVFPNTPSGNSSYEYFKKVVDFHKIESTAGAISFSHFNNGEVWVALLTLLYIDVLATT
GTLYTMAEIGGFVNEEGEFEGEYIAYMVDAGSTIVASTLGVSPIATFVESSAGIREGGRTGLTAIIVGFYFLLS
LFFAPLIASVPPWAIGPSLVMVGVLMMKVVKDIDWNNIKEAVPAFVTMVLMPLTYSISNGIIGGIGMYIAL
SLYDSMVTGVNWLMKMRKMVVKEQNQVSAANADQRIEII-

>10. PH5
MAEDLERPLLGPENFSREGIDLEKLPLEQVFEELRTSKEGLSDEDAEERLNIFGPNKLEEKRENKFIKFLGFM
WNPLSWVMEAAAIMAIALANGGGQGPDWQDFLGIVCLLLINSTISFIEENNAGNAAAALMARLAPRTKV
LRDGRWQEKDAAILVPGDIISIKLGDIIPADARLLEGDPLKVDQSALTGESLPVTKKTGDEVFSGSTCKHGEI
EAIVIATGVNSFFGKAAHLVDSTQVTGHFQK[VRKENSVILLLDSTIW]VLASIGNFCICSIAMGMILEIIIMFP
VQNRSYRTGINNLLVLLIGGIPIAMPTVLSVTLAIGSHRLSQQGAITKRMTAIEEMAGMDVLCSDKTGTLTL
NRLTIDRNLIEVFQKDMDKDMVVLLAARASRLENQDAIDAAVINMLADPKEARANIREVHFLPFNPVDKR
TAITYIDSDGKWYRASKGAPEQILTLCQEKQQIAAKVHTIIDKFAERGLRSLAVSFQEIPENSKESPGGPWQ
FCGLLPLFDPPRHDSAETIRRALNLGVCVKMITGDQLAIAKETGRRLGMGTNMYPSCSLFGRDKDETEALP
VDELIEKADGFAGVFPEHKYEIVKILQMNEHVVGMTGDGVNDAPALKKADIGIAVDDATDAARSAADLVL
TEPGLSVIVSAVLTSRAIFQRMKNYTIYAVSITIRIVLGFMLLALIWKYDFPPFMVLIIAILNDGTIMTISKDRV
KPSPRPDSWKLNEIFATGVVLGTYLALVTVLFYWLADSTQFFEAHFHVKSLSGSSEEMSSAVYLQVSIISQA
LIFVTRSQSWSFAERPGALLMFAFVVAQLVATLIAVYAHINFASVRGIGWGWAGVIWLYSLIFYIPLDIIKFA
VCYALTGEAWNLLFDKKTAFTSKKDYGREDREAQWVLSQRSLQGVISSEFEPRSRRPSMIAEQAKRRAEIT
RLRELYTLRGHIESVARLKNLDLNKIQTAHTVLIFESA-

>11. HMA5
MEANGKDDLKKPLLQDPDAVAVNVAPLNDSRNKKIRTLLFKVNGITCASCSNSIESALGKLKGIESATVSPL
QGQAIVKYVPELISAIKIKEAVEDTGFLVDEFPEQDIAICRIRIKGMACTSCSESVERALSMLDGVKKAVVGLS
LEEAKVHFDPTVTSTSRIVEAIEDAGFGADIISSGSDLNKVHFKLDGISSPDDFTAIQCCLEALEGVNTVEINQ
QEHIVIISYEPDIIGPRTLMQCIQEAGHGSSTYCASLYIPPRQREIEKEHEIRSYRNLFLWSCLFSVPVFVFSMV
LPMLPPYGNWLEYKVFNMLTVGLLLKWILCTPVQFVIGRRFYAGSYHALRRKSANMDVLIALGTNAAYFY
SVYIMIKALTSNSFEGQDFFETSPMLISFILLGKYLEVLAKGKTSDALAKLTELAPETAYLLTLDGAGNIISETEI
STQLIQKNDVLKIVPGAKVPVDGVVINGQSYVNESMITGEARPVSKKTGDKVIGGTVNENGCVLIKATHIG
SETALSQIVQLVEAAQLARAPVQKLADQISRFFVPTVVLVAVVTWLGWFIPGELGVYPSSWIPKGMSIFEL
ALQFGISVLVVACPCALGLATPTAIMVATGKGASQGVLIKGGNALEKAHKVKLVVFDKTGTLTVGKPTVVS
AVLFSNFSMEEFCDITISAEANSEHPIAKAVVDHAKKLRQTRGAENEYHTEIEEFEVHAGAGVSGKVGERRI
LVGNRRLMHAFNVPVTNEVENYISENEQLARTCVLVAVDGKIAGAFAVTDPVKPDAARVVSFLHSMDITS
VMVTGDNWATAKAIASEVGIQMVFAETDPLGKADKIKELQ[ASFSLLFPDSVHFLSIKLI]LKGTPVAMVGD
GINDSPALVAADVGMAIGAGTDVAIEAADIVLIKSNLEDVVTALDLSRKTMSRIRLNYVWALGYNVLGMP
VAAGVLFPFTGIRLPPWLAGACMAASSISVVCSSLLLQSYKKPLHTRVN-

>13. Phy4.2
MEFLIMIMMFNMLVPQIFSQVNDYHIMQFVLGWNPSFCSKPDTKPCVNPVPENFTIHGLWPANAYANS
LYLKMPKSWKERQDLKNDMALLNQDQNLVNSLHHVWPSVVPKFPAELFWRHEWAKHGFGIRAQIDVK
TYFEAATRIHDTMIAINGKTNLKGYFTGVGIQPGKSVTVSQLGRALNPLVNGIDIRCYNNGTHNFLLEVIFCL
DKSSLYSFISCQRAYSRTGQVQLTGSCVLKDPLYLPSQ- [IKVEWIQSISAKWN]

>14. Phy3
[NALVLIQVSIIVLSSFYELSHGEVGIENVSDEPVLESLMFKIGHKRPPPSRPPKQQKPSPPRPVKSPPPSPPLP
PPPVKSPPSWTPPPSAHAQPPPPH] MVPTVVSDYHLMQYVLQWVPTTCMKEPNCKASFLNPQENVFSL
HGLWPANSTGYSLNCSEPKLTTLLNVRNVWRGDQTLKLKLQKVWSNLLGPDEGFWIHEWKKHGFCTNTI
IKDVDYFKAAITINNMIVKGSTNNLIGYLKAVGIYPSNSFHSKTDIESALYPLVGKNNKVYVSCEKIDNQVHL
KEIYLCLDKSLQKFISCPQPHANRGCGPSPKNIVIPAF-

>15. F3’5’H
MVLLSELAAATLIFLTTHIFISTLLSITNGRRLPPGPRGWPLIGALPLLGAMPHVSLAKMAKKYGAIMYLKVG
TCGMAVASTPDAAKAFLKTLDINFSNRPPNAGATHLAYGAQDMVFAHYGPRWKLLRKLSNLHMLGGKA
LENWANVRANELGHMLKSMFDMSREGERVVVAEMLTFAMANMIGQVILSKRVFVNKGVEVNEFKDM
VVELMTTAGYFNIGDFIPCLAWMDLQGIEKGMKRLHKKFDALLTKMFDEHKATSCERKGKQDFLDCVME
NRDNSEGERLSTTNIKALLL[VCSKY] NLFTAGTDTSSSAIEWALAEMMKNPAILKKAQAEMDQVIGNNRR
LLESDIPNLPYLRAICKETFRKHPSTPLNLPRISNEPCIVDGYYIPKNTRLSVNIWAIGRDPEVWENPLEFYPER
FLSGRNSKIDPRGNDFELIPFGAGRRICAGTRMGIVMVEYILGTLVHSFDWKLPSEVIELNMEEAFGLALQK
AVPLEAMVTPRLPIDVYAPLG-

>17. Nutcracker
MPADNSSAMNDNSTGSGEASVSSSGNQAKETSKKKRNLPGMPDPDAEVIALSPTTLLATNRFVCEICSKG
FQRDQNLQLHRRGHNLPWKLRQRSSKEVKKRVYVCPEPSCVHHDPSRALGDLTGIKKHFCRKHGEKKWK
CDKCSKKYAVQSDLKAHSKICGTREYKCDCGTLFSRRDSFITHRAFCDALAQESAKALPEEPPNANEEPNN
QAIASPSPPPTPPVLPAATESQPPPSLPPPPAASMPITPAAVPPSTAVISFVSPVQNPVSAELRESSNSDSAG
TGTNHVMEETAALVSLTGSCSSSSNSCSNGSTSSSVFGSVFASSTASGSMPSQAPGFIGLFQAMAPDRVA
EMAPPSSTEPISLCLAMSQGSSIFRPAGQERRQYAPAPQPAMSATALLQKAAQMGAAATSSSLLRGFGV
MSSTSSSHGQQEWNGRPLDPDGASLAAGLGLGLPCDAGSGLKELMLGTPSVFGPKHPTLDLLGLGMAA
NGGPSPGLSALITSMGSNLDMASFGSADFSGKDLGRNS-

>19. ETC3 / TRY
MDQNLHHQPKIMHRCCSHE[VVSFVLIWKT]EVNSMEWEFISMTKQEEDLIYRMHKLVGDRWGLIAGRI
PGRTAEEIERFWIMRHSDGFAHKRRQLRKV-

>21. RING protein
MDESNQFSGDASTSEQVQVIDDDDRRPPVPFPERNSNDVFHNAPYYPVSNFRANDFPGNDNYEGAVPL
TRDDTWSCFVVVLTFWFFVSMTLILGVYGPADLQLGPNSSILIKRNPLFVKNMKVEEMDDANNRPTVYGF
YQNPTLDVFTTSHEAYKTSLPPITNKEWIYNLNQGSQINITYTVNSSSSNSLVLIIAQGSEGYAQWLEDPSYP
NTTLSWNVIHGSGTISQDIGKSSTYYVAVGNLNSGVVQVRLDIRVKALLYNTTGSYYECNLRQKECSMRFF
FPHGNAALLTSPPQRPDTANGMWSIKISYGPRWITYLFGVGGMSFLIWLVFWYLNKMQSTRQEDIRDPV
GPTESQQTPLLSHKDDDLASWGSSYDGLSQDDEYNEDGLDLPASQGKQVNDGEFSNNIRRLCAICFDAP
RDCFFLPCGHCASCFECGTRIAEAAGTCPICRRPMK KVRRIYTV-

>22. ATL72
MCQHGLPRSILSNSTQTNNDNSISSSCPLILLVALLLLLLCAIAVKAIIRCVLRYNGRFPFESEDSATARLASNG
LNKDELRKIQVIVYEPGMKMVSVINECPICLGEFQHGEKLRMLP[E]VCNHGFHVKCIDKWLSSHSSCPTCR
QLLLLNPVCQA-

>25. LRR kinase
MDKQLRFKKILALVFLMLSISFSEQLQSSQVHTLLRIQNLLNYPTALNNWNNDMDFCTTEPSSNVTIVCYE
ESITQLHIIGEKGASQLPTNFSLDSFVTTLVKLPSLKVLRLVSLGLWGPLPRKLSRLSSLEILDLSSNFIHSDIPQE
ISSLTNLQTLILDGNRFTGRLPKGIGSLLVLAVMSVKNNSLHGDLPDTLRNLHNIRVLSLSRNNFTGDVPDLS
NLGNLQFLDLEHNSLGPKFPQISSKIESIVLRNNKFTAGIPEKIQSYNQLEHMDISSNRFMGPFPLYLLSLPSI
TYLNVAGNKFTGMFFEDNRCNAGLDFVDLSDNLLTGRLPSCLFTGFKHRIVHFSNNCLATGDRSQHPYSF
CRNEALAVGILPRHQKQKQASKVVLALIICGSTTGGVVLVCASILIVKNLLAKKAARKTPTRIILENASSAYTS
KLFTDARYITQSMKLGTLSLPSYRTFSLEELEVATNNFDTASLMGDQMYRGKLKDGSYIAIRCLKMERSNSS
PNLMHHIELISKLRHHHLVSSLGHCFDCSLDDSSVSRIFLVFEYVPNGTLRSWISDKHAKRRLTWTRRIAAAI
GVARGIQFLHTGIVPGIFSNNLKITDILLDQNHVAKLCSYNLPVVAEKESSQNIPSGSKELKSPRANNEEKLD
VFDFGVILLEIISGRQINTKNEALVIQNQLEESIRANKMSRMNVVDPAIRNSCSNESLKTIIEICYRCLLQDTEE
RPSVEDVLWNLQFASQVQDAWTRDSSSSDGSPISPPLVMQ-

```
>26. NAK kinase
MGNCCPKPVNNHPSSVKLSSCPAVSSTAPTFKKPTINNNNNSTVQGEIRPVGSSDGGVKGQLVPSNGKII
TPNLKMFSLGELKSATRNFRPDTVLGEGGFGRVFKGWVDEKTFAPSRVGVGMAVAVKKSNPDSDQGHK
EWQ[FSKEIDYKRICGVLQ]AEVDFLGKFSHSNLVKLIGYCWEEKEFLLVYEHMQKGSLERHLFRKGAETLS
WNTRLKIAIGAARGLTFLHTTEKQVIYRDFKAANILLDEDYNAKLSDFGLAKLGPINGNSHVSTRVVGTYGY
AAPEYVASGHLYVNSDVYGFGVVLLEIITGLRVLDMNRLTGQHDLVDWAKPILPDKRKLRTIMDSRLQGQ
YPIHAAFQVAQLILQCLESDPKTRPPMEGVLECLEQCNSVKMNPRESKGNHSKEHRYHQSGDHWNHNY
HPQRSPLHVDKTGSGDVIGRKAHHSPINRSY-

>31. 3-HIB-hydrolase
MKHFATCKAWWQVLVEEKGSCRTVTLNRPSILNALNYFMVSRLFQLYKSWEDDPAVRFVVLKGNGRAF
CSGGDVVQTYNLIKK[GKTDTSSLP]GNIDECKEFYRKVYSIAYVTGIYSKPQVALLNGITMGGGAAISIPGTF
RVATEKTVFANPETLIGLHPDAGASFYLSRLPGYLGEYLGLTGDKLSGEEMMSCGLATHYSLSERLPLIEEQL
GKLITDDPSVIESSLAKFGDIVHPDQMSVLKRIETVNKIFSHETVEEIVDCLESEVAKTADAWCISTLKRLREIS
PLSLKVSLRSIREGRFQTLEQCIIREYRMTLQAFFGQITHDYCEGVRARIVDKDLSPKWNPPSLEHVTNDMV
DEYFSPLTAVEPDLELPTQQKKAVA-

>32. NAGS
[MAMLKLNHGWNLSMGKKYDLFIKGGDKFLRFEGKRTNYVKDRLCLGCCSGTDHNGKVIIMNENYHSLS
TAAAAVEEGEAEEFVNMMREIQPYFVAHRGSTFVLVLSAEIIDSPYLSSILK]MLITSLKRPSFIQIKQFHPPTF
NISKFPIKNDFKEKQKMLSFGVKCNAIGETENGIDVVSNAIKDDIFVGFFREAWPYFLAHRGSIFVVLISAEIV
DSTTHLDPLLMDISLLHGLGIKFVLVPGTHVQIDRFLAERGSEPKYVGRYRVTDPDSLMAAMDAAGRTRL
MIEAKLSPGPSLTGVRRHGENSRWHDGVSVTSGNFLAAKRRGVAEGIDYAATGEVKKIDVSRIRERLDQD
SIVLLSNLGYSSSGEVLNCNTYEVATACALALGAEKLICIIDGPILDESGRLIRFLTLQDADMLVRKRAAQSEIA
ANYVKAVSQEDFNCLDHNGSNGSISSYNMNGFSEKYSVTFQNGVGFDNGNGLWSSEQGFAIGGQERLS
RLNGYLSELAAAAFVCRGGVQRVHLLDGTIGGVLLKELFQRDGVGTMVASDLYEGTRMARLSDIPGIKQLL
QPLEESGTLIRRSEEELVEALHSFIVVEREGQIIASAALFPYFEDKSGEVAAIAVSPDCRGQGQGDKLLDYIE
KASNLGLQMLFLLTTRTADWFVRRGFSECSIDHIPARRRKKINLSRRSKYYMKKLLPDRSGIRFDSPLS-
[ESVHQKSFRS-]

>34. A/b-hydrolase
MIGKTVILFIVFLALAYQAIQPPPPKICGSPDGPPITAPRVRLSDGRHLAYKEQGVPKHQAKYKIVFIHGFDC
CRHDVAFASSLSPDVIETLGIYIVSFDRPGYGESDPHPQRTPKSLAHDIEELADQLKLGTKFYVTGFSMGGQ
AVWGCLKYIPYRLAGAALLTPVTNYWWPGFSANLTKKSYYDIPPPDLWTLRVSHYLPWLTYWWNTQNFF
PASSVVAHSPDIFLTQDKLLAPKFATSQEPYRAQIRQQGEFESLHRDMIIGIKTWEFDPMDLENPFHNDDG
SVHIWQGDEDGLVPVILQRYIAQRLPWIQYHELKGSGHLFPYADGMGDKIMKTFLLGETFVS-

>35. Oxoglutarate-dependent dioxygenase
MASQDVKVPTIDFCTPELKPGTAQWESTKTQVFQALQEYGCFEAIYDKVPNEMQEALFDKLEETLPLEAK
LKEYLEKPHSRYEGQIPHLPFYEHMHAADLLLPESVETFANTFWRDGNSAFCNMVKSYTKTLMELDEMVK
RMVSESLGLQNYIDALLDPDFFLLRCTHYKATQGGDTNDSLGPHTDGSFLTIVKQNQAGLQVLKENGEWI
EFNVSPNSYIILSADSFMAWTNGRLHSPTHRVTMAGDTDRFSMQLFSFPKPGCTMETPKELVDGEHLLLF
KPYDIFGYYEYIMSGAKAGFGLKCYCGV-
```

```
>35. Oxoglutarate-dependent dioxygenase
MSLQQSHFAAAAVSPPSRAVVYDKHGPPDTVTRLTEVPPVEITKNDVCVKMLAAPINPADINKIEGVYPD
RTPLPIIGGGEGVGEIHSVGSSVKDLTPGDLVMHSAPYVQGTWQTYIVQDQSLWHKIDKSTPIEYAATVSV
NPVSALRMLDEFVTLKPGDSIVQNGATSVVGRCVIQLAKVRGIQIINIIRDNPGSDEVKEKLIKLGADKVFTE
SELDGKRVKSLLGDIPEPILGLNCIGGNAASLVIKFLK[ENKRVLFR]QGGTMVTYGGMSKEPISVSTSNFIFK
DLTLKGFWLQEWKLDQAKYRDWIDYLLALARTGKLKYE[YDLS]MEILPFDDFYTALEKAMGKQGSRGKQ
LLKF-
```

```
>38. Rolling stone
```

Here was a frame-shift, so black parts are the same amino acids in annotation and possible improved annotation. Yellow amino acids are based on the same DNA sequence, but read in another frame. Red amino acids are deleted in improved, whereas green amino acids are added in the improved version.

```
MAAGILIQWYDFICFAIIGIAMFVALYMLCCCKEHHPRYQTLTGQDDQNIDGTSPPMLAIQEIREKRHWR
ERRTIFGQIGTIISAYGCVLYYSNKAPYGNVINEEGSLPKDMKNSVCTRSEILNKAKGTMIKLQGYYCTSGGV
VPQENTKPPQDAQGVKKNRGTVADSRKQVNAASNSGTGNGQSQAHISKGSEWISPTKSRKSAVILTDIV
FWFIILPFSSNTHLRLNLLMGGMHSLNIILLLLDTSLNALVALPFSGTFNTVGSSVVLLFGAGSPALLWNLLV
ACDFENLNFL-
```

```
MAAGILIQWYDFICFAIIGIAMFVALYMLCCCKEHHPRYQTLTGQDDQNIDGTSPPMLYESLQGHPGDQ
GEAALERASYHLWTSCWRGIHPAWMLATRFLSFVVLCMFLTWDVLLYNCSIFVYYTEIGTIISAYGCVLYYS
NKAPYGNVINEEGSLPKDMKNSVCTRSEILNKAKGTMIKLQGQYHYEETVIQQRESFWNYLLQILYQTCA
GAVILTDIVFWFIILPFSSNTHLRLNLLMGGMHSLNIILLLLDTSLNALPFPWRTGLAYFMLWSCCYVIFQW
VIHACGFMWWPYPFLELSTPWAPLWYFCLALVHLPCYGIYWLLATLKISTFSRLFPQAFVKY-
```

```
>40. Chitin synthase
MDNHQFIYPILVSIVTATLGLVSFTLCIAAEFKKTKKKDLRFDGKLCYLPGSVAFELGIAALICLVIAQVIGNLFI
CRNFFSRNQPNNSKSKKKPSLNGFICGLS[W]CFGMAIILLGAASSMNKSQSLGEGWVDSECYIVKDGVYA
VSAFLVLVTMGSTLGPVILKIKKKNQVKQDGNIPSITS-
```

APPENDIX C. PRIMER SEQUENCES

Table 4. Primer sequences

Gene	Forward / reverse	Sequence
Phy3	Forward	TAAGGAACGTGTGGAGAGGG
	Reverse	TCCATTCATGTATCCAGAACCC
Phy4	Forward	TTGGAAGGAGCGGCAGGATT
	Reverse	GCCCACTCGTGTCTCCAAAAT
Rab18	Forward	TGTCACAAGGCGAGACACTT
	Reverse	TACAGCCCTTTCGCTATCCC
Wat1-related protein - middle	Forward 1 (used for overlap)	AAGGACACAATTACTATGGGAAG
	Reverse 1	TCACTTTCTGAGGAGTTGGTATT
	Forward 2	GCACTCATATAATCTTCGTGCT
	Reverse 2 (used for overlap)	TGGTGACGACATTGATGATAG
Wat1-related protein – 1st half	Forward	GCAGGGACATTGTTGTGTGT
	Reverse	CAGGAGGTGAAGATGGCGTAT
Wat1-related protein – 2nd half	Forward	TTCCCAGCACAACTAGGCAA
	Reverse	GGAGATAGCTTCTTGGGCCT

APPENDIX D. GEL ELECTROPHORESIS FOR 'WAT1-RELATED PROTEIN'

I wanted to check whether the 'Wat1-related mRNA' consisted of one or two proteins. To check this, the primers mentioned in Table 4 as 'used for overlap' were used; they are located in the last part of the first half and in the start of the second half. So if the primers are able to make a product in the cDNA, this means that the mRNA consists of both halves. If this does not happen, the mRNA does not contain both halves. To check if the primers worked properly, we also checked whether there was a band in the genomic DNA. The results are not very good, but as you can see in Figure 12, there is no band present in the cDNA, whereas there is a band present in the genomic DNA.

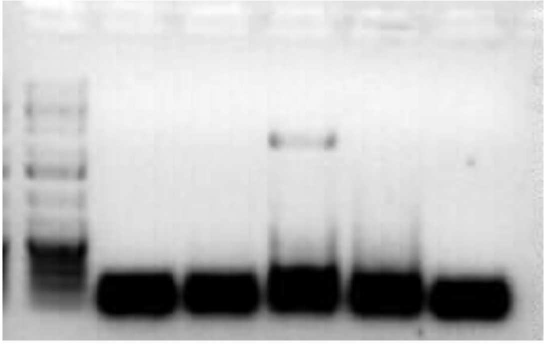


Table 12. Gel electrophoresis for 'Wat1-related protein'
The first band is the GeneRuler for control. The second band is the wild-type stage 4-5 AN1 and the third band is the wild-type stage 4-5 PH3. The third and fourth band are both genomic DNA; three has line D187, four has line D215. The last band is water.